EC968

Panel Data Analysis

Steve Pudney
ISER

Lecture 2: Linear regression for panel data

- Within-group ("fixed effects") regression
- Asymptotics for short panels
- Random effects regression
- Testing the zero covariance assumption

Linear regression for panel data

The "standard" panel data model is:

$$y_{it} = \mathbf{z}_i \boldsymbol{\alpha} + \mathbf{x}_{it} \boldsymbol{\beta} + u_i + \varepsilon_{it}$$

We have observations indexed by $t = 1 \ldots T_i$, $i = 1 \ldots n$.

- A pooled regression of y on $\mathbf{z}$ and $\mathbf{x}$ ignores the individual effect u, and therefore isn't appropriate.
- The u_i can be captured using dummy variables. Construct a set of n dummy variables $D1_{it} \ldots Dn_{it}$, where:

$$Dr_{it} = 1 \text{ if } i = r \text{ and } 0 \text{ otherwise, for } r = 1 \ldots n$$

Thus Dr_{it} tells us whether observation i, t relates to person r.

- The model is now:

$$y_{it} = \mathbf{z}_i \boldsymbol{\alpha} + \mathbf{x}_{it} \boldsymbol{\beta} + u_1 D1_{it} + \ldots + u_n Dn_{it} + \varepsilon_{it}$$

Thus $u_1 \ldots u_n$ are now seen as the coefficients of a set of n dummy variables.

Shortcut calculation of the dummy variable regression

The Frisch-Waugh theorem on partitioned regression tells us that a multiple regression of y on $(\mathbf{z}, \mathbf{x})$ and $(D1 \ldots Dn)$ can be done in two stages:

Stage 1: regress y on $(D1 \ldots Dn)$ and each of the variables in $(\mathbf{z}, \mathbf{x})$ on $(D1 \ldots Dn)$; replace y and $(\mathbf{z}, \mathbf{x})$ by their residuals from these regressions $\Rightarrow y^*$ and $(\mathbf{z}^*, \mathbf{x}^*)$

Stage 2: regress y^* on $(\mathbf{z}^*, \mathbf{x}^*)$

It can be shown that, in our case, the residuals y^* and $(\mathbf{z}^*, \mathbf{x}^*)$ are:

$$y_{it}^* = y_{it} - \bar{y}_i$$

$$\mathbf{x}_{it}^* = \mathbf{x}_{it} - \bar{\mathbf{x}}_i$$

$$\mathbf{z}_i^* = \mathbf{z}_i - \bar{\mathbf{z}}_i \equiv \mathbf{0}$$

Thus, least-squares dummy variables (LSDV) is equivalent to a regression of $y_{it} - \bar{y}_i$ on $\mathbf{x}_{it} - \bar{\mathbf{x}}_i$, with $\mathbf{z}$ eliminated from the model (since $\mathbf{z}$ is collinear with $D1 \ldots Dn$).

Another interpretation of LSDV

Start differently, by thinking how we can cope with u_i

We don't know its statistical properties, so let's try to eliminate it from the model. We can eliminate it in various ways, for example:

Time differencing: $y_{it} - y_{it-1} = (\mathbf{x}_{it} - \mathbf{x}_{it-1})\boldsymbol{\beta} + \varepsilon_{it} - \varepsilon_{it-1}$

or

Within-group transform: $y_{it} - \bar{y}_i = (\mathbf{x}_{it} - \bar{\mathbf{x}}_i)\boldsymbol{\beta} + \varepsilon_{it} - \bar{\varepsilon}_i$

The Frisch-Waugh theorem tells us that the within-group approach is the most efficient in the least squares sense

A note on terminology

Different names are commonly used for this one estimation method:

- Least squares dummy variables (LSDV)
- Within-group regression
- Fixed-effects regression
- Covariance analysis regression

- "LSDV" refers to the method of derivation using explicit dummy variables;
- "within-group" refers to the type of data transform implied by the method;
- "fixed effects" is common but very poor terminology which suggests (wrongly, in the case of sample survey data) that the u_i are fixed parameters
- "covariance analysis" reflects the origins of the method as a generalisation of analysis of variance used in agricultural experiments

