PANEL DATA Models

Some BASIC considerations [CT, p. 230]

(1) Usually observed at regular intervals in time. - There are some exceptions, though.

(2) BALANCED $T_i = T$ for all i

- no missing time observations for ANY individual. sample size = NT

UNBALANCED At least 1 missing time obs for a subset of individuals

sample size < NT

(3) small N - long T

large N - short T

large N - large T

The distinction is important for estimation and inference

(4) Model errors are likely correlated

Units = Individuals

Usually, we expect errors for a unit (i) to be correlated. The units may be uncorrelated with one another though.
⇒ "CLUSTER"

Units = Countries states, etc

It is also possible that there may be correlation ACROSS individuals. Macro events could affect all countries.

Either way, GLS or some Robust cov. estimation is required to get consistent std errors, tests, and CI.

(5) Identification depends on type of regressor parameters of
(a) time invariant regressors e.g., gender, Race may not be estimable.

(b) time trends won't vary by individual. – or macro effects.

(6) Some coeffs may vary across individual and/or time.

(7) Fixed effects estimation is used A lot in Econ Because endogenous regressors can be used if reg are correlated ONLY with the individual's time-invariant characteristics.

May other disciplines rely on Random effects, which are often inconsistent if regressors are not completely exogenous.

(8) Dynamic panels are sometimes estimated. – They include lagged dependent variables As regressors.

Basic Panel Models

Individual Effect.

$$y_{it} = \alpha_i + x_{it}^T \beta + u_{it}$$

$i = 1, \ldots, N$

$t = 1, \ldots, T$

$x_{it}^T = \{ \; x_{2it} \; x_{3it} \cdots x_{12it} \}$

$\beta^T = (\; , \beta_2, \ldots, \beta_k)$

α_i individual effect.
<u>time-invariant</u>.

Here, the individual effect <u>can</u> be correlated with the variables in x_{it}. Basically, you have a parameter, α_i, to estimate for each individual (fixed)

It assumes that

$$E\left[u_{it} \mid \alpha_i, x_{it}\right] = 0$$

Random Effects

$\alpha_i \sim (0, \sigma_\alpha^2)$ is random and unobservable

$$y_{it} = x_{it}^T \beta + \bar{\beta}_1 + \varepsilon_{it}$$

$$\varepsilon_{it} = (u_{it} + \alpha_i)$$

α_i, being random, cannot be correlated with x_{it} in this case. However, you can estimate parameters and ME for time-invariant regressors.

Pooled or Population Averaged

$$y_{it} = \alpha + \underset{\sim}{x}_{it}^T \underline{\beta} + u_{it}$$

LS consistent as long as individual effects (in u_{it}) are <u>NOT</u> correlated with $\underset{\sim}{x}_{it}$

with this one, you'll want to use "cluster" std errors since it is likely that errors within an individual unit are correlated.

Two-way fixed Effects

$$y_{it} = \alpha_i + \gamma_t + x_{it}^T \beta + u_{it}$$

time dummies (drop 1)
+ individual dummies (no const in x_{it})

Mixed linear Models

$$y_{it} = \alpha_i + x_{it}^T \beta_i + u_{it}$$

$$(\alpha_i, \beta_i^T)^T \sim (\beta, \Sigma)$$

For long panels, you could run separate regressions for each i.