5.5 Random effects estimates

The within-group estimator requires very few assumptions and is often a powerful method for dealing with unobserved variables which may be correlated with the observed variables. It does, however, have several drawbacks. First, one cannot estimate the parameters on *any* fixed effect, whether observed or not. Second, it is difficult to make out of sample predictions because the estimates are conditional on a set of fixed characteristics α_i. Third, estimates of the α_i are not consistent unless $T \to \infty$.

An alternative approach treats differences between individuals as part of the error term. Instead of Equation (1), we write

$$y_{it} = \alpha + \mathbf{x}'_{it}\boldsymbol{\beta} + \nu_i + u_{it} \tag{10}$$

The error term consists of $(\nu_i + u_{it})$. One of the standard OLS assumptions is now violated, because the error terms will be correlated within groups. To estimate this model we therefore use a *generalised least squares model* known as the random effects estimator. This estimator is equivalent to estimating

$$y_{it} - \theta\bar{y} = (1-\alpha)\theta + (\mathbf{x}_{it} - \theta\bar{\mathbf{x}}_i)\boldsymbol{\beta} + ((1-\theta)\nu_i + (u_{it} - \theta\bar{u}_i)) \tag{11}$$

where θ is a function of σ_ν^2 and σ_u^2. If the variance of the ν_i is zero, then $\theta = 0$ and the model gives the same results as OLS.

We can estimate this model using the `xtreg` with the `re` option:

```
xtreg lnpay d2 female age training, re
```

```
Random-effects GLS regression              Number of obs      =       9621
Group variable (i): pid                    Number of groups   =       5780

R-sq:  within  = 0.0500                     Obs per group: min =          1
       between = 0.2450                                    avg =        1.7
       overall = 0.2404                                    max =          2

Random effects u_i ~ Gaussian              Wald chi2(4)       =    2073.61
corr(u_i, X)       = 0 (assumed)           Prob > chi2        =     0.0000

------------------------------------------------------------------------------
      lnpay |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         d2 |   .0783663   .0054965    14.26   0.000     .0675933    .0891394
     female |  -.7464051   .0201466   -37.05   0.000    -.7858917   -.7069186
   training |   .1124054   .0094395    11.91   0.000     .0939043    .1309065
        age |   .0159121   .0008505    18.71   0.000     .0142452    .0175791
      _cons |   6.264126   .0336539   186.13   0.000     6.198165    6.330086
-------------+----------------------------------------------------------------
    sigma_u |  .72260848
    sigma_e |  .23352156
        rho |  .90543999   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```