Coefficient estimates

The within-group regression is:

$$\hat{\boldsymbol{\beta}} = \mathbf{W}_{xx}^{-1}\mathbf{w}_{xy} = \boldsymbol{\beta} + \mathbf{W}_{xx}^{-1}\mathbf{w}_{x\varepsilon}$$

where $\mathbf{W}_{xx}$, $\mathbf{w}_{xy}$ and $\mathbf{w}_{x\varepsilon}$ are within-group moment matrices:

$$\mathbf{W}_{xx} = n^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T_i}\left(\mathbf{x}_{it} - \bar{\mathbf{x}}_i\right)'\left(\mathbf{x}_{it} - \bar{\mathbf{x}}_i\right)$$

$$\mathbf{w}_{x\varepsilon} = n^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T_i}\left(\mathbf{x}_{it} - \bar{\mathbf{x}}_i\right)'\left(\varepsilon_{it} - \bar{\varepsilon}_i\right)$$

If $\mathbf{x}_{it}$ and ε_{it} are uncorrelated, $\mathrm{E}(\mathbf{w}_{x\varepsilon}) = \mathbf{0}$, so:

$$E\hat{\boldsymbol{\beta}} = \boldsymbol{\beta}$$

Residuals

There are two residuals for the within-group regression:

$$\hat{e}_i = \bar{y}_i - \bar{\mathbf{x}}_i \hat{\boldsymbol{\beta}}$$

$$\hat{\varepsilon}_{it} = (y_{it} - \bar{y}_i) - (\mathbf{x}_{it} - \bar{\mathbf{x}}_i)\hat{\boldsymbol{\beta}} = y_{it} - \mathbf{x}_{it}\hat{\boldsymbol{\beta}} - \hat{e}_i$$

$\hat{e}_i$ is an estimate of $\mathbf{z}_i\boldsymbol{\alpha} + u_i$; $\hat{\varepsilon}_{it}$ is an estimate of ε_{it}

Since $\hat{\varepsilon}_{it}$ is the residual from a multiple regression, its sample variance is an unbiased estimator of σ_ε^2 under the classical assumptions of independent sampling of individuals and:

$$E\varepsilon_{it} = 0 \,; \quad E\varepsilon_{it}{}^2 = \sigma_\varepsilon^2$$

$$E\mathbf{x}_{is}\varepsilon_{it} = \mathbf{0} \quad \text{for all } i, s, t$$

$$E\varepsilon_{is}\varepsilon_{it} = 0 \ \text{ for all } i, s \neq t$$

Estimation of α

The residual $\hat{e}_i$ can be written:

$$\hat{e}_i = \bar{y}_i - \bar{\mathbf{x}}_i \hat{\boldsymbol{\beta}} = \left(\mathbf{z}_i \boldsymbol{\alpha} + \bar{\mathbf{x}}_i \boldsymbol{\beta} + u_i + \bar{\varepsilon}_i\right) - \bar{\mathbf{x}}_i \hat{\boldsymbol{\beta}}$$
$$= \mathbf{z}_i \boldsymbol{\alpha} + u_i + \bar{\varepsilon}_i - \bar{\mathbf{x}}_i \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)$$

Since $\hat{e}_i$ is an estimate of $\mathbf{z}_i\boldsymbol{\alpha} + u_i$, we could regress it on $\mathbf{z}_i$ to estimate $\boldsymbol{\alpha}$. (Use T_i repeated observations on the group means for individual i, to weight individuals appropriately). This gives:

$$\hat{\boldsymbol{\alpha}} = \mathbf{B}_{zz}^{-1} \mathbf{b}_{z\hat{e}}$$

where $\mathbf{B}_{xx}$ *etc.* are between-group cross-product matrices:

$$\mathbf{B}_{zz} = \sum_{i=1}^{n} \sum_{t=1}^{T_i} \bar{\mathbf{z}}' \bar{\mathbf{z}} = \sum_{i=1}^{n} T_i \bar{\mathbf{z}}' \bar{\mathbf{z}}; \quad \mathbf{b}_{z\hat{e}} = \sum_{i=1}^{n} \sum_{t=1}^{T_i} \bar{\mathbf{z}}' \hat{e}_i$$

