# pantarei: three-way decomposition of variability in longitudinal, time-series-cross-sections, and multilevel data

Maria Elena Bontempi

Department of Economics, University of Bologna
mariaelena.bontempi@unibo.it

XVIII Italian Stata Conference, Milan, September 25, 2025

# The context I

Panel data (longitudinal and time-series-cross-sections) and multilevel data are widely used, the former in econometrics, the latter in *psychometrics* (psychology, sociology, education, criminology, biostatistics, etc., Bou and Satorra, 2017).

Different methodological approaches and aims emerge, reflecting the peculiarities of each field.

Different language, or different interpretation of the same language: fixed effects, random effects, multilevel models, mixed models.

We often deal with complex data and "social" structures, but these dataset can all be considered as grouped data:

- In econometrics, standard panel data have a structure with repeated observations over time nested in individuals, companies, countries.
- In psychometrics, multilevel models are used to analyse hierarchical data, such as *individual responses* (level-1 students or trees/plots) collected *within groups* (level-2 schools and background, species and ecological regions). Multilevel modelling also deals with cross-classifications or supergroups.
- Clustering and cross-classified effects tend to create dependencies between observations and imbalanced data structures: non-consecutive time periods or different number of units in each group.

# The context II

A preliminary and crucial step of any empirical research is to study the nature and relevance of the components that influence the variability of the variables, particularly the dependent variable.

The study of variability at various levels is interesting in itself and can provide iindications on the directions in which further investigation would be worthwhile.

Quantifying the unexplained variability with an unconditional (empty) model represents a preliminary step in assessing what we still do not understand about the differences in the values of the response variable between and within clusters.

Once explanatory variables have been introduced, we can obtain an indication of the progress make in explaining the variability of the response variable and the variability that remains unexplained.

Furthermore, quantifying the different types of variabilities in the response variable and the possible explanatory variables is a fundamental guide for the modelling and estimation strategies to be employed (Gibbons et al., 2014).

# The context III

- The presence of a significant clustering excludes the possibility of implementing POLS:
  $\longrightarrow$ The assumption of independent replications is violated (McNeish and Kelley, 2019; McNeish, 2023). Instead, the dependence of observations within clusters is of fundamental interest because it reflects essential aspects of the social world, such as the effect of attending the same school or living in the same area, working in the same company.
  $\longrightarrow$ Each nesting level is associated with a variability that has a distinct interpretation. For example, the growth of trees is influenced by their individual characteristics but also by the characteristics of the species and the ecological regions).
  $\longrightarrow$ Failure to distinguish between these different sources of variability would lead to incorrect conclusions due to logical errors such as:
  - The *ecological fallacy* (Robinson, 1950), which interprets associations at the higher level as pertaining to the lower level: conclusions about individuals are drawn on the basis of group-level data analysis, but the relationships observed at the group level are not necessarily valid at the individual level.
  - The *atomistic fallacy* (Alker, 1969), which instead interprets associations at the individual level as pertaining to the higher level.
  - In the "panel" jargon, the fallacy consists of confusing aggregate effects with individual effects, resulting from confounding the within-group (individuals) and between-group (higher level) effects.

# The context IV

- The absence of within-cluster variability excludes the possibility of implementing an FE estimation, particularly if the interest is on measurable higher-level explanatory variables, such as the role of education, race or gender on the wages of workers inside different companies.
- Considerably different between and within variabilities may suggest avoiding the weighted average implementd by RE/GLS and instead estimating the between and within effects of explanatory variables on the dependent variable separately, exploiting the Mundlak (1978)'s approach (also useful for testing endogeneity because of heterogeneity).

# The aim of the new command I

The new command pantarei has a name that derives from the ancient Greek expression πάντα ῥεῖ (Heraclitus of Ephesus, c. 535 – c. 475 BCE, pre-Socratic Greek philosopher?) to indicate that "everything flows", in the sense that heterogeneity and variations are central and even the methodologies for analysing panel data are constantly changing.
It breaks down the total variability into:

(1) The component due to level-2, individuals, cross-section or variability between clusters.

(2) The component due to level-1 observations nested in level-2, the variability within clusters.

The within variability is further decomposed into:

(2a) The residual variability once level-2 and level-1 heterogeneities have been taken into account, called within level-2-level-1 variability.

(2b) The variability due to the level-1 effect, called between level-1 variability.

# The aim of the new command II

- It is suitable for longitudinal, cross-sectional time-series and multilevel data.
- The standard panels investigated by the FE approach are the default ones; the `re` option exploits the GLS/RE approach; the `ml` option uses the ML and mixed model (hierarchical and with cross-effects).
- Two levels are considered, but it is possible to replicate the command on the same dataset by changing the levels considered, particularly in multilevel datasets.
- The joint tests on the statistical significance of the heterogeneities at the levels considered are reported.
- In the default FE setting, the levels are considered including dummy variables, and the F tests for the joint null hypothesis that all levels have the same mean are reported.
- In the GLS/RE case, levels are considered random and their variance parameters are tested using Breusch and Pagan's Lagrange multiplier tests; a mixed model is also implemented for level-1 effects which are estimated and tested as in the FE case.
- In the ML multilevel models, the levels are again considered random and tested using likelihood-ratio tests that compare the fitted models with and without random effects. A mixed model is also implemented for level-1 effects, which are estimated and tested as in FE case. In addition, a cross-effects specification is used whereby level-1 is not simply nested within level-2 clusters, but a single artificail level-3 super-clusteris assumed within which both level-2 and level-1 are nested (a trick from Rasbash and Goldstein, 1994; Goldstein, 1987).

# The aim of the new command III

- The dataset can be investigated with the `common` option, which computes the decomposition of variability for the subsample of non-missing observations of all variables.
- The default FE model has the `ur` option, which implements. for each variable, the Elliott et al. (1996) `dfgls` unit root test on the level-1 common factor that may drive all level-2 clusters. This is a proxy of the common correlated effect, CCE, which is both estimated from the residuals of each variable estimated against level-1 dummies and proxied by computing the level-1 averages (Pesaran, 2006; Brown et al., 2021). The longer the time span, the more relevant the stationarity feature of the common factor used to account for cross-sectional dependence becomes.
- A graph of the time series of the estimated and computed common factors for each variable is saved in the researcher's working directory as a file `UR_varname.gph`.
- The `ml` option predicts, from the cross-effects model, the empirical Bayes estimates (i.e. posterior, shrunken, or best linear unbiased predictions, BLUPs) of the effects together with their associated standard errors, and produces caterpillar plots of the effects, which are saved for each variable as files `pillaru03lev1_varname.gph` (level-1 effects) and `pillaru02lev2_varname.gph` (level-2 effects).

# The aim of the new command IV

- The new command is flexible and extends the existing commands `xtsum`, which only works for datasets that can be sorted with `tsset` or `xtset` and does not consider the role of CCE inside the within variability, and `loneway`, which allows only one level at a time.
- It also simplifies the use of post-regression results available after `mixed`. For example, a similar reason led to the creation of `pisamixed.ado`, contained in the zipped file PISA2022_Stata_PublicCodes and available for download from the website `https://www.oecd.org/en/data/datasets/pisa-2022-database.html#data`, which assists in the use of mixed calibrated on PISA data.
- The variance decomposition performed by the new command is consistent across different approaches to data, as it calculates the percentages of variability for each level as the ratio of partial sums of squares, rather than as VPCs and ICCs given by variance ratios (which are not adequately comparable due to the different degrees of freedom used in the calculation of variances).

# Where we would like to go I

A model for standard panel:

$$y_{it} = \alpha + \sum_{k=1}^{K} \beta_k x_{kit} + \left(\boldsymbol{\theta}'\mathbf{w}_i\right) + \mu_i + \tau_t + \varepsilon_{it}$$

In the econometric literature, FE and RE denote two type of estimators of $\boldsymbol{\beta}$ in a model where the unobserved random variable, $\mu_i$, varies across units. We select FE or RE depending on whether $\mu_i$ is allowed to correlate or not (at unit level) with the time varying variables $\mathbf{x}_{it}$.

Very often we have endogeneity because of heterogenenity, $\mathbb{E}(\mathbf{x}_{it}\mu_i) \neq 0$ and only FE provides unbiased estimates of $\boldsymbol{\beta}$. However, the unbiased within transformation cannot provide estimates of $\boldsymbol{\theta}$ for $\mathbf{w}_i$, the time-invariant/changing very little over time observables, when they are of prime theoretical interest.

The useful extension is Mundlak (1978) and Chamberlain (1980) discussed in the tribute to the live and work of Marc Nerlove (Baltagi and Mátyás, 2025) and Bellemare and Millimet (2025) is:

$$y_{it} = \boldsymbol{\beta}'\mathbf{x}_{it} + \boldsymbol{\theta}'\mathbf{w}_i + \tau_t + \boldsymbol{\pi}_x'\bar{\mathbf{x}}_{i.} + e_i + \varepsilon_{it}$$

where the "between" equation $\alpha + \mu_i = \alpha_i = \boldsymbol{\pi}_x'\bar{\mathbf{x}}_{i.} + e_i$ with $\bar{\mathbf{x}}_{i.} = T_i^{-1}\sum_{t=1}^{T_i}\mathbf{x}_{it}$, is used to captures the sample variation across units.

# Where we would like to go II

A multilevel hierarchical model (Goldstein, 2011; Bryk and Raudenbush, 2002; Hox and Roberts, 2011; Snijders and Bosker, 2012) with level-1 individuals $i = 1, \ldots, n_j$ nested within level-2 groups $j = 1, \ldots, J$ is:

$$
\begin{aligned}
y_{tij} &= \beta_{0j} + \sum_{k=1}^{K} \beta_{kj} x_{tijk} + \sum_{s=1}^{T} \gamma_s D_{st} + \varepsilon_{tij} \\
\beta_{0j} &= \beta_{00} + \boldsymbol{\beta}'_{01} \mathbf{w}_j + u_{0j} \\
\beta_{kj} &= \beta_{k0} + \boldsymbol{\beta}'_{k1} \mathbf{w}_j + u_{kj} \quad \text{for} k = 1, \ldots, K
\end{aligned}
$$

- $\beta_{00}$ is the grand mean of $y$ across all groups $j$; the mean of $y$ for group $j$ is $\beta_{00} + u_{0j}$ so that the distances between groups averages respect to the grand average $\beta_{00}$ represent the between-group variability.
- $\beta_{k0}$ is the grand mean slope across all groups, and $\beta_{k0} + u_{kj}$ is the role of group $j$ on the effect of $x_{tijk}$. The target is understanding how effects vary.
- Different types of effects incorporated at different levels: $u_{0j}$ and $u_{kj}$ are considered as random while $\gamma_s$ can be a fixed effect for $D_{st}$, the dummy for period $s$, which is the best choice when time is sparse and period effects are nuisance controls.
- $\beta_{kj}$ slope for variable $k$ in group $j$; $\mathbf{x}_{tij} = (x_{tij1}, \ldots, x_{tijK})'$ vector of unit- and time-varying covariates (usually centered).

