8. Panel data II: random effects and clustered data

LPO 8852: Regression II

Sean P. Corcoran

Panel data: fixed effects models

In Lecture 7, we used panel data to address omitted variables bias due to unobserved heterogeneity (u_i):

$$y_{it} = \beta_0 + \beta_1 x_{it} + u_i + e_{it}$$

i is a group or individual with multiple observations t, and $Cov(x_{it}, u_i) \neq 0$. (NOTE: switching notation here—u_i was c_i in FE lecture)

Estimation methods:

- Fixed effects "within" regression (LSDV; `xtreg, fe`; or `areg`)
- First-difference or long-difference

Key assumption: *strict exogeneity*, no within- or cross-period correlation between e_{it} and x_{it}.

Panel data: fixed effects models

Advantages:

- Unobserved u_i can be correlated with the explanatory variables
- β_1 is estimated using *within-group* (i) variation in x, y

Disadvantages:

- Cannot estimate slope coefficients for time-invariant x
- Fixed effects "remove" a lot of the variation in y
- The "within" model is less efficient (higher standard errors)
- There may be more measurement error (and attenuation bias) when relying on within-group *changes* vs. levels
- Group intercepts use up a lot of degrees of freedom

Random effects

The fixed effects model allows u_i to be correlated with x_{it}. An alternative conception of u_i is as a *random* effect, uncorrelated with x_{it}.

$$y_{it} = \beta_0 + \beta_1 x_{it} + \underbrace{u_i + e_{it}}_{v_{it}}$$

Think of v_{it} as a *composite* error consisting of a between-group component (u_i) common to all observations within the group and a within-group component (e_{it}). It is assumed u_i and e_{it} are independent of one another and:

$$u_i \sim N(0, \sigma_u^2)$$
$$e_{it} \sim N(0, \sigma_e^2)$$

Sometimes called a "random intercepts" model.

Random effects

If u_i is uncorrelated with x_{it}, then the composite error term v_{it} is uncorrelated with x_{it}. (We already assumed e_{it} is uncorrelated with x_{it}). This means the OLS estimator for β_1 will be unbiased and consistent.

Note: estimation of this model does *not* involve estimating the u_i's as parameters as in the LSDV model.

Random effects

The composite error term v_{it} is not, however, i.i.d.:

$$Corr(v_{it}, v_{is}) = \rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2} \text{ for } s \neq t$$

The common error for observations in group i (u_i) results in correlation between the composite error in period t (v_{it}) and in period s (v_{is}).

Random effects

The composite error term v_{it} is not, however, i.i.d.:

$$Corr(v_{it}, v_{is}) = \rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2} \text{ for } s \neq t$$

The common error for observations in group i (u_i) results in correlation between the composite error in period t (v_{it}) and in period s (v_{is}).

This means OLS is consistent but not efficient, and that traditional standard error formulas assuming i.i.d. errors are incorrect. The ratio above (ρ) is called the **intra-class correlation** (more on this later).

Estimation using GLS (details later): `xtreg, re`.

Success for All example

- Success for All is a whole-school literacy intervention.
- Borman et al. (2005) conducted a randomized evaluation of SFA in 2001-02 and 2002-03 (21 treatment schools and 20 control).
- A *cluster-randomized* design with randomization at the school level.
- The data used by Murnane & Willett (*ch7_sfa.dta*) include grade 1 only. The outcome of interest is *wattack*, the student's score on a "Word-Attack" test.

Next slide: an "unconditional" model with no x_{it} estimates variance components σ_u^2 and σ_e^2 and the intra-class correlation ρ.

Random effects with `xtreg`

```
. xtreg wattack, re i(schid)

Random-effects GLS regression                   Number of obs     =      2,334
Group variable: schid                           Number of groups  =         41

R-sq:                                           Obs per group:
     within  = 0.0000                                         min =         10
     between = 0.0000                                         avg =       56.9
     overall = 0.0000                                         max =        134

                                                Wald chi2(0)      =          .
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =          .

------------------------------------------------------------------------------
     wattack |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   477.5356   1.447118   329.99   0.000     474.6994    480.3719
-------------+----------------------------------------------------------------
     sigma_u |  8.8705267
     sigma_e |  17.725757
         rho |  .20027618   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

This example: Success for All impact evaluation (from Murnane & Willett). $\sigma_u^2 = 8.87^2 = 78.7$ and $\sigma_e^2 = 17.73^2 = 314.35$. $\rho = 0.200$.

loneway

`loneway` (one-way ANOVA) is another handy command for estimating variance components and ICC. (Note the difference in σ_u and ρ from `xtreg, re`. With unbalanced panels, these will differ slightly).