Estimation of $\hat{\boldsymbol{\alpha}}$

Rewrite $\hat{\boldsymbol{\alpha}}$ as:

$$\hat{\boldsymbol{\alpha}} = \mathbf{B}_{zz}^{-1}\mathbf{b}_{z\hat{e}} = \boldsymbol{\alpha} + \mathbf{B}_{zz}^{-1}\mathbf{b}_{zu} + \mathbf{B}_{zz}^{-1}\mathbf{b}_{z\varepsilon} - \mathbf{B}_{zz}^{-1}\mathbf{B}_{zx}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)$$

But $\hat{\boldsymbol{\beta}}$ is unbiased and we assume $\mathbf{z}_i$ is uncorrelated with ε_{it} , so:

$$E\hat{\boldsymbol{\alpha}} = \boldsymbol{\alpha} + E\left(\mathbf{B}_{zz}^{-1}\mathbf{b}_{zu}\right)$$

Thus $\hat{\boldsymbol{\alpha}}$ is only unbiased if u_i and $\mathbf{z}_i$ are uncorrelated.

Estimation of σ_u^2

One way is to use the between-group regression. Replace each observation by the individual mean:

$$\bar{y}_i = \mathbf{z}_i\boldsymbol{\alpha} + \bar{\mathbf{x}}_i\boldsymbol{\beta} + u_i + \bar{\varepsilon}_i \ , \qquad i = 1\ldots n;\, t = 1\ldots T_i$$

Estimator: $$\begin{pmatrix} \hat{\boldsymbol{\alpha}} \\ \hat{\boldsymbol{\beta}} \end{pmatrix} = \begin{pmatrix} \mathbf{B}_{zz} & \mathbf{B}_{zx} \\ \mathbf{B}_{xz} & \mathbf{B}_{xx} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{b}_{zy} \\ \mathbf{b}_{xy} \end{pmatrix}$$

The residual variance is an estimate of $\bar{T}\sigma_u^2 + \sigma_\varepsilon^2$ so:

$$\hat{\sigma}_u^2 = \frac{s_B^2 - s_W^2}{\bar{T}}$$

where s_B^2 and s_W^2 are the b-g and w-g residual variances and $\bar{T}$ is the mean no. of observations per individual.

Note that $\hat{\sigma}_u^2$ may be negative!

Asymptotics for short panels

For panel data arising from repeated surveys, n is usually much larger than $T = \max(T_i)$. This suggests using asymptotic theory based on $n \to \infty$, with all T_i fixed.

Incidental parameters problem: If we regard the unobserved effects $u_1 \ldots u_n$ as parameters to be estimated, then the dimension of the parameter space $\to \infty$ as $n \to \infty$. Standard asymptotic theory doesn't work in this case.

Consistency of within-group estimator:

$$\begin{aligned}
\operatorname*{plim}_{n\to\infty} \hat{\boldsymbol\beta}_W &= \boldsymbol\beta + \operatorname*{plim}_{n\to\infty}\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T_i}(\mathbf{x}_{it}-\bar{\mathbf{x}}_i)'(\mathbf{x}_{it}-\bar{\mathbf{x}}_i)\right)^{-1} \\
&\quad \times \operatorname*{plim}_{n\to\infty}\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T_i}(\mathbf{x}_{it}-\bar{\mathbf{x}}_i)'(\varepsilon_{it}-\bar{\varepsilon}_i)\right) \\
&= \boldsymbol\beta + \left(\operatorname*{plim}_{n\to\infty} \mathbf{W}_{xx}\right)^{-1} \times \mathbf{0} \quad = \boldsymbol\beta
\end{aligned}$$