# Where we would like to go III

- The $\mathbf{w}_j$ vector of time-invariant group-level covariates (usually centered) can be included in $\beta_{0j}$ and $\beta_{kj}$ so that $\boldsymbol{\beta}_{01}$ and $\boldsymbol{\beta}_{k1}$ are the vectors of contextual effects induced by $\mathbf{w}_j$ on intercept and slope $k$, respectively.
- Multilevel models can also be integrated with the Mundlak (1978)'s approach.

# Where do we start I

Unconditional "empty" models are:

FE/LSDV

$$y_{it} = \alpha + \underbrace{\sum_{j=1}^{N-1} \mu_j D_{ji}}_{\text{lev2 heterog.}} + \underbrace{\sum_{s=1}^{T-1} \tau_s \Theta_{st}}_{\text{lev1 heterog.}} + \varepsilon_{it}$$

The $\mu_i$ term is "unobserved heterogeneity" or unobserved "permanent/stable" effects specific to each level-2 cluster $i$ *included in the sample* and nesting level-1 $t$.

Adding a dummy variable for each unit (minus one baseline) is equivalent to within-transform the data.

Within-transformation prevents the estimation of $\mathbf{w}_i$ time-invariant variables.

Unobserved "unstable" factors are in the error term $\varepsilon_{it}$, with $\mathbb{E}(\varepsilon_{it}) = 0$ and $Var(\varepsilon_{it}) = \sigma^2_\varepsilon$.

Time dummies capture the role of CCE as a component of within-units variability, and it is equivalent to compute temporal averages.

RE/Random Intercept Model

$$y_{it} = \alpha + \sum_{s=1}^{T-1} \tau_s \Theta_{st} + \underbrace{\mu_i + \varepsilon_{it}}_{\nu_{it}}$$

The $\nu_{it}$ two-way error term includes "unobserved heterogeneity" which is condensed by variability $\sigma^2_\mu$. The two components have $\mathbb{E}(\varepsilon_{it}) = \mathbb{E}(\mu_i) = 0$, $Var(\varepsilon_{it}) = \sigma^2_\varepsilon$ and $Cov(\varepsilon_{it}, \mu_i) = 0$.

The $\mu_i$ is the result of a random draw from a distribution, with some units higher (lower) on $y$; each draw, added to the grand mean of $y$, $\alpha$, gives the observed intercept of unit $i$. The interest is *on the population* of clusters, beyond the specific clusters in the sample.

In this "random effects" framework, the $\mathbf{w}_i$ time-invariant variables can be estimated.

Time effects can be considered as fixed or random as well.

# Where do we start II

Two-level hierarchical model (Goldstein, 2011; Bryk and Raudenbush, 2002; Hox and Roberts, 2011; Snijders and Bosker, 2012) with level-1 units $i = 1, \ldots, n_j$ nested within level-2 groups $j = 1, \ldots, J$:

$$
\begin{aligned}
y_{ij} &= \beta_{0j} + \varepsilon_{ij} \\
\beta_{0j} &= \beta_{00} + u_{0j}
\end{aligned}
$$

The term $\beta_{00}$ is the grand mean of $y_{ij}$ across all groups $j$; the mean of $y$ for group $j$ is $\beta_{00} + u_{0j}$ so that the distances between groups averages respect to the grand average $\beta_{00}$ represent the between-group variability.
The distances between each observation $y_{ij}$ and the group mean represents the within-group between-individual variability.

Two-way cross-classified model (Goldstein, 2011; Rasbash and Goldstein, 1994; Rasbash and Browne, 2001; Rasbash and Browne, 2008; Beretvas, 2011) where level-1 units $i = 1, \ldots, n_{(j_1,j_2)}$ are clustered in more that one level-2 classification and the classification units are not nested within each other. The cross-classified factors are $j_1 = 1, \ldots, J_1$ and $j_2 = 1, \ldots, J_2$ at the same level-2; the pair or combination $(j_1, j_2)$ indexes the cells of the cross-classification:

$$
\begin{aligned}
y_{i(j_1,j_2)} &= \beta_{0(j_1,j_2)} + + \varepsilon_{i(j_1,j_2)} \\
\beta_{0(j_1,j_2)} &= \beta_{000} + u_{0j_1} + u_{0j_2} + u_{00j_1 \times j2}
\end{aligned}
$$

The term $\beta_{000}$ is the overall mean of $y_{i(j_1,j_2)}$ across all level-2 $j_1$, all level-2 $j_2$ and units $i$.

# Where do we start III

The random effects (with covariances typically assumed to be zero) are:

$u_{0j} \sim \mathcal{N}(0, \sigma^2_{u_0})$ with $\sigma^2_{u_0}$ between-group variance.

$\varepsilon_{tij} \sim \mathcal{N}(0, \sigma^2_{\varepsilon})$, with $\sigma^2_{\varepsilon}$ within-group between-units variance.

$u_{00j_1} \sim \mathcal{N}(0, \sigma^2_{u_{00j_1}})$, with $\sigma^2_{u_{00j_1}}$ between level-2 $j_1$ variance having adjusted for differences between level-2 $j_2$.

$u_{00j_2} \sim \mathcal{N}(0, \sigma^2_{u_{00j_2}})$, with $\sigma^2_{u_{00j_2}}$ between level-2 $j_2$ variance having adjusted for differences between level-2 $j_1$.

$\varepsilon_{i(j_1,j_2)} \sim \mathcal{N}(0, \sigma^2_{\varepsilon})$, with $\sigma^2_{\varepsilon}$ level-1 variance within the cells defined by level-2 $j_2$ $j_1$ pairings.

$u_{00j_1 \times j2} \sim \mathcal{N}(0, \sigma^2_{u_{00j_1 \times j2}})$ random interaction effect between the two cross-classified factors; without sufficiently large within-cell sample sizes, it is hard to separate the additive variances from the variance within the cells of the cross-classification and thus this last random effect is most typically set to zero.

# Where do we start IV

FE/RE models can be estimated with *ji* instead of *it*. Multilevel models can be estimated with *ti* instead of *ij*.

FE/RE models are estimated by inverting the indices in *ti*, and multilevel models can be estimated by inverting the indices in *ji* to capture the variability at each level and test the significance of the effect of each level.

In multilevel models we can use a mixed specification in which *i* of *ij* and *t* of *ti* are considered fixed effects captured by a set of dummy variables, i.e., by $\sum_{n=1}^{n_j} \gamma_{un} D_{ni}$ for units within each cluster *j* and $\sum_{s=1}^{T} \gamma_{ms} D_{st}$ for measurement occasions *t* within each cluster *i*, respectively.

Of course, there are a huge number of possible model specifications if each coefficient can be either fixed or random. This number increases with the number of clusters (level-1, level-2, level-3, ...). Strategies change if the panel structure is hierarchical or cross-classified.

The new `pantarei` command presents the possible specifications for two-level models and considers the possibility that there are few units within groups or that temporal observations are few and scattered, so that dummy variables are just controls for noise and panel imbalance in a mixed framework.

# The syntax I

pantarei list_of_variables [,options]

- options: the default is FE model; ur can be added to test for stationarity of CCE and obtain the plot of CCE. re implements the RE/GLS model; ml implements the multilevel and cross-classified models. common can be added to the default, re and ml to implement the models on the sub-sample of non missing values.
- list_of_variables must follow the order level-2, level-1, variables that we want to investigate.
- It is a good idea to generate level-2 and level-1 indices as progressive numbers (e.g. egen lev2=(group) level-2-index, label); this is not necessary for temporal level-1 index already in the form e.g. 2005, .., 2025 or 2020q1, .., 2025q4.
- The command can be executed under conditions by using the sequence preserve hence if or in condition hence pantarei hence restore.
- Time series operators (l. or d.) are not allowed, but lagged and first-differenced variables can be generated before implementing the pantarei command.

# A macro panel I

```
pantarei id year tint dlpc
```

```
-------------------------------- Level 2 (i): id   Level 1 (t): year       -----------------------
                                 Hypothetical full sample, N x T: 209
                                 N: 11    T: 19
-----------------------------------------------------------------------------------------------------
Variable                         |  Mean   Std.Dev.  % SS |  Observations   |  Test
---------------------------------+-------------------------+-----------------+-----------------------
tint      Overall xit-x..        | .09459  .03436         | N x T       209| id effects:
       1-Between xi.-x..         |         .02169   36.41 | N-min        11|  F(10, 180)   =  41.89
       2-Within xit-xi.          |         .02808   63.59 | N-bar     11.00|      Prob > F =   0.00
       ------ of which: ------   |        ----------------| N-max        11| year effects:
        2a-res xit-xi.-x.t+x..   |         .01461   15.64 | T-min        19|  F(18, 180)   =  30.65
        2b-between x.t-x..       |         .02439   47.95 | T-bar     19.00|      Prob > F =   0.00
         (2b in % of 2)          |                 (75.40)| T-max        19| Balancedness  = 100.00
-----------------------------------------------------------------------------------------------------
dlpc      Overall xit-x..        | .04837  .04363         | N x T       198| id effects:
       1-Between xi.-x..         |         .02583   32.02 | N-min        11|  F(10, 170)   =  26.01
       2-Within xit-xi.          |         .03693   67.98 | N-bar     11.00|      Prob > F =   0.00
       ------ of which: ------   |        ----------------| N-max        11| year effects:
        2a-res xit-xi.-x.t+x..   |         .02149   20.93 | T-min        18|  F(17, 170)   =  22.49
        2b-between x.t-x..       |         .03072   47.06 | T-bar     18.00|      Prob > F =   0.00
         (2b in % of 2)          |                 (69.22)| T-max        18| Balancedness  =  94.74
-----------------------------------------------------------------------------------------------------
```