```
. loneway wattack schid

        One-way Analysis of Variance for wattack: word attack posttest

                                          Number of obs =        2,334
                                              R-squared =       0.2185

    Source                SS         df      MS            F     Prob > F
-------------------------------------------------------------------------
Between schid         201450.43      40    5036.2607     16.03     0.0000
Within schid          720466.21   2,293    314.20244
-------------------------------------------------------------------------
Total                 921916.63   2,333    395.16358

         Intraclass       Asy.
         correlation      S.E.       [95% Conf. Interval]
         ------------------------------------------------
            0.20993     0.04402       0.12366     0.29621

         Estimated SD of schid effect              9.137203
         Estimated SD within schid                 17.72576
         Est. reliability of a schid mean           0.93761
              (evaluated at n=56.56)
```

Random effects with `xtreg`

```
. xtreg wattack sfa ppvt, re i(schid)

Random-effects GLS regression                   Number of obs     =      2,334
Group variable: schid                           Number of groups  =         41

R-sq:                                           Obs per group:
     within  = 0.1101                                         min =         10
     between = 0.3960                                         avg =       56.9
     overall = 0.1820                                         max =        134

                                                Wald chi2(2)      =     308.21
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000

------------------------------------------------------------------------------
     wattack |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         sfa |   3.440921   2.297268     1.50   0.134    -1.061642    7.943485
        ppvt |   .4851754   .0278075    17.45   0.000     .4306737    .5396771
       _cons |   432.0475   2.972263   145.36   0.000      426.222     437.873
-------------+----------------------------------------------------------------
     sigma_u |  6.9082397
     sigma_e |  16.725172
         rho |  .14574141   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

This regression: includes the treatment indicator (*sfa*) and one covariate (*ppvt*). Note changes in σ_u and σ_e, ρ. The residual variability is reduced with the inclusion of x's.

Random effects

Class size and passing rates in TX (see previous panel data lecture):

```
. xtreg avgpassing avgclass, re i(campus)

Random-effects GLS regression                   Number of obs     =     16,062
Group variable: campus                          Number of groups  =      4,326

R-sq:                                           Obs per group:
     within  = 0.0018                                         min =          1
     between = 0.0098                                         avg =        3.7
     overall = 0.0060                                         max =          4

                                                Wald chi2(1)      =       2.74
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0978

------------------------------------------------------------------------------
  avgpassing |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    avgclass |  -.0442893   .0267548    -1.66   0.098    -.0967277    .0081491
       _cons |   76.21828   .5503649   138.49   0.000     75.13959    77.29698
-------------+----------------------------------------------------------------
     sigma_u |  12.391941
     sigma_e |  6.4870883
         rho |  .78490199   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Random effects

Compare to fixed effects: very different slope coefficient estimate.

```
. xtreg avgpassing avgclass, fe i(campus)

Fixed-effects (within) regression               Number of obs     =     16,062
Group variable: campus                          Number of groups  =      4,326

R-sq:                                           Obs per group:
     within  = 0.0018                                         min =          1
     between = 0.0098                                         avg =        3.7
     overall = 0.0060                                         max =          4

                                                F(1,11735)        =      21.30
corr(u_i, Xb)  = -0.1189                        Prob > F          =     0.0000

------------------------------------------------------------------------------
  avgpassing |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    avgclass |  -.1339024   .0290105    -4.62   0.000    -.1907678   -.0770371
       _cons |   78.09211   .5590819   139.68   0.000     76.99621      79.188
-------------+----------------------------------------------------------------
     sigma_u |  12.997022
     sigma_e |  6.4870883
         rho |  .80056238   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(4325, 11735) = 13.83                Prob > F = 0.0000
```

Random vs. fixed effects

- The RE model is biased and inconsistent if the FE assumptions are more appropriate (correlation between x_{it} and u_i).
- If the RE assumptions hold (no correlation between x_{it} and u_i), both RE and FE are *consistent*. They should give "similar" answers in large samples, but the FE model will be *inefficient* (larger standard errors).
- A sufficiently large difference in point estimates suggests the FE assumption is probably correct and RE is inconsistent.
- The **Hausman test** is a formal test of this.

Hausman test

First use `estimates store` to save your `fe` and `re` estimates. Name them FE and RE, for example.