Example of panel data estimation

The Stata command *xtreg* computes within-group and between-group regressions

Example: within- and between-group regressions of log earnings on age, year of birth and time, allowing for unobserved individual effects:

gen age=year-cohort

gen logearn=ln(w_hr)

xtreg logearn age cohort year, fe

xtreg logearn age cohort year, be

Stata output: within-group regression

```
. xtreg logearn age cohort year, fe

Fixed-effects (within) regression               Number of obs      =     21124
Group variable (i): pid                         Number of groups   =      5859

R-sq:  within  = 0.1255                         Obs per group: min =         1
       between = 0.0027                                        avg =       3.6
       overall = 0.0064                                        max =        11

                                                F(1,15264)         =   2191.42
corr(u_i, Xb)  = -0.4165                        Prob > F           =    0.0000

------------------------------------------------------------------------------
     logearn |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0302855    .000647    46.81   0.000     .0290174    .0315536
      cohort |  (dropped)
        year |  (dropped)
       _cons |   .9004189   .0249395    36.10   0.000     .8515345    .9493033
-------------+----------------------------------------------------------------
     sigma_u |  .62294342
     sigma_e |  .24397194
         rho |  .86701327   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(5858, 15264) =    17.98         Prob > F = 0.0000
```

Stata output: between-group regression

```
. xtreg logearn age cohort, be

Between regression (regression on group means)  Number of obs      =     21124
Group variable (i): pid                         Number of groups   =      5859

R-sq:  within  = 0.1255                         Obs per group: min =         1
       between = 0.0027                                        avg =       3.6
       overall = 0.0081                                        max =        11

                                                F(2,5856)          =      7.92
sd(u_i + avg(e_i.))=  .5556311                  Prob > F           =    0.0004

------------------------------------------------------------------------------
     logearn |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0039101   .0026781     1.46   0.144    -.0013401    .0091602
      cohort |   .0010323   .0024038     0.43   0.668    -.0036801    .0057446
       _cons |  -.2244308   4.808276    -0.05   0.963    -9.650426    9.201565
------------------------------------------------------------------------------
```

Important points

- The within-group R^2 is much higher than the between-group R^2

  $\Rightarrow$ the covariate *age* (and/or *year* and/or *cohort*) "explains" a reasonable amount of the pay variation over time for a given individual

  $\Rightarrow$ but pay differences between individuals are not closely related to age and cohort

- The large coefficient differences between the within- and between-group age coefficients suggest that a single regression model with classical assumptions doesn't fit the evidence very well

‘Random effects’ GLS & ML estimation

- In general, since individuals are sampled at random from the population, u_i (and all other variables) are random: so “random effects” is tautological
- Extract the overall mean from u_i:

$$y_{it} = \alpha_0 + \mathbf{z}_i \boldsymbol{\alpha} + \mathbf{x}_{it} \boldsymbol{\beta} + u_i + \varepsilon_{it}$$

- We may choose to assume that u_i is mean-independent of $\mathbf{z}_i$ and $\mathbf{X}_i$ (implying also zero correlation):

$$\mathrm{E}(u_i \mid \mathbf{z}_i, \mathbf{X}_i) = 0$$

- Assume homoskedasticity and uncorrelatedness

$$\mathrm{E}(u_i^2 \mid \mathbf{z}_i, \mathbf{X}_i) = \sigma_u^2 \; ; \; \mathrm{E}(u_i \, \varepsilon_{it} \mid \mathbf{z}_i, \mathbf{X}_i) = 0 \quad \forall \, t$$

- Then write the composite random disturbance as:

$$v_{it} = u_i + \varepsilon_{it}$$

- What is the covariance matrix of the process $\{v_{it}\}$?