# A macro panel II

pantarei id year tint dlpc, common ur

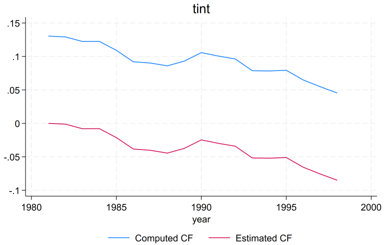
```
-------------------------------- Level 2 (i): id   Level 1 (t): year       -----------------------
                                 Hypothetical full sample, N x T: 209
                                 N: 11    T: 19
-----------------------------------------------------------------------------------------------
Variable                        |  Mean   Std.Dev.  % SS |  Observations   |  Test
--------------------------------+------------------------+----------------+-----------------------
tint     Overall xit-x..        | .09329   .03421        | N x T       198| id effects:
       1-Between xi.-x..        |          .02137   35.66 | N-min        11|  F(10, 170)    =  37.83
       2-Within xit-xi.         |          .02816   64.34 | N-bar     11.00|      Prob > F =   0.00
       ------ of which: ------  |         ----------------| N-max        11| year effects:
        2a-res xit-xi.-x.t+x..  |          .01474   16.03 | T-min        18|  F(17, 170)    =  30.15
        2b-between x.t-x.. (CF) |           .0244   48.32 | T-bar        18|      Prob > F =   0.00
         (2b in % of 2)         |                 (75.09)| T-max     18.00| Balancedness   =  94.74
         Estimated CF x.t       |                        |                | H0: UR - pval =   0.47
         Computed  CF x.t       |                        |                | H0: UR - pval =   0.47
-----------------------------------------------------------------------------------------------
dlpc     Overall xit-x..        | .04837   .04363        | N x T       198| id effects:
       1-Between xi.-x..        |          .02583   32.02 | N-min        11|  F(10, 170)    =  26.01
       2-Within xit-xi.         |          .03693   67.98 | N-bar     11.00|      Prob > F =   0.00
       ------ of which: ------  |         ----------------| N-max        11| year effects:
        2a-res xit-xi.-x.t+x..  |          .02149   20.93 | T-min        18|  F(17, 170)    =  22.49
        2b-between x.t-x.. (CF) |          .03072   47.06 | T-bar        18|      Prob > F =   0.00
         (2b in % of 2)         |                 (69.22)| T-max     18.00| Balancedness   =  94.74
         Estimated CF x.t       |                        |                | H0: UR - pval =   0.08
         Computed  CF x.t       |                        |                | H0: UR - pval =   0.08
-----------------------------------------------------------------------------------------------
```

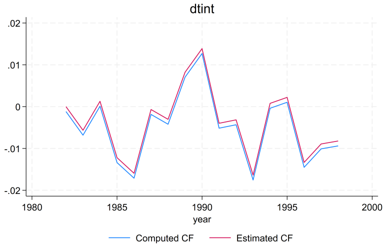
# A macro panel III

```
g dtint=d.tint
g ddlpc=d.dlpc
pantarei id year dtint ddlpc, common ur
```

```
-------------------------------- Level 2 (i): id   Level 1 (t): year       -----------------------
                                 Hypothetical full sample, N x T: 209
                                 N: 11    T: 19
---------------------------------------------------------------------------------------------------
Variable                         |  Mean   Std.Dev.  % SS |  Observations   |  Test
---------------------------------+------------------------+----------------+-----------------------
dtint     Overall xit-x..        | -.005   .01157         | N x T       187| id effects:
       1-Between xi.-x..         |        .001859    2.36 | N-min        11|  F(10, 160)    =   0.76
       2-Within xit-xi.          |         .01175   97.64 | N-bar     11.00|      Prob > F =   0.67
       ------ of which: ------   |        ----------------| N-max        11| year effects:
        2a-res xit-xi.-x.t+x..   |        .008803   49.83 | T-min        17|  F(16, 160)    =   9.59
        2b-between x.t-x.. (CF)  |        .008222   47.81 | T-bar        17|      Prob > F =   0.00
         (2b in % of 2)          |                 (48.97)| T-max     17.00| Balancedness   =  89.47
         Estimated CF x.t        |                        |                | H0: UR - pval =   0.01
         Computed  CF x.t        |                        |                | H0: UR - pval =   0.01
---------------------------------------------------------------------------------------------------
ddlpc     Overall xit-x..        |-.00584  .01455         | N x T       187| id effects:
       1-Between xi.-x..         |        .002577    2.87 | N-min        11|  F(10, 160)    =   0.82
       2-Within xit-xi.          |         .01474   97.13 | N-bar     11.00|      Prob > F =   0.61
       ------ of which: ------   |        ----------------| N-max        11| year effects:
        2a-res xit-xi.-x.t+x..   |         .01175   56.10 | T-min        17|  F(16, 160)    =   7.31
        2b-between x.t-x.. (CF)  |        .009583   41.03 | T-bar        17|      Prob > F =   0.00
         (2b in % of 2)          |                 (42.24)| T-max     17.00| Balancedness   =  89.47
         Estimated CF x.t        |                        |                | H0: UR - pval =   0.03
         Computed  CF x.t        |                        |                | H0: UR - pval =   0.03
---------------------------------------------------------------------------------------------------
```

# A macro panel IV




CCE for the levels of the interest rate

CCE for the first differences of the interest rate

# A macro panel V

```
pantarei id year tint dlpc, common re
```

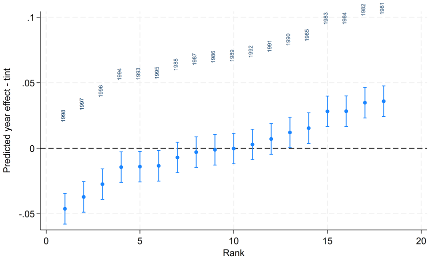
```
--------------------------------- Level 2 (j): id   Level 1 (i): year       -----------------------
                                  Hypothetical full sample, Ni x Nj: 209
                                  Nj: 11   Ni: 19
----------------------------------------------------------------------------------------------------
Variable                         |  Mean   Std.Dev.  % SS |  Observations   |  Test
---------------------------------+-------------------------+----------------+-----------------------
tint      Overall xij-x..        | .09329   .03421         | Ni x Nj     198| id effects:
       1-Between xj-x..          |          .02278   40.54 | Nj-min       11|  Chi2(1)       = 170.98
       2-Within xij-xj           |          .02707   59.46 | Nj-bar    11.00|   Prob > Chi2 =   0.00
       ------ of which: ------   |         ----------------| Nj-max       11| year effects:
        2a-res xij-xj-xi+x..     |          .01474   16.03 | Ni-min       18|  Chi2(1)       = 184.31
        2b-between xi-x..        |          .02314   43.44 | Ni-bar    18.00|   Prob > Chi2 =   0.00
        (2b in % of 2)           |                  (73.05)| Ni-max       18| Balancedness   =  94.74
----------------------------------------------------------------------------------------------------
   -- year as fixed effects: --  |                         |                |
       3-Between xj-x..          |          .02137   35.66 |                | id effects:
       4-Within xij-xj           |          .02816   64.34 |                |  Chi2(1)       = 759.32
       ------ of which: ------   |         ----------------|                |   Prob > Chi2 =   0.00
       4a-res xij-xj-xi+x..      |          .01474   16.03 |                | year effects:
       4b-between xi-x..         |           .0244   48.32 |                |  Chi2(17)      = 512.54
        (4b in % of 4)           |                  (75.09)|                |   Prob > Chi2 =   0.00
----------------------------------------------------------------------------------------------------
```

# A macro panel VI

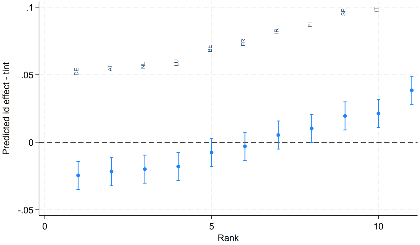
```
pantarei id year tint dlpc, common ml
```

```
-------------------------------- Level 2 (j): id   Level 1 (i): year       -----------------------
                                 Hypothetical full sample, Ni x Nj: 209
                                 Nj: 11   Ni: 19
--------------------------------------------------------------------------------------------------
Variable                         |  Mean   Std.Dev.  % SS |  Observations   |  Test
---------------------------------+-------------------------+-----------------+------------------------
tint     Overall xij-x..         | .09329   .03421          | Ni x Nj     198| id effects:
       1-Between xj-x..          |          .02392   44.68 | Nj-min       11|  Chi2(1)      =  51.32
       2-Within xij-xj           |          .02611   55.32 | Nj-bar    11.00|   Prob > Chi2 =   0.00
       ------ of which: ------   |         ----------------| Nj-max       11| year effects:
        2a-res xij-xj-xi+x..     |          .01405   14.57 | Ni-min       18|  Chi2(1)      =  71.58
        2b-between xi-x..        |          .02241   40.75 | Ni-bar    18.00|   Prob > Chi2 =   0.00
        (2b in % of 2)           |                  (73.66)| Ni-max       18| Balancedness  =  94.74
--------------------------------------------------------------------------------------------------
   -- year as fixed effects: --  |                         |                 |
       3-Between xj-x..          |          .02137   35.66 |                 | id effects:
       4-Within xij-xj           |          .02816   64.34 |                 |  Chi2(1)      = 180.57
       ------ of which: ------   |         ----------------|                 |   Prob > Chi2 =   0.00
       4a-res xij-xj-xi+x..      |          .01405   14.57 |                 | year effects:
       4b-between xi-x..         |          .02477   49.77 |                 |  Chi2(17)     = 563.80
        (4b in % of 4)           |                  (77.36)|                 |   Prob > Chi2 =   0.00
--------------------------------------------------------------------------------------------------
   -- year as crossed-effect: -- |                         |                 |
       3-Between xj-x..          |           .0205   34.81 |                 | id effects:
       4-Within xij-xj           |          .02753   65.19 |                 |  Chi2(1)      = 163.35
       ------ of which: ------   |         ----------------|                 |   Prob > Chi2 =   0.00
       4a-res xij-xj-xi+x..      |          .01474   16.99 |                 | year effects:
       4b-between xi-x..         |          .02367   48.19 |                 |  Chi2(1)      =  71.58
        (4b in % of 4)           |                  (73.93)|                 |   Prob > Chi2 =   0.00
--------------------------------------------------------------------------------------------------
```

# A macro panel VII




Level-3 (cross-effects) pillar-plot for the interest rate




Level-2 pillar-plot for the interest rate

# One of the most widely used longitudinal panels I

In abdata (Arellano and Bond, 1991; Blundell and Bond, 1998) there are no missing values, so we can skip the common option.

`pantarei id year n w k`

```
--------------------------------- Level 2 (i): id   Level 1 (t): year        -----------------------
                                  Hypothetical full sample, N x T: 1260
                                  N: 140   T: 9
----------------------------------------------------------------------------------------------------
Variable                         |  Mean   Std.Dev.  % SS |  Observations   |  Test
---------------------------------+------------------------+-----------------+-----------------------
n        Overall xit-x..         | 1.056    1.342         | N x T      1031| id effects:
       1-Between xi.-x..         |          1.331    97.90 | N-min        35|  F(139, 883)  = 428.33
       2-Within xit-xi.          |          .2092     2.10 | N-bar    114.56|      Prob > F =   0.00
       ------ of which: ------   |        ----------------| N-max       140| year effects:
        2a-res xit-xi.-x.t+x..   |          .1731     1.43 | T-min         7|  F(8, 883)     =  52.24
        2b-between x.t-x..       |          .1169     0.68 | T-bar      7.36|      Prob > F =   0.00
         (2b in % of 2)          |                 (32.12)| T-max         9| Balancedness  =  81.83
----------------------------------------------------------------------------------------------------
w        Overall xit-x..         | 3.143     .263         | N x T      1031| id effects:
       1-Between xi.-x..         |          .2503    90.01 | N-min        35|  F(139, 883)  =  75.38
       2-Within xit-xi.          |         .08939     9.99 | N-bar    114.56|      Prob > F =   0.00
       ------ of which: ------   |        ----------------| N-max       140| year effects:
        2a-res xit-xi.-x.t+x..   |         .07837     7.61 | T-min         7|  F(8, 883)     =  34.51
        2b-between x.t-x..       |         .04302     2.38 | T-bar      7.36|      Prob > F =   0.00
         (2b in % of 2)          |                 (23.82)| T-max         9| Balancedness  =  81.83
----------------------------------------------------------------------------------------------------
k        Overall xit-x..         | -.4416   1.514         | N x T      1031| id effects:
       1-Between xi.-x..         |          1.503    97.97 | N-min        35|  F(139, 883)  = 396.39
       2-Within xit-xi.          |          .2322     2.03 | N-bar    114.56|      Prob > F =   0.00
       ------ of which: ------   |        ----------------| N-max       140| year effects:
        2a-res xit-xi.-x.t+x..   |          .2037     1.55 | T-min         7|  F(8, 883)     =  34.39
        2b-between x.t-x..       |          .1116     0.48 | T-bar      7.36|      Prob > F =   0.00
         (2b in % of 2)          |                 (23.76)| T-max         9| Balancedness  =  81.83
----------------------------------------------------------------------------------------------------
```

# One of the most widely used longitudinal panels II

Option ur gives the error message

`The lenght of time series could be not enough to compute UR test (at least 9 periods are required) Error code r(127) (or r(2000) or r(498)) maxlag() invalid - invalid number, missing not allowed`

Anyway, we can produce the pillar-plots with the option `ml`

```
pantarei id year n w k, ml
```

# One of the most widely used longitudinal panels III

```
-------------------------------- Level 2 (j): id   Level 1 (i): year      ----------------------
                                 Hypothetical full sample, Ni x Nj: 1260
                                 Nj: 140   Ni: 9
------------------------------------------------------------------------------------------------
Variable                        |   Mean   Std.Dev.  % SS |  Observations   |  Test
--------------------------------+-------------------------+-----------------+-------------------
n        Overall xij-x..        |  1.056    1.342         | Ni x Nj    1031| id effects:
       1-Between xj-x..         |           1.327   97.20 | Nj-min       35|  Chi2(1)      =3032.80
       2-Within xij-xj          |           .2412    2.80 | Nj-bar   114.56|   Prob > Chi2 =   0.00
       ------ of which: ------  |         ----------------| Nj-max      140| year effects:
        2a-res xij-xj-xi+x..    |           .1724    1.42 | Ni-min        7|  Chi2(1)      =   2.50
        2b-between xi-x..       |           .1671    1.38 | Ni-bar     7.36|   Prob > Chi2 =   0.03
        (2b in % of 2)          |                 (49.38)| Ni-max        9| Balancedness  =  81.83
------------------------------------------------------------------------------------------------
   -- year as fixed effects: -- |                         |                 |
       3-Between xj-x..         |           1.331   97.90 |                 | id effects:
       4-Within xij-xj          |           .2092    2.10 |                 |  Chi2(1)      =3355.32
       ------ of which: ------  |         ----------------|                 |   Prob > Chi2 =   0.00
       4a-res xij-xj-xi+x..     |           .1724    1.42 |                 | year effects:
       4b-between xi-x..        |            .118    0.69 |                 |  Chi2(8)      = 422.78
        (4b in % of 4)          |                 (32.73)|                 |   Prob > Chi2 =   0.00
------------------------------------------------------------------------------------------------
   -- year as crossed-effect: -- |                        |                 |
       3-Between xj-x..         |           1.329   97.72 |                 | id effects:
       4-Within xij-xj          |           .2175    2.28 |                 |  Chi2(1)      =3335.63
       ------ of which: ------  |         ----------------|                 |   Prob > Chi2 =   0.00
       4a-res xij-xj-xi+x..     |           .1731    1.43 |                 | year effects:
       4b-between xi-x..        |           .1308    0.85 |                 |  Chi2(1)      =   2.50
        (4b in % of 4)          |                 (37.20)|                 |   Prob > Chi2 =   0.03
------------------------------------------------------------------------------------------------
```

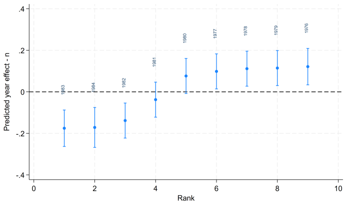
# One of the most widely used longitudinal panels IV

```
-----------------------------------------------------------------------------------------------------
w        Overall xij-x..        |  3.143     .263          | Ni x Nj     1031| id effects:
       1-Between xj-x..         |             .2521   91.33 | Nj-min        35|  Chi2(1)       =1665.09
       2-Within xij-xj          |            .08329    8.67 | Nj-bar    114.56|   Prob > Chi2 =    0.00
       ------ of which: ------  |          ----------------| Nj-max       140| year effects:
        2a-res xij-xj-xi+x..    |            .07801    7.54 | Ni-min         7|  Chi2(1)       =    4.25
        2b-between xi-x..       |            .02966    1.13 | Ni-bar      7.36|   Prob > Chi2 =    0.01
        (2b in % of 2)          |                   (13.05)| Ni-max         9| Balancedness  =   81.83
-----------------------------------------------------------------------------------------------------
   -- year as fixed effects: -- |                          |                 |
       3-Between xj-x..         |             .2503   90.01 |                 | id effects:
       4-Within xij-xj          |            .08938    9.99 |                 |  Chi2(1)       =1883.55
       ------ of which: ------  |          ----------------|                 |   Prob > Chi2 =    0.00
       4a-res xij-xj-xi+x..     |            .07801    7.54 |                 | year effects:
       4b-between xi-x..        |            .04362    2.45 |                 |  Chi2(8)       = 276.27
        (4b in % of 4)          |                   (24.50)|                 |   Prob > Chi2 =    0.00
-----------------------------------------------------------------------------------------------------
   -- year as crossed-effect: -- |                          |                 |
       3-Between xj-x..         |             .2438   89.12 |                 | id effects:
       4-Within xij-xj          |            .09128   10.88 |                 |  Chi2(1)       =1863.04
       ------ of which: ------  |          ----------------|                 |   Prob > Chi2 =    0.00
       4a-res xij-xj-xi+x..     |            .07838    7.95 |                 | year effects:
       4b-between xi-x..        |            .04672    2.93 |                 |  Chi2(1)       =    4.25
        (4b in % of 4)          |                   (26.94)|                 |   Prob > Chi2 =    0.01
-----------------------------------------------------------------------------------------------------
```

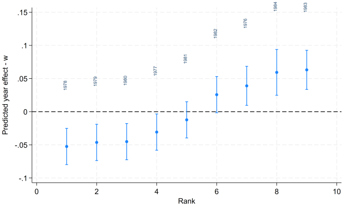
# One of the most widely used longitudinal panels V

```
-------------------------------------------------------------------------------------------------
k        Overall xij-x..         | -.4416    1.514         | Ni x Nj    1031| id effects:
      1-Between xj-x..           |           1.505   98.14 | Nj-min       35|  Chi2(1)      =3063.15
      2-Within xij-xj            |           .2217    1.86 | Nj-bar   114.56|   Prob > Chi2 =    0.00
      ------ of which: ------    |         ----------------| Nj-max      140| year effects:
       2a-res xij-xj-xi+x..      |           .2028    1.54 | Ni-min        7|  Chi2(1)      =    0.23
       2b-between xi-x..         |           .0905    0.32 | Ni-bar     7.36|   Prob > Chi2 =    0.16
       (2b in % of 2)            |                 (17.14)| Ni-max        9| Balancedness  =   81.83
-------------------------------------------------------------------------------------------------
   -- year as fixed effects: --  |                         |                |
      3-Between xj-x..           |           1.503   97.97 |                | id effects:
      4-Within xij-xj            |           .2322    2.03 |                |  Chi2(1)      =3288.44
      ------ of which: ------    |         ----------------|                |   Prob > Chi2 =    0.00
      4a-res xij-xj-xi+x..       |           .2028    1.54 |                | year effects:
      4b-between xi-x..          |           .1132    0.50 |                |  Chi2(8)      =  278.52
       (4b in % of 4)            |                 (24.44)|                |   Prob > Chi2 =    0.00
-------------------------------------------------------------------------------------------------
   -- year as crossed-effect: -- |                         |                |
      3-Between xj-x..           |           1.498   97.78 |                | id effects:
      4-Within xij-xj            |           .2422    2.22 |                |  Chi2(1)      =3267.04
      ------ of which: ------    |         ----------------|                |   Prob > Chi2 =    0.00
      4a-res xij-xj-xi+x..       |           .2037    1.56 |                | year effects:
      4b-between xi-x..          |           .1306    0.67 |                |  Chi2(1)      =    0.23
       (4b in % of 4)            |                 (29.91)|                |   Prob > Chi2 =    0.16
-------------------------------------------------------------------------------------------------
```

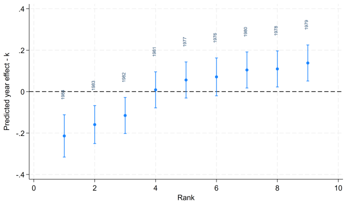
# One of the most widely used longitudinal panels VI