Random effects covariance structure

Variances & covariances (conditional on $\mathbf{z}_i$, $\mathbf{X}_i$) :

$$\text{var}(v_{it}) = \sigma_u^2 + \sigma_\varepsilon^2 ; \quad \text{cov}(v_{it}, v_{is}) = \sigma_u^2 \quad \forall\, s \neq t$$

Define the $T_i \times 1$ vector $\mathbf{v}_i$ with elements $v_{i1} \ldots v_{iT}$. Note that $\mathbf{v}_i$ and $\mathbf{v}_j$ are independent for $i \neq j$. The covariance matrix of $\mathbf{v}_i$ is:

$$\boldsymbol{\Omega}_i = \sigma_\varepsilon^2\, \mathbf{I} + \sigma_u^2\, \mathbf{E}$$

where **I** is the identity matrix and **E** is a matrix with each element equal to 1, both of order $T_i \times T_i$.

Lemma: the inverse of $\boldsymbol{\Omega}_i$ is:

$$\boldsymbol{\Omega}_i^{-1} = \frac{1}{\sigma_\varepsilon^2}\left(\mathbf{I} - \frac{T_i \sigma_u^2}{\sigma_\varepsilon^2 + T_i \sigma_u^2}\left(T_i^{-1}\mathbf{E}\right)\right) = \frac{1}{\sigma_\varepsilon^2}\left(\mathbf{M}_W + \psi_i \mathbf{M}_B\right)$$

Within- and between-group transformations

$$\mathbf{\Omega}_i^{-1} = \frac{1}{\sigma_\varepsilon^2}\left(\mathbf{M}_W + \psi_i \mathbf{M}_B\right)$$

The **M**-matrices are:

$$\mathbf{M}_W = \mathbf{I} - T_i^{-1}\mathbf{E}$$

$$\mathbf{M}_B = T_i^{-1}\mathbf{E}$$

$\mathbf{M}_W$ is the $T_i \times T_i$ idempotent matrix that transforms a $T_i \times 1$ vector of data to within-group mean deviation form;

$\mathbf{M}_B$ is the idempotent transformation to a $T_i \times 1$ vector of repeated means (the between-group transform).

The scalar $\psi_i = \sigma_\varepsilon^2 / \left(\sigma_\varepsilon^2 + T_i \sigma_u^2\right)$ reflects the relative size of $T_i {\sigma_u}^2$ and σ_ε^2.

Generalised Least Squares

For simplicity, subsume $\mathbf{z}_i$ within $\mathbf{x}_{it}$. Then GLS is:

$$\hat{\boldsymbol{\beta}}_{GLS} = \left(\sum_{i=1}^{n} \mathbf{X}_i ' \boldsymbol{\Omega}_i^{-1} \mathbf{X}_i \right)^{-1} \sum_{i=1}^{n} \mathbf{X}_i ' \boldsymbol{\Omega}_i^{-1} \mathbf{y}_i$$

$$= \left(\sum_{i=1}^{n} \left[\mathbf{W}_{xxi} + \psi_i \mathbf{B}_{xxi} \right] \right)^{-1} \sum_{i=1}^{n} \left[\mathbf{w}_{xyi} + \psi_i \mathbf{b}_{xyi} \right]$$

where $\mathbf{W}_{xxi} = \sum_{t=1}^{T_i} (\mathbf{x}_{it} - \bar{\mathbf{x}}_i)'(\mathbf{x}_{it} - \bar{\mathbf{x}}_i), \quad \mathbf{B}_{xxi} = T_i \, \bar{\mathbf{x}}_i ' \bar{\mathbf{x}}_i$, etc.

So GLS uses both within-group and between-group variation, but weights them according to the relative sizes of $\sigma_\varepsilon^2 + T_i \sigma_u^2$ and σ_ε^2.