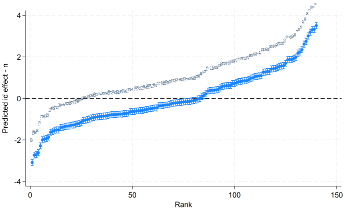

Level-3 (cross-effects) pillar-plot for log(employees)




Level-3 (crossed-effects) pillar-plot for log(wage)




Level-3 (cross-effects) pillar-plot for log(capital)




Level-2 pillar-plot for log(employees)

# Considerations I

- Cross-classified multilevel models are applicable to macro panels, although the latter are usually considered hierarchical structures with periods nested within individuals in a balanced manner (all units measured over the same time span and all starting in the same period).
- In fact, it may be appropriate to consider a cross-classification of measurement occasions treated as random in order to capture events that can affect all units. This is an alternative way of considering CCEs induced by shocks to the international business cycle and different capacities to adapt to joint monetary policy choices (such as euro adoption).
- In a balanced macro long $T$ panel, fixed time effects or cross-time effects are relevant and capable of capturing a broken trend that tends to be non-stationary for the interest rate (an I(1) process).
- Also, cross-classified models allow us to assess the effects of panel imbalance, in the form of panel dropout or occasional non-measurement, as in `abdata`.
- Multilevel models show that in this longitudianl short $T$ unbalanced panel, time considered as random is more relevant, particularly for variable wage $w$, capturing entries and exits of units from the sample as well as the role of inflation.
- Although the UR test requires at least 9 periods and cannot be implement on `abdata`, pillar plots allow for some considerations, such as the tendency of employees $n$ to decrease when wages $w$ increase in 1983-1984, the decline in wages $w$ as a consequence of the second oil shock in 1979-1980; a stable decrease of capital stock $k$ over the period.

# Very unbalanced micro panel I

Data from de Jong et al. (2011)

```
egen id=group(gvkey), label
tsset id yeara
Panel variable: id (unbalanced)
 Time variable: yeara, 1986 to 2001, but with gaps
        Delta: 1 unit
xtdes
      id:  1, 2, ..., 5449                                  n =       5449
   yeara:  1986, 1987, ..., 2001                            T =         16
           Delta(yeara) = 1 unit
           Span(yeara)  = 16 periods
           (id*yeara uniquely identifies each observation)

Distribution of T_i:   min      5%     25%       50%       75%     95%     max
                         1       1       1         3         8      14      16

     Freq.  Percent    Cum. |  Pattern
 ---------------------------+------------------
      384      7.05    7.05 |  ...............1
      176      3.23   10.28 |  1111111111111111
      167      3.06   13.34 |  1...............
      149      2.73   16.08 |  ..............1.
      136      2.50   18.57 |  11..............
      132      2.42   20.99 |  ............1...
      131      2.40   23.40 |  .............1..
      113      2.07   25.47 |  111.............
       98      1.80   27.27 |  ...........1....
     3963     72.73  100.00 | (other patterns)
 ---------------------------+------------------
     5449    100.00         |  XXXXXXXXXXXXXXXX
```

# Very unbalanced micro panel II

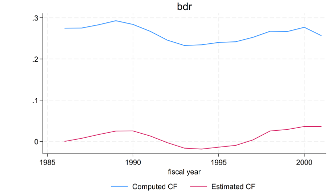
pantarei id yeara bdr mdr, common ur

```
-------------------------------- Level 2 (i): id   Level 1 (t): yeara      ----------------------
                                 Hypothetical full sample, N x T: 87184
                                 N: 5449   T: 16
-----------------------------------------------------------------------------------------------
Variable                         |   Mean   Std.Dev.  % SS |  Observations   |  Test
---------------------------------+-------------------------+-----------------+-------------------
The lenght of time series could be not enough to compute UR test
(at least 9 periods are required)
Error code r(127) (or r(2000) or r(498))
bdr      Overall xit-x..         |  .2608    .2139         | N x T     27762| id effects:
       1-Between xi.-x..         |           .1826   72.90 | N-min      1528|  F(5448, 22298)=  11.21
       2-Within xit-xi.          |           .1242   27.10 | N-bar   1735.12|     Prob > F =   0.00
       ------ of which: ------   |         ----------------| N-max      2056| yeara effects:
        2a-res xit-xi.-x.t+x..   |            .123   26.56 | T-min         1|  F(15, 22298) =  30.50
        2b-between x.t-x.. (CF)  |          .01631    0.54 | T-bar         5|     Prob > F =   0.00
         (2b in % of 2)          |                  (2.01) | T-max     16.00| Balancedness  =  31.84
         Estimated CF x.t        |                         |                 | H0: UR - pval =   0.00
         Computed  CF x.t        |                         |                 | H0: UR - pval =   0.04
-----------------------------------------------------------------------------------------------
The lenght of time series could be not enough to compute UR test
(at least 9 periods are required)
Error code r(127) (or r(2000) or r(498))
mdr      Overall xit-x..         |  .2689    .2459         | N x T     27762| id effects:
       1-Between xi.-x..         |           .2086   71.93 | N-min      1528|  F(5448, 22298)=  11.12
       2-Within xit-xi.          |           .1453   28.07 | N-bar   1735.12|     Prob > F =   0.00
       ------ of which: ------   |         ----------------| N-max      2056| yeara effects:
        2a-res xit-xi.-x.t+x..   |            .141   26.40 | T-min         1|  F(15, 22298) =  93.99
        2b-between x.t-x.. (CF)  |          .03282    1.67 | T-bar         5|     Prob > F =   0.00
         (2b in % of 2)          |                  (5.95) | T-max     16.00| Balancedness  =  31.84
         Estimated CF x.t        |                         |                 | H0: UR - pval =   0.13
         Computed  CF x.t        |                         |                 | H0: UR - pval =   0.09
-----------------------------------------------------------------------------------------------
```

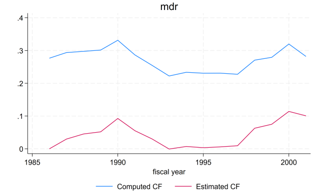
# Very unbalanced micro panel III

```
pantarei id yeara bdr mdr, common ml
```

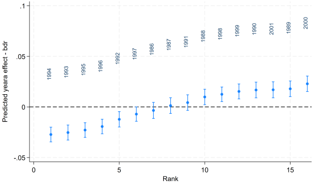
# Very unbalanced micro panel IV



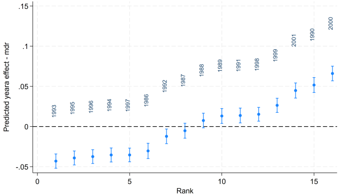

CCE for debt-ratio, book value




CCE for debt-ratio, market value




Level-3 (crossed-effects) pillar-plot for debt-ratio, book value




Level-3 (crossed-effects) pillar-plot for debt-ratio, market value)

# Very unbalanced micro panel V

The ur option provides a warming when the panel is unbalanced, but the test is implemented for those companies with more than 9 periods.

The debt-to-market value ratio is non-stationary.

Some breaks occured: the oil shock (Iraq invaded Kuwait, causing an increase in the price of crude oil, which sparked fears of inflation) and the credit crisis (bankruptcy of hundreds of US credit insititutions, more restrictive lending by banks, restrictive policy by the FED to curb inflation) in 1990, which caused the S&P 500 to fall by about 14% from July to October 1990.

In 2000 the dot-com bubble burst, with technology stocks being overvalued (the Nasdaq grew by almost 400% between 1995 and March 2000, driven largely by unprofitable internet companies), between June 1999 and May 2000, the FED raised the federal funds rate from 4.75% to 6.5% to cool the overheated economy, and companies overinvesting in IT systems before 2000, only to slow down their investments afterward. This caused the Nasdaq Composite to fall by around 50% and the S&P 500 to fall by 9%.

# Quarterly data & micro panel I

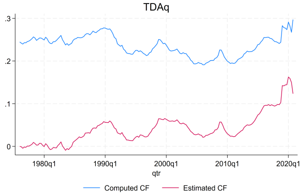
The pantarei command also works, e.g., with quarterly data. An example from COMPUSTAT data over the 1975q1-2020q4 (Bontempi and Bottazzi, 2025), again for debt ratios at book and market values.

```
pantarei id qtr TDAq TDMq, ur
```

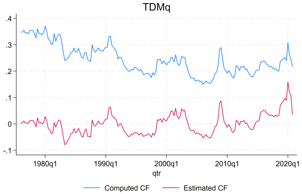
```
-------------------------------- Level 2 (i): id   Level 1 (t): qtr       -----------------------
                                 Hypothetical full sample, N x T: 2281140
                                 N: 12673   T: 180
----------------------------------------------------------------------------------------------------
Variable                         |   Mean   Std.Dev.  % SS |  Observations   |  Test
---------------------------------+-------------------------+-----------------+----------------------
TDAq      Overall xit-x..        |  .2327     .211         | N x T     436256| id effects:
       1-Between xi.-x..         |          .1681    63.48 | N-min        483|  F(12179, 423897)=  62.96
       2-Within xit-xi.          |          .1293    36.52 | N-bar    2423.64|      Prob > F =    0.00
       ------ of which: ------   |        ----------------| N-max       3812| qtr effects:
        2a-res xit-xi.-x.t+x..   |          .1269    35.14 | T-min          1|  F(179, 423897)=  93.59
        2b-between x.t-x.. (CF)  |         .02493     1.39 | T-bar         36|      Prob > F =    0.00
         (2b in % of 2)          |                  (3.80) | T-max     180.00| Balancedness  =  19.12
         Estimated CF x.t        |                         |                 | H0: UR - pval =    0.81
         Computed  CF x.t        |                         |                 | H0: UR - pval =    0.22
----------------------------------------------------------------------------------------------------
TDMq      Overall xit-x..        |  .2314    .2396         | N x T     424603| id effects:
       1-Between xi.-x..         |          .1933    65.07 | N-min        479|  F(12146, 412277)=  65.32
       2-Within xit-xi.          |          .1437    34.93 | N-bar    2358.91|      Prob > F =    0.00
       ------ of which: ------   |        ----------------| N-max       3779| qtr effects:
        2a-res xit-xi.-x.t+x..   |          .1393    32.81 | T-min          1|  F(179, 412277)= 149.00
        2b-between x.t-x.. (CF)  |           .035     2.12 | T-bar         35|      Prob > F =    0.00
         (2b in % of 2)          |                  (6.08) | T-max     180.00| Balancedness  =  18.61
         Estimated CF x.t        |                         |                 | H0: UR - pval =    0.00
         Computed  CF x.t        |                         |                 | H0: UR - pval =    0.28
----------------------------------------------------------------------------------------------------
```

```
pantarei id qtr TDAq TDMq, common ml
```

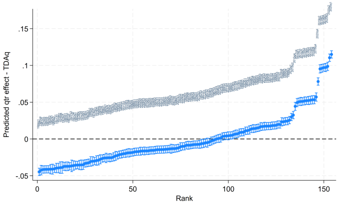
# Quarterly data & micro panel II



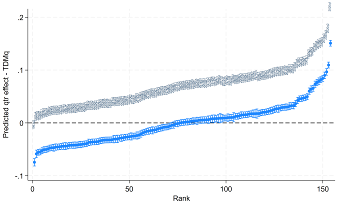

CCE for debt-ratio, book value quarterly




CCE for debt-ratio, market value quarterly




Level-3 (cross-effects) pillar-plot for debt-ratio, book value quarterly




Level-3 (cross-effects) pillar-plot for debt-ratio, market value quarterly)

# Another preliminar analysis I

Data taken from Bontempi and Kean (2025) Do efficient institutions discourage foreign direct investment (FDI) in developing countries rich in natural resources?