Note that $\lim_{T_i \to \infty} \psi_i = 0$, so between-group variation is unimportant in a long panel

Feasible GLS

Separate out **z** and **x** again. It can be shown that GLS is equivalent to the following procedure:

(1) Transform the data:

$$y_{it}^{+} = y_{it} - \theta_i \bar{y}_i \; ; \quad \mathbf{z}_i^{+} = (1-\theta_i)\mathbf{z}_i \; ; \quad \mathbf{x}_{it}^{+} = \mathbf{x}_{it} - \theta_i \bar{\mathbf{x}}_i$$

where:

$$\theta_i = 1 - \sqrt{\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T_i \sigma_u^2}}$$

(2) Regress y_{it}^{+} on $\left(\mathbf{z}_i^{+}, \mathbf{x}_{it}^{+}\right)$, pooling all observations

The variance parameters σ_ε^2 and σ_u^2 can be estimated from the within-group and between-group regression residuals.

Maximum likelihood

The log-likelihood function is:

$$L(\alpha_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma_\varepsilon^2, \sigma_u^2) = const - \frac{1}{2}\sum_{i=1}^{n} \ln \det \boldsymbol{\Omega}_i - \frac{1}{2}\sum_{i=1}^{n} \mathbf{v}_i' \boldsymbol{\Omega}_i^{-1} \mathbf{v}_i$$

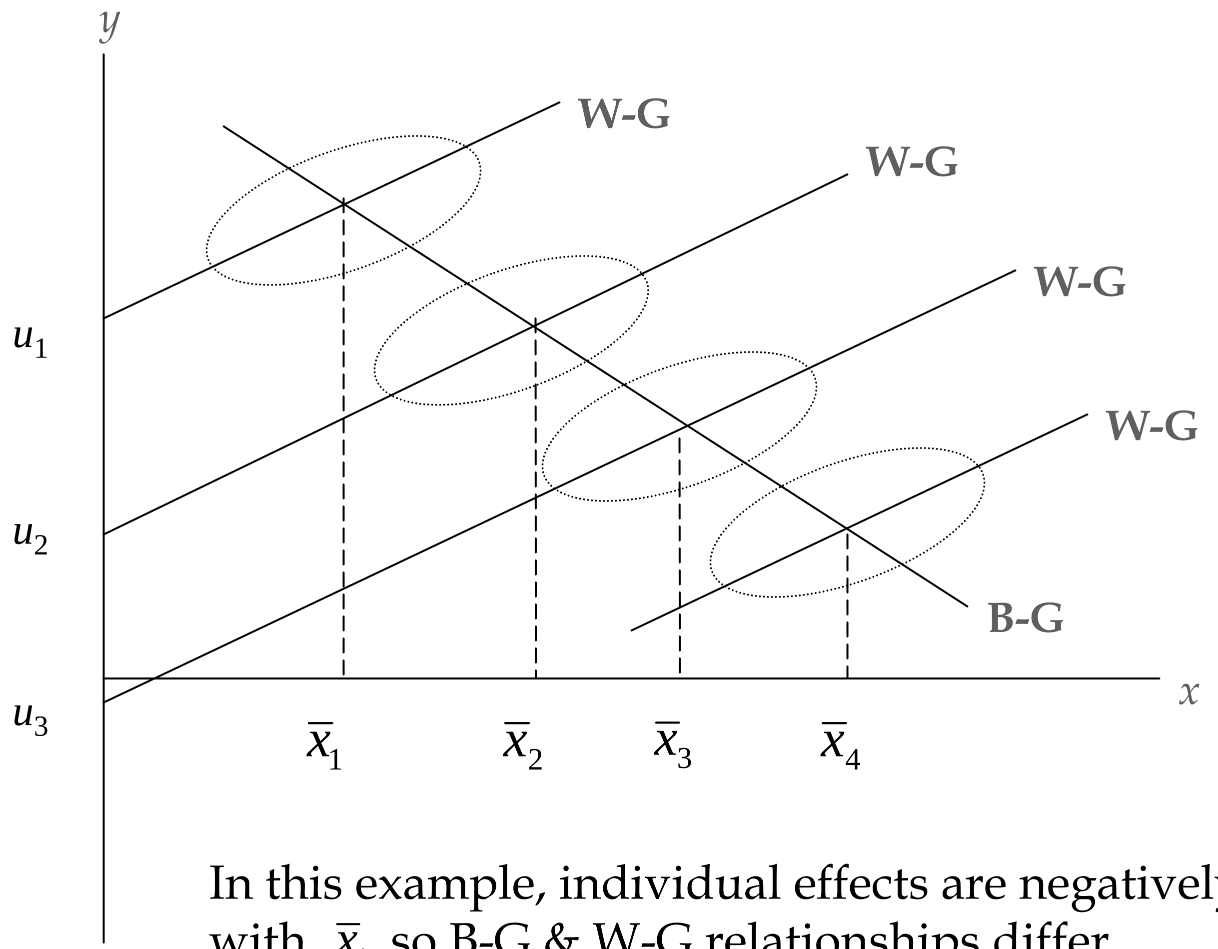
This can be maximised numerically to estimate all parameters simultaneously

ML and feasible GLS are asymptotically equivalent as $n \to \infty$, with each T_i fixed.

In Stata, the command *xtreg* has various options:

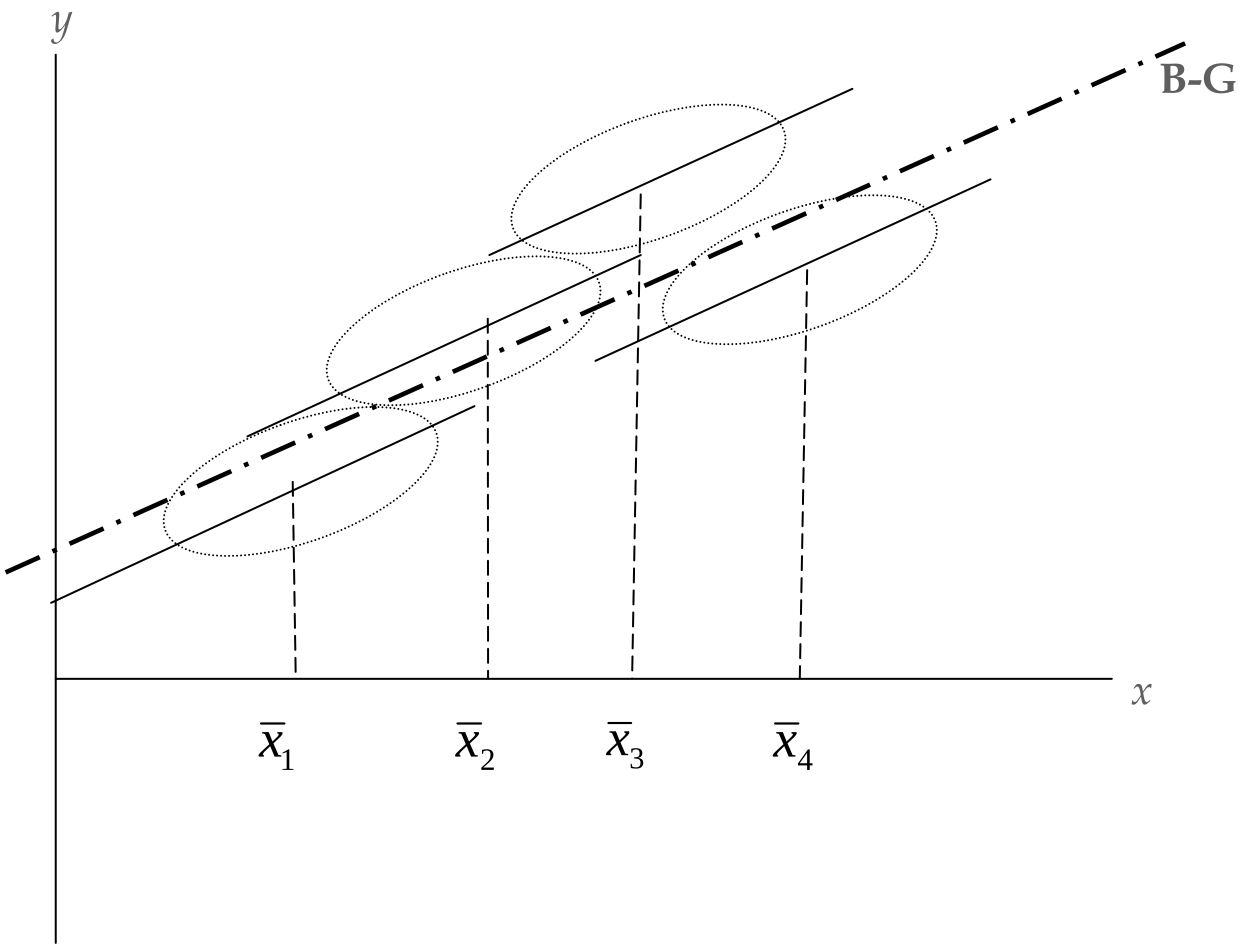
,fe for within-group
,be for between-group
,re for random effects (feasible GLS)
,mle for random effects (ML)