One of the main findings is that the emerging of democracy stimulates FDI regardless of the level of natural resources, provided that the host countries are already governed democratically to some extent.

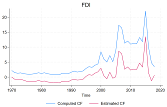
Conversely, countries characterised by the risk of expropriation or long-standing governments (often associated with more autocratic regimes) do not discourage FDI aimed at exploiting natural resources.

# Another preliminar analysis II

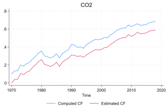
| | CO2 | FDI | DEM | RISK | GOV | FE | ME |
|---|---|---|---|---|---|---|---|
| **between** $i$ | 92.69 | 58.39 | 65.37 | 44.86 | 18.51 | 81.75 | 79.00 |
| **between** $t$ | 0.84 | 0.44 | 9.15 | 28.66 | 43.86 | 0.59 | 0.84 |
| **residual within** | 6.47 | 41.17 | 25.48 | 26.49 | 37.63 | 17.55 | 20.16 |
| **UR** | yes | yes | yes | yes | no | no | yes |

Table: Managing `pantarei` results

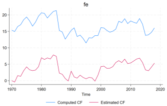
# Another preliminar analysis III



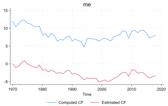

FDI



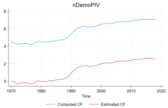

CO2 emissions



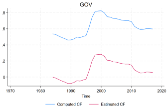

Fuels



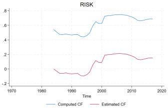

Minerals




Democracy

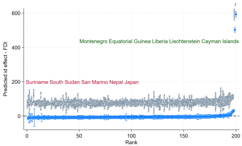



Gov. stability

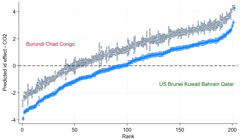



Risk of expropriation

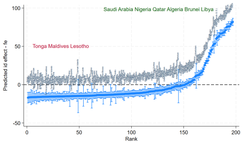
# Another preliminar analysis IV



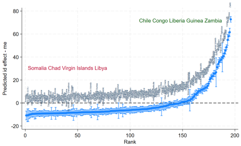

FDI



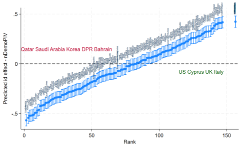

CO2 emissions



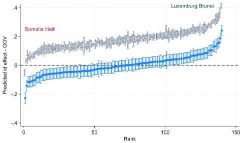

Fuels



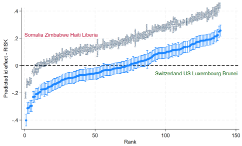

Minerals




Democracy




Gov. stability




Risk of expropriation

# A classical multilevel dataset I

Data taken from Abrevaya (2006) studying the effect fo smoking on the weight (in grams) of single newborns. Smoking status was determined by the answer to the question on the birth certificate regarding tobacco consumption during pregnancy. Smoking may vary from one pregnancy to another.

```
egen mom=group(momid)
```

```
pantarei mom idx birwt cigs smoke, common ml
```

# A classical multilevel dataset II

```
--------------------------------- Level 2 (j): mom   Level 1 (i): idx       -----------------------
                                  Hypothetical full sample, Ni x Nj: 11934
                                  Nj: 3978   Ni: 3
----------------------------------------------------------------------------------------------------
Variable                         |   Mean   Std.Dev.  % SS |  Observations   |  Test
---------------------------------+--------------------------+-----------------+-----------------------
birwt     Overall xij-x..        |   3470    527.1         | Ni x Nj     8604| mom effects:
       1-Between xj-x..          |           448.4   72.35 | Nj-min       648|  Chi2(1)       =1315.66
       2-Within xij-xj           |             378   27.65 | Nj-bar  2868.00|   Prob > Chi2 =    0.00
       ------ of which: ------   |         ----------------| Nj-max      3978| idx effects:
        2a-res xij-xj-xi+x..     |           375.6   27.29 | Ni-min         2|  Chi2(1)       =   22.99
        2b-between xi-x..        |           38.84    0.36 | Ni-bar      2.16|   Prob > Chi2 =    0.00
        (2b in % of 2)           |                  (1.31) | Ni-max         3| Balancedness   =   72.10
----------------------------------------------------------------------------------------------------
   -- idx as fixed effects: --   |                          |                 |
       3-Between xj-x..          |           448.6   72.40 |                 | mom effects:
       4-Within xij-xj           |           377.7   27.60 |                 |  Chi2(1)       =1336.67
       ------ of which: ------   |         ----------------|                 |   Prob > Chi2 =    0.00
       4a-res xij-xj-xi+x..      |           375.6   27.29 |                 | idx effects:
       4b-between xi-x..         |           36.05    0.31 |                 |  Chi2(2)       =   55.14
        (4b in % of 4)           |                  (1.13) |                 |   Prob > Chi2 =    0.00
----------------------------------------------------------------------------------------------------
   -- idx as crossed-effect: --  |                          |                 |
       3-Between xj-x..          |             369   63.97 |                 | mom effects:
       4-Within xij-xj           |           377.7   36.03 |                 |  Chi2(1)       =1336.68
       ------ of which: ------   |         ----------------|                 |   Prob > Chi2 =    0.00
       4a-res xij-xj-xi+x..      |           375.7   35.63 |                 | idx effects:
       4b-between xi-x..         |           35.57    0.40 |                 |  Chi2(1)       =   22.99
        (4b in % of 4)           |                  (1.10) |                 |   Prob > Chi2 =    0.00
----------------------------------------------------------------------------------------------------
```

# A classical multilevel dataset III

```
-----------------------------------------------------------------------------------------------------
smoke    Overall xij-x..          |  .1399    .3469          | Ni x Nj    8604| mom effects:
       1-Between xj-x..           |           .3188    84.43 | Nj-min      648|  Chi2(1)       =3330.07
       2-Within xij-xj            |           .1867    15.57 | Nj-bar  2868.00|   Prob > Chi2 =    0.00
       ------ of which: ------    |         ----------------| Nj-max     3978| idx effects:
        2a-res xij-xj-xi+x..      |           .1867    15.57 | Ni-min        2|  Chi2(1)       =    0.00
        2b-between xi-x..         |        1.04e-09     0.00 | Ni-bar     2.16|   Prob > Chi2 =    0.50
        (2b in % of 2)            |                   (0.00) | Ni-max        3| Balancedness  =   72.10
-----------------------------------------------------------------------------------------------------
   -- idx as fixed effects: --    |                          |                |
       3-Between xj-x..           |           .3188    84.43 |                | mom effects:
       4-Within xij-xj            |           .1867    15.57 |                |  Chi2(1)       =3325.32
       ------ of which: ------    |         ----------------|                |   Prob > Chi2 =    0.00
       4a-res xij-xj-xi+x..       |           .1867    15.57 |                | idx effects:
       4b-between xi-x..          |         .003505     0.01 |                |  Chi2(2)       =    0.74
        (4b in % of 4)            |                   (0.04) |                |   Prob > Chi2 =    0.69
-----------------------------------------------------------------------------------------------------
   -- idx as crossed-effect: --   |                          |                |
       3-Between xj-x..           |           .2942    82.21 |                | mom effects:
       4-Within xij-xj            |           .1867    17.79 |                |  Chi2(1)       =3330.07
       ------ of which: ------    |         ----------------|                |   Prob > Chi2 =    0.00
       4a-res xij-xj-xi+x..       |           .1867    17.79 |                | idx effects:
       4b-between xi-x..          |        2.32e-08     0.00 |                |  Chi2(1)       =    0.00
        (4b in % of 4)            |                   (0.00) |                |   Prob > Chi2 =    0.50
-----------------------------------------------------------------------------------------------------
```

# A classical multilevel dataset IV

```
---------------------------------------------------------------------------------------------------
cigs     Overall xij-x..          |  2.269    8.904          | Ni x Nj    8604| mom effects:
       1-Between xj-x..           |           7.288    67.00 | Nj-min      648|  Chi2(1)      = 752.75
       2-Within xij-xj            |           6.976    33.00 | Nj-bar  2868.00|   Prob > Chi2 =   0.00
       ------ of which: ------    |         ----------------| Nj-max     3978| idx effects:
        2a-res xij-xj-xi+x..      |           6.977    33.00 | Ni-min        2|  Chi2(1)      =   0.00
        2b-between xi-x..         |        4.77e-07     0.00 | Ni-bar     2.16|   Prob > Chi2 =   0.50
        (2b in % of 2)            |                   (0.00) | Ni-max        3| Balancedness  =  72.10
---------------------------------------------------------------------------------------------------
   -- idx as fixed effects: --    |                          |                |
       3-Between xj-x..           |           7.288    66.99 |                | mom effects:
       4-Within xij-xj            |           6.977    33.01 |                |  Chi2(1)      = 750.21
       ------ of which: ------    |         ----------------|                |   Prob > Chi2 =   0.00
       4a-res xij-xj-xi+x..       |           6.977    33.00 |                | idx effects:
       4b-between xi-x..          |           .1018     0.01 |                |  Chi2(2)      =   0.15
        (4b in % of 4)            |                   (0.03) |                |   Prob > Chi2 =   0.93
---------------------------------------------------------------------------------------------------
   -- idx as crossed-effect: --   |                          |                |
       3-Between xj-x..           |           5.587    54.40 |                | mom effects:
       4-Within xij-xj            |           6.975    45.60 |                |  Chi2(1)      = 752.75
       ------ of which: ------    |         ----------------|                |   Prob > Chi2 =   0.00
       4a-res xij-xj-xi+x..       |           6.977    45.60 |                | idx effects:
       4b-between xi-x..          |        .0000403     0.00 |                |  Chi2(1)      =   0.00
        (4b in % of 4)            |                   (0.00) |                |   Prob > Chi2 =   0.50
---------------------------------------------------------------------------------------------------
```

# A classical multilevel dataset V

Results similar to Rabe-Hesketh and Skrondal (2012). There is an average of 2.16 births per mother.

The random intercept or level-2 residual (a mother-specific error component which remains constant across births) represents the combined effects of omitted mother characteristics or unobserved heterogeneity at the mother level and ranges from 64 to 72%. This also represents the estimated intraclass correlation capturing the correlation (equal to 0.72) between the birth weights of siblings or tells us that 72% of the variance in birth weight is shared between siblings.

Birth weight and smoking vary more between mothers (clusters of mothers as units of analysis) than within mothers (level-1 newborns as units of analysis).

As the mother effect is shared by all responses by the same mother, it induces a within-mother dependence of 28-36% in birth weight and 16-18% in smoking behaviour (among the residuals). This indicates that birth weight and smoking behaviour are not stable across pregnancies and that weight is less stable than smoking behaviour.

The between level-1 variability of weight is very low (0.4) (the between level-1 variability of smoking is zero). This indicates that weight and smoking behaviour are not related to birth order (firstborn, secondborn and thirdborn).

# A cross-classified multilevel dataset I

A typical cross-classified data structure is the conifer system investigated by Ols et al. (2020) and Bontempi, Mairesse, et al. (2025), based on 10.643 trees sampled across the systematic grid of the continuous French national forest inventory (NFI) over the 2006–2016 period.

A national 1 km $\times$ 1 km grid is used; every year, one tenth of the grid is inventoried and 5,442 plots classified as forests by the photo interpretation of aerial images are visited in the field in order to deliver fully renewed information on forests. Sampling plots are temporary, specie-homogeneous and consist in 25-m radius circles within which two dominants trees, chosen within the main tree species encountered on the plot, are measured (241 plots, 4% of the total, have only one tree). The panel is not a strict three-level hierarchy. Rather it is a cross-classified multilevel model where trees are nested within the cells formed by the cross-classification of species-by-ecological regions.

Therefore, the growth of tree-plots (measured by radial increment or diameter in metres and logs) can be influenced by both species and ecological regions, which are both sources of variation when considering tree-plots nested in the cross-classification of both species and ecological regions. Hence, there are two separate two-level hierarchies which cross one another:
(1) a tree-plot-within-specie hierarchy, and
(2) a tree-plot-within-ecological-region hierarchy.

# A cross-classified multilevel dataset II

Table: Distribution of Tree Species by Ecological regions

| | 1-Crystalline Oceanic West | 2-Semi-Oceanic Central North | 3-Semi-Continental East | 4-Vosges | 5-Oceanic South West | 6-Massif Central | 7-Alps | 8-Mediterranean | Total |
|---|---|---|---|---|---|---|---|---|---|
| Douglas fir | 0 | 551 (A) (48.7) | 0 | 0 | 0 | 1,145 (A) (35.9) | 0 | 0 | 1,696 (15.9) |
| Norway spruce | 0 | 0 | 337 (A) (100) | 303 (47.8) | 0 | 730 (A) (22.9) | 279 (19.3) | 0 | 1,649 (15.5) |
| European larch | 0 | 0 | 0 | 0 | 0 | 0 | 341 (23.6) | 0 | 341 (3.2) |
| Aleppo pine | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 455 (100) | 455 (4.3) |
| Corsican pine | 482 (A) (44.5) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 482 (4.5) |
| Maritime pine | 602 (A) (55.5) | 0 | 0 | 0 | 2,366 (A) (100) | 0 | 0 | 0 | 2,968 (27.9) |
| Scots pine | 0 | 581 (A) (51.3) | 0 | 0 | 0 | 725 (22.7) | 826 (57.1) | 0 | 2,132 (20.0) |
| Silver fir | 0 | 0 | 0 | 331 (52.2) | 0 | 589 (18.5) | 0 | 0 | 920 (8.6) |
| **Total** | 1,084 (10.19) | 1,132 (10.64) | 337 (3.17) | 634 (5.96) | 2,366 (22.23) | 3,189 (29.96) | 1,446 (13.59) | 455 (4.28) | 10,643 (100) |

*Note*: (A) for afforested trees.

# A cross-classified multilevel dataset III

```
egen treespid=group(treesp), label
egen ecoregid=group(ecoreg), label
pantarei ecoregid treespid ri1 diameter13 lri1 ldiam, ml
```

# A cross-classified multilevel dataset IV

```
--------------------------------- Level 2 (j): ecoreg   Level 1 (i): treesp  -----------------------
                                  Hypothetical full sample, Ni x Nj: 9464952
                                  Nj: 2968   Ni: 3189
----------------------------------------------------------------------------------------------------
Variable                         |   Mean   Std.Dev.  % SS |  Observations   |  Test
---------------------------------+-------------------------+----------------+-----------------------
ri1      Overall xij-x..         | .00285  .002085         | Ni x Nj   10643| ecoreg effects:
       1-Between xj-x..          |        .0004631    4.31 | Nj-min      341|  Chi2(1)      =1393.53
       2-Within xij-xj           |         .002041   95.69 | Nj-bar  1330.38|   Prob > Chi2 =   0.00
       ------ of which: ------   |         ----------------| Nj-max     2968| treesp effects:
        2a-res xij-xj-xi+x..     |         .001851   78.67 | Ni-min      337|  Chi2(1)      =2232.66
        2b-between xi-x..        |        .0009197   17.02 | Ni-bar  1330.38|   Prob > Chi2 =   0.00
        (2b in % of 2)           |                 (17.79)| Ni-max     3189| Balancedness  =   0.11
----------------------------------------------------------------------------------------------------
   -- treesp as fixed effects: --|                         |                |
       3-Between xj-x..          |        .0007936   12.67 |                | ecoreg effects:
       4-Within xij-xj           |         .001949   87.33 |                |  Chi2(1)      = 224.55
       ------ of which: ------   |         ----------------|                |   Prob > Chi2 =   0.00
       4a-res xij-xj-xi+x..      |         .001851   78.67 |                | treesp effects:
       4b-between xi-x..         |         .000656    8.66 |                |  Chi2(7)      =1191.27
        (4b in % of 4)           |                  (9.92) |                |   Prob > Chi2 =   0.00
----------------------------------------------------------------------------------------------------
   -- treesp as crossed-effect: --|                        |                |
       3-Between xj-x..          |        .0004482    4.34 |                | ecoreg effects:
       4-Within xij-xj           |         .001969   95.66 |                |  Chi2(1)      = 240.59
       ------ of which: ------   |         ----------------|                |   Prob > Chi2 =   0.00
       4a-res xij-xj-xi+x..      |         .001851   84.49 |                | treesp effects:
       4b-between xi-x..         |        .0007192   11.17 |                |  Chi2(1)      =2232.66
        (4b in % of 4)           |                 (11.68)|                |   Prob > Chi2 =   0.00
----------------------------------------------------------------------------------------------------
```

# A cross-classified multilevel dataset V

```
----------------------------------------------------------------------------------------------------
diame~13  Overall xij-x..        |  .3329    .1364          | Ni x Nj   10643| ecoreg effects:
   -- treesp as fixed effects: --|                          |                |
       3-Between xj-x..          |          .03751    6.62  |                | ecoreg effects:
       4-Within xij-xj           |           .1319   93.38  |                |  Chi2(1)      = 223.72
       ------ of which: ------   |         ----------------|                |   Prob > Chi2 =    0.00
       4a-res xij-xj-xi+x..      |           .1259   85.02  |                | treesp effects:
       4b-between xi-x..         |          .04216    8.36  |                |  Chi2(7)      =1042.74
        (4b in % of 4)           |                  (8.95)  |                |   Prob > Chi2 =    0.00
----------------------------------------------------------------------------------------------------
   -- treesp as crossed-effect: --|                         |                |
       3-Between xj-x..          |          .03676    5.81  |                | ecoreg effects:
       4-Within xij-xj           |           .1385   94.19  |                |  Chi2(1)      = 236.12
       ------ of which: ------   |         ----------------|                |   Prob > Chi2 =    0.00
       4a-res xij-xj-xi+x..      |           .1259   77.78  |                | treesp effects:
       4b-between xi-x..         |          .06179   16.42  |                |  Chi2(1)      =1404.65
        (4b in % of 4)           |                 (17.43)|                |   Prob > Chi2 =    0.00
----------------------------------------------------------------------------------------------------
```

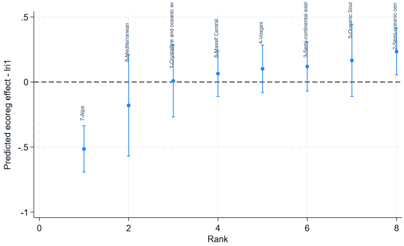
# A cross-classified multilevel dataset VI

```
---------------------------------- Level 2 (j): ecoreg   Level 1 (i): treesp  -----------------------
                                   Hypothetical full sample, Ni x Nj: 9464952
                                   Nj: 2968   Ni: 3189
-----------------------------------------------------------------------------------------------------
Variable                          |   Mean   Std.Dev.  % SS  | Observations   |  Test
----------------------------------+--------------------------+----------------+-----------------------
lri1     Overall xij-x..          | -6.151     .825          | Ni x Nj   10643| ecoreg effects:
       1-Between xj-x..           |           .2522     8.18 | Nj-min      341|  Chi2(1)       =2414.64
       2-Within xij-xj            |           .7908    91.82 | Nj-bar  1330.38|   Prob > Chi2 =    0.00
       ------ of which: ------    |         ----------------| Nj-max     2968| treesp effects:
        2a-res xij-xj-xi+x..      |           .6796    67.78 | Ni-min      337|  Chi2(1)       =3370.56
        2b-between xi-x..         |           .4324    24.04 | Ni-bar  1330.38|   Prob > Chi2 =    0.00
        (2b in % of 2)            |                  (26.18)| Ni-max     3189| Balancedness  =    0.11
-----------------------------------------------------------------------------------------------------
   -- treesp as fixed effects: --|                          |                |
       3-Between xj-x..           |           .4012    20.70 |                | ecoreg effects:
       4-Within xij-xj            |           .7349    79.30 |                |  Chi2(1)       = 660.74
       ------ of which: ------    |         ----------------|                |   Prob > Chi2 =    0.00
       4a-res xij-xj-xi+x..       |           .6796    67.78 |                | treesp effects:
       4b-between xi-x..          |           .2994    11.52 |                |  Chi2(7)       =1818.63
        (4b in % of 4)            |                  (14.53)|                |   Prob > Chi2 =    0.00
-----------------------------------------------------------------------------------------------------
   -- treesp as crossed-effect: --|                          |                |
       3-Between xj-x..           |           .2575     9.60 |                | ecoreg effects:
       4-Within xij-xj            |           .7393    90.40 |                |  Chi2(1)       = 680.28
       ------ of which: ------    |         ----------------|                |   Prob > Chi2 =    0.00
       4a-res xij-xj-xi+x..       |           .6798    76.38 |                | treesp effects:
       4b-between xi-x..          |           .3112    14.02 |                |  Chi2(1)       =3370.56
        (4b in % of 4)            |                  (15.51)|                |   Prob > Chi2 =    0.00
-----------------------------------------------------------------------------------------------------
```