Within- & between-group relationships:
correlated individual effects



In this example, individual effects are negatively correlated with $\overline{x}_i$, so B-G & W-G relationships differ

Within- & between-group relationships:
uncorrelated individual effects



Testing the hypothesis of uncorrelated effects

The random effects estimator (and any estimator that uses between-group variation) is only consistent as $n \to \infty$ if the following hypothesis is true:

$$H_0\text{:}\quad \mathrm{E}(u_i \mid \mathbf{z}_i, \mathbf{X}_i) = 0$$
$$H_1\text{:}\quad \mathrm{E}(u_i \mid \mathbf{z}_i, \mathbf{X}_i) \neq 0$$

It is important to test H_0. There are many equivalent ways of doing so:

(1) Hausman parameter contrast test:

$$\left(\hat{\boldsymbol{\beta}}_W - \hat{\boldsymbol{\beta}}_{GLS}\right)'\left[\mathrm{cov}\left(\hat{\boldsymbol{\beta}}_W\right) - \mathrm{cov}\left(\hat{\boldsymbol{\beta}}_{GLS}\right)\right]^{-1}\left(\hat{\boldsymbol{\beta}}_W - \hat{\boldsymbol{\beta}}_{GLS}\right)$$
$$\sim \chi^2(k_x) \text{ under } H_0$$

(2) Mundlak approach: estimate the model

$$y_{it} = \alpha_0 + \mathbf{z}_i\boldsymbol{\alpha} + \mathbf{x}_{it}\boldsymbol{\beta} + \bar{\mathbf{x}}_i\boldsymbol{\gamma} + u_i + \varepsilon_{it}$$

by GLS and test H_0: $\boldsymbol{\gamma} = \mathbf{0}$. (NB: $\hat{\boldsymbol{\beta}}_{GLS} \equiv \hat{\boldsymbol{\beta}}_W$ in this case)

Example: BHPS feasible GLS RE model

```
. xtreg logearn age cohort, re

Random-effects GLS regression                   Number of obs      =     21124
Group variable (i): pid                         Number of groups   =      5859

R-sq:  within  = 0.1255                         Obs per group: min =         1
       between = 0.0011                                        avg =       3.6
       overall = 0.0131                                        max =        11

Random effects u_i ~ Gaussian                   Wald chi2(2)       =   2109.10
corr(u_i, X)       = 0 (assumed)                Prob > chi2        =    0.0000

------------------------------------------------------------------------------
     logearn |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0288609   .0006306    45.77   0.000      .027625    .0300969
      cohort |   .0226111    .000841    26.88   0.000     .0209627    .0242595
       _cons |    -43.445   1.664017   -26.11   0.000    -46.70641   -40.18359
-------------+----------------------------------------------------------------
     sigma_u |  .52813738
     sigma_e |  .24397993
         rho |  .82412387   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Example: BHPS Hausman test

```
. hausman within re

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     within         re         Difference          S.E.
-------------+----------------------------------------------------------------
         age |    .0302855     .0288609        .0014245        .0001446
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(1) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       97.05
                Prob>chi2 =      0.0000
```