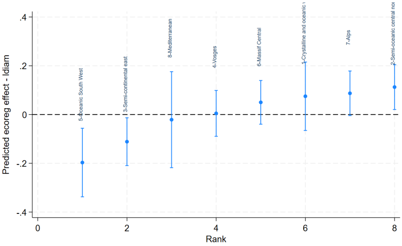
# A cross-classified multilevel dataset VII

```
----------------------------------------------------------------------------------------------------
ldiam    Overall xij-x..          | -1.196    .4623          | Ni x Nj   10643| ecoreg effects:
   -- treesp as fixed effects: --|                          |                |
       3-Between xj-x..          |           .1261    6.51 |                | ecoreg effects:
       4-Within xij-xj           |           .4471   93.49 |                |  Chi2(1)      = 219.08
       ------ of which: ------   |         ----------------|                |   Prob > Chi2 =   0.00
       4a-res xij-xj-xi+x..      |           .4322   87.30 |                | treesp effects:
       4b-between xi-x..         |           .1229    6.19 |                |  Chi2(7)      = 750.28
        (4b in % of 4)           |                  (6.62) |                |   Prob > Chi2 =   0.00
----------------------------------------------------------------------------------------------------
   -- treesp as crossed-effect: --|                          |                |
       3-Between xj-x..          |           .1181    5.30 |                | ecoreg effects:
       4-Within xij-xj           |           .4672   94.70 |                |  Chi2(1)      = 230.75
       ------ of which: ------   |         ----------------|                |   Prob > Chi2 =   0.00
       4a-res xij-xj-xi+x..      |           .4323   81.02 |                | treesp effects:
       4b-between xi-x..         |           .1898   13.68 |                |  Chi2(1)      =1130.98
        (4b in % of 4)           |                 (14.45)|                |   Prob > Chi2 =   0.00
----------------------------------------------------------------------------------------------------
```

# A cross-classified multilevel dataset VIII

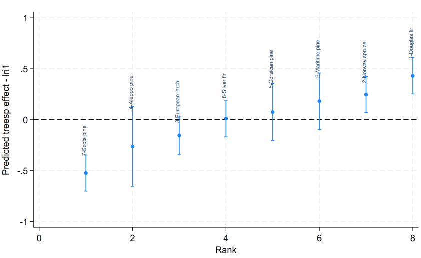



Ecological regions for radial increment in logs

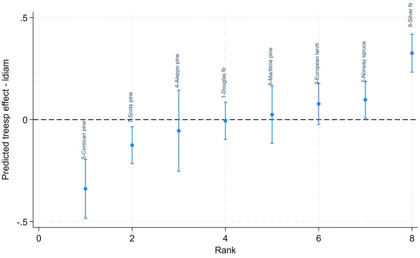



Ecological regions for diameter in logs

# A cross-classified multilevel dataset IX




Species for radial increment in logs




Species for diameter in logs

# A cross-classified multilevel dataset X

A different picture emerges when we compare radial increment and diameter (in logs). Diameter can be affected by measurement errors (inaccuracy) because the bark is also included, but this does not grow in the same way as wood (rings).

We see that 14.02-13.68% of the variation in growth lies between species, while 9.60-5.30% lies between ecological regions, and 76.38-81.02% lies between tree-plots. Thus, there are stronger growth disparities across the 8 species than there are across the 8 ecoregs. This finding is surprising given that the number of ecoregs in the data is equal to the number of species.

In terms of ICCs, we note that homogeneity or correlation in growth at the species level is greater than at the ecological region level (here coincide with VPCs). The greatest homogeneity is found in trees of the same species living in the same ecological region, computed as 23.62-18.98%.

These results guide the modelling of conditional models, in which adjusted ICCs are based on the residuals of the conditional model rather that on the observations of the dependent variable alone, allowing us to assess how the addition of explanatory variables can modify the unexplained variability and similarity between and within the various clusters.

Thank you for your attention!

# References I


Abrevaya, J. (2006). "Estimating the effect of smoking on birth outcomes using a matched panel data approach". In: Journal of Applied Econometrics 21.2, pp. 489–519.

Alker, H. S. (1969). "A typology of ecological fallacies". In: Quantitative Ecological Analysis. Ed. by M. Dogan and S. Rokkan. MIT Press, Mass.

Arellano, M. and S. R. Bond (1991). "Some tests of specification for panel data: Monte Carlo Evidence and an application to employment equations". In: Review of Economic Studies 58.2, pp. 277–297.

Baltagi, B. H. and L. Mátyás (2025). Seven decades of econometrics and beyond. A tribute to the life and work of Marc Nerlove. Springer Cham.

Bellemare, M. F. and D. L. Millimet (2025). "Retrospectives: Yair Mundlak and the fixed effects estimator". In: Journal of Economic Perspectives 39:1, pp. 261–274.

Beretvas, S. N. (2011). "Cross-classified and multiple membership models". In: Handbook of Advanced Multilevel Analysis. Ed. by J. J. Hox and J. K. Roberts. Routledge, Taylor & Francis Group, New York and Hove.

Blundell, R. W. and S. R. Bond (1998). "Initial conditions and moment restrictions in dynamic panel data models". In: Journal of Econometrics 87, pp. 115–143.

Bontempi, M. E. and L. Bottazzi (2025). "Heterogeneity of covenant violations and corporate financial behavior". In: WP DSE.

# References II


Bontempi, M. E. and H. Kean (2025). "Do Good Institutions Defer FDI to Natural Resource-Rich Developing Countries? A Multilevel Dynamic Analysis". In: WP DSE.

Bontempi, M. E., J. Mairesse, and C. Ols (2025). "Resilience of French conifers to climate change". In: WP DSE.

Bou, J. C. and A. Satorra (2017). "Univariate Versus Multivariate Modeling of Panel Data: Model Specification and Goodness-of-Fit Testing". In: Organizational Research Methods 21:1, pp. 1–47.

Brown, N. L., P. Schmidt, and J. M. Wooldridge (2021). "Simple Alternatives to the Common Correlated Effects Model". In: arXiv:2112.01486 [econ.EM].

Bryk, A. and S. Raudenbush (2002). Hierarchical Linear Models, Advanced Quantitative Techniques. SAGE, New York.

Chamberlain, G. (1980). "Analysis of covariance with qualitative data". In: Review of Economic Studies 47, pp. 225–238.

de Jong, a., M. J. C. Verbeek, and P. Verwijmeren (2011). "Firms' debt–equity decisions when the static tradeoff theory and the pecking order theory disagree". In: Journal of Banking & Finance 35, pp. 1303–1314.

Elliott, G., T. J. Rothenberg, and J. H. Stock (1996). "Efficient tests for an autoregressive unit root". In: Econometrica 64, pp. 813–836.

# References III


- Gibbons, S., H. G. Overman, and P. Pelkonen (2014). “Area Disparities in Britain: Understanding the contribution of people vs. place through variance decomposition”. In: Oxford Bulletin of Economics and Statistics 75, pp. 745–763.
- Goldstein, H. (1987). Multilevel Models in Education and Social Research. Charles Griffin & Co., Oxford University Press.
- Goldstein, H. (2011). Multilevel Statistical Models. John Wiley & Sons, Ltd, 4th Ed.
- Hox, J. J. and J. K. Roberts (2011). Handbook of Advanced Multilevel Analysis. Routledge, Taylor & Francis Group, New York and Hove.
- McNeish, D. (2023). “A practical guide to selecting and blending approaches for clustered data: Clustered errors, multilevel models, and fixed-effect models”. In: Psychological Methods. DOI: https://doi.org/10.1037/met0000620.
- McNeish, D. and K. Kelley (2019). “Fixed effects models versus mixed effects models for clustered data: reviewing the approaches, disentangling the differences, and making recommendations”. In: Psychological Methods 24, pp. 20–35.
- Mundlak, Y. (1978). “On the pooling of time series and cross-sectional data”. In: Econometrica 46, pp. 69–86.
- Ols, C., J.-C. Hervé, and J.-D. Bontemps (2020). “Recent growth trends of conifers across Western Europe are controlled by thermal and water constraints and favored by forest heterogeneity”. In: Science of the Total Environment 742, 140453.

# References IV


- Pesaran, M. H. (2006). "Estimation and inference in large heterogeneous panels with a multifactor error structure". In: Econometrica 74.4, pp. 967–1012.
- Rabe-Hesketh, S. and A. Skrondal (2012). Multilevel and Longitudinal Modeling Using Stata, Third edition. Stata Press Publication.
- Rasbash, J. and W. J. Browne (2001). "Modelling non-hierarchical structures". In: Multilevel Modelling of Health Statistics. Ed. by A. H. Leyland and H. Goldstein. Chichester: Wiley.
- Rasbash, J. and W. J. Browne (2008). "Non-hierarchical multilevel models". In: Handbook of Multilevel Analysis. Ed. by D. J. and M. E. New York: Springer.
- Rasbash, J. and H. Goldstein (1994). "Efficient analysis of mixed hierarchical and cross-classified random structures using a multilevel model". In: Journal of Educational and Behavioral Statistics 19, pp. 337–350.
- Robinson, W. S. (1950). "Ecological correlation and the behaviour of individuals". In: American Sociological Review 15, pp. 351–357.
- Snijders, T. and R. Bosker (2012). Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Analysis, 2nd edition. Sage.