```
xtreg avgpassing avglcass, fe i(campus)
estimates store FE
xtreg avgpassing avgclass, re i(campus)
estimates store RE
hausman FE RE
```

Hausman test

Null hypothesis: RE assumptions hold, both estimators consistent but RE is efficient. Alternative: RE assumptions do *not* hold and the RE estimator is inconsistent. In the TX example we can reject H_0:

```
. hausman FE RE

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |       FE           RE         Difference          S.E.
-------------+----------------------------------------------------------------
    avgclass |   -.1339024    -.0442893       -.0896131        .0112156
------------------------------------------------------------------------------
                         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)
                          =       63.84
                Prob>chi2 =      0.0000
```

Review of GLS

In a linear regression with known heteroskedasticity, we can transform the original data and apply OLS to the transformed data. E.g.:

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

with $\text{Var}(u_i) = k_i \sigma_u^2$. The GLS transformation divides the data by $\sqrt{k_i}$. Observations with greater variance get *less* weight. The transformed model satisfies homoskedasticity.

GLS estimation of random effects models

The random effects model with one covariate is:

$$y_{it} = \beta_0 + \beta_1 x_{it} + \underbrace{u_i + e_{it}}_{v_{it}}$$

GLS estimation again involves a transformation. Let:

$$\theta = 1 - \sqrt{\frac{\sigma_e^2}{\sigma_e^2 + T\sigma_u^2}}$$

(and note the term under the square root looks like but is different from the ICC). T is the number of observations per group, assuming a balanced panel.

GLS estimation of random effects models

The transformations of y_{it} and x_{it} are:

$$y_{it} - \theta \bar{y}_i$$

$$x_{it} - \theta \bar{x}_i$$

and OLS is estimated on the transformed model:

$$y_{it} - \theta \bar{y}_i = \beta_0(1-\theta) + \beta_1(x_{it} - \theta \bar{x}_i) + (v_{it} - \theta \bar{v}_i)$$

GLS estimation of random effects models

The transformations of y_{it} and x_{it} are:

$$y_{it} - \theta \bar{y}_i$$

$$x_{it} - \theta \bar{x}_i$$

and OLS is estimated on the transformed model:

$$y_{it} - \theta \bar{y}_i = \beta_0(1-\theta) + \beta_1(x_{it} - \theta \bar{x}_i) + (v_{it} - \theta \bar{v}_i)$$

The transformed y_{it} and x_{it} are *quasi-demeaned*. If $\theta = 1$, we have the demeaned (within) model.

GLS estimation of random effects models

θ is not known so it must first be estimated with consistent estimators for σ_e^2 and σ_u^2. Then, $\hat{\theta}$ is used in OLS estimation ("feasible GLS").

$$\hat{\theta} = 1 - \sqrt{\frac{\hat{\sigma}_e^2}{\hat{\sigma}_e^2 + T\hat{\sigma}_u^2}}$$

Consistent estimators for σ_u^2 and σ_e^2 can be obtained using pooled OLS or fixed effects residuals.

GLS estimation of random effects models

One method for estimating σ_u^2 and σ_e^2: note that

$$v_{it} = u_i + e_{it}$$

$$v_{it} v_{is} = (u_i + e_{it})(u_i + e_{is})$$

$$E(v_{it} v_{is}) = \underbrace{E(u_i^2)}_{\sigma_u^2} + \underbrace{E(u_i e_{is})}_{0} + \underbrace{E(u_i e_{it})}_{0} + \underbrace{E(e_{it} e_{is})}_{0}$$

Get the composite residuals $\hat{v}_{it}$ using pooled OLS. The square of the RMSE in this regression estimates σ_v^2. The within-group covariance in $\hat{v}_{it}$ (the sample analog of $E(v_{it} v_{is})$ above) provides a consistent estimate of σ_u^2. Then, $\hat{\sigma}_e^2 = \hat{\sigma}_v^2 - \hat{\sigma}_u^2$. See problem set.

GLS estimation of random effects models

$$y_{it} - \theta\bar{y}_i = \beta_0(1-\theta) + \beta_1(x_{it} - \theta\bar{x}_i) + (v_{it} - \theta\bar{v}_i)$$

$$\theta = 1 - \sqrt{\frac{\sigma_e^2}{\sigma_e^2 + T\sigma_u^2}}$$

Notice the transformation subtracts a *fraction* of the within-group mean, where the fraction depends on σ_e^2, σ_u^2, and T.

- When $\theta = 0$, the model reduces to pooled OLS
- When $\theta = 1$, the model reduces to fixed effects (within)
- So, the value of θ is indicative of which model RE is closer to

GLS estimation of random effects models

$$y_{it} - \theta\bar{y}_i = \beta_0(1-\theta) + \beta_1(x_{it} - \theta\bar{x}_i) + (v_{it} - \theta\bar{v}_i)$$

$$\theta = 1 - \sqrt{\frac{\sigma_e^2}{\sigma_e^2 + T\sigma_u^2}}$$

Notice the transformation subtracts a *fraction* of the within-group mean, where the fraction depends on σ_e^2, σ_u^2, and T.

- When $\theta = 0$, the model reduces to pooled OLS
- When $\theta = 1$, the model reduces to fixed effects (within)
- So, the value of θ is indicative of which model RE is closer to

θ gets closer to 1 as between-group variation σ_u^2 grows relative to within-group variation σ_e^2, and as the number of time periods T grows.

GLS estimation of random effects models

Can request $\hat{\theta}$ in `xtreg, re`:

```
. xtreg avgpassing avgclass, re i(campus) theta

Random-effects GLS regression                   Number of obs      =     16,062
Group variable: campus                          Number of groups   =      4,326

R-sq:                                           Obs per group:
     within  = 0.0018                                         min =          1
     between = 0.0098                                         avg =        3.7
     overall = 0.0060                                         max =          4

                                                Wald chi2(1)       =       2.74
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =     0.0978

------------------- theta --------------------
  min      5%       median        95%      max
0.5362   0.6529     0.7468      0.7468   0.7468

------------------------------------------------------------------------------
  avgpassing |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    avgclass |  -.0442893   .0267548    -1.66   0.098    -.0967277    .0081491
       _cons |   76.21828   .5503649   138.49   0.000     75.13959    77.29698
-------------+----------------------------------------------------------------
     sigma_u |  12.391941
     sigma_e |  6.4870883
         rho |  .78490199   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

This uses the original unbalanced panel, so $\hat{\theta}$ varies with group size.

GLS estimation of random effects models

Can request $\hat{\theta}$ in `xtreg, re`:

```
. xtreg avgpassing avgclass, re i(campus) theta

Random-effects GLS regression                   Number of obs     =     14,796
Group variable: campus                          Number of groups  =      3,699

R-sq:                                           Obs per group:
     within  = 0.0020                                         min =          4
     between = 0.0138                                         avg =        4.0
     overall = 0.0061                                         max =          4

                                                Wald chi2(1)      =       2.97
corr(u_i, X)     = 0 (assumed)                  Prob > chi2       =     0.0848
theta            = .73287384

------------------------------------------------------------------------------
  avgpassing |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    avgclass |  -.0484254   .0280999    -1.72   0.085    -.1035003    .0066494
       _cons |   76.51251   .5742248   133.24   0.000     75.38705    77.63797
-------------+----------------------------------------------------------------
     sigma_u |  11.706021
     sigma_e |  6.4897977
         rho |  .76490175   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

This uses the balanced panel, so $\hat{\theta}$ is constant.

GLS estimation of random effects models

It is useful to consider the error term in the quasi-demeaned model:

$$v_{it} - \theta \bar{v}_i = (1-\theta)u_i + (e_{it} - \theta \bar{e}_i)$$

Suppose the RE assumption that u_i is uncorrelated with x_{it} does *not* hold. As $\theta \rightarrow 1$, the u_i component of the error term diminishes in importance, the RE estimator tends toward the FE estimator, and any bias associated with RE tends to zero.

MLE estimation of random effects models

Random effects models can also be estimated using maximum likelihood in which case all parameters of the model (β's, σ's) are estimated jointly:

```
. xtreg avgpassing avgclass, mle i(campus)

Fitting constant-only model:
Iteration 0:   log likelihood = -53584.523
Iteration 1:   log likelihood = -53584.523

Fitting full model:
Iteration 0:   log likelihood = -53674.187
Iteration 1:   log likelihood = -53583.763
Iteration 2:   log likelihood = -53582.969
Iteration 3:   log likelihood = -53582.969

Random-effects ML regression                    Number of obs     =     14,796
Group variable: campus                          Number of groups  =      3,699

Random effects u_i ~ Gaussian                   Obs per group:
                                                              min =          4
                                                              avg =        4.0
                                                              max =          4

                                                LR chi2(1)        =       3.11
Log likelihood  = -53582.969                    Prob > chi2       =     0.0780

------------------------------------------------------------------------------
  avgpassing |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    avgclass |  -.0496391   .0281539    -1.76   0.078    -.1048197    .0055415
       _cons |   76.53576   .5755876   132.97   0.000     75.40763    77.66389
-------------+----------------------------------------------------------------
    /sigma_u |    11.8066   .1481004                      11.51987    12.10047
    /sigma_e |   6.492198   .0436102                      6.407283    6.578237
         rho |   .7678329   .0051631                      .7575916    .7778289
------------------------------------------------------------------------------
LR test of sigma_u=0: chibar2(01) = 1.2e+04            Prob >= chibar2 = 0.000
```

Getting estimates of u_i

As with `xtreg, fe`, one can obtain the $\hat{u}_i$ estimates of the group random effects. Unlike `fe`, these are not coefficient estimates but rather estimated from residuals. The random effects $\hat{u}_i$ can be calculated in two ways:

- Maximum likelihood (following `xtreg, mle`)
- Empirical Bayes / shrinkage approach: the Best Linear Unbiased Predictors (BLUPs)

Shrinkage approach: multiply $\hat{u}_i$ by a shrinkage factor $\hat{R}_i = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \frac{\hat{\sigma}_e^2}{T_i}}$

where T_i is the number of observations in group i. Examples on next 3 slides.

Getting estimates of u_i: MLE

```
. xtreg avgpassing avgclass, re mle i(campus)

Fitting constant-only model:
Iteration 0:   log likelihood = -53584.523
Iteration 1:   log likelihood = -53584.523

Fitting full model:
Iteration 0:   log likelihood = -53674.187
Iteration 1:   log likelihood = -53583.763
Iteration 2:   log likelihood = -53582.969
Iteration 3:   log likelihood = -53582.969

Random-effects ML regression                    Number of obs     =     14,796
Group variable: campus                          Number of groups  =      3,699

Random effects u_i ~ Gaussian                   Obs per group:
                                                              min =          4
                                                              avg =        4.0
                                                              max =          4

                                                LR chi2(1)        =       3.11
Log likelihood  = -53582.969                    Prob > chi2       =     0.0780

------------------------------------------------------------------------------
  avgpassing |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    avgclass |  -.0496391   .0281539    -1.76   0.078    -.1048197    .0055415
       _cons |   76.53576   .5755876   132.97   0.000     75.40763    77.66389
-------------+----------------------------------------------------------------
    /sigma_u |    11.8066   .1481004                      11.51987    12.10047
    /sigma_e |   6.492198   .0436102                      6.407283    6.578237
         rho |   .7678329   .0051631                      .7575916    .7778289
------------------------------------------------------------------------------
LR test of sigma_u=0: chibar2(01) = 1.2e+04           Prob >= chibar2 = 0.000

. predict uhat1, u

. sum uhat1

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       uhat1 |     14,796    8.39e-09    12.24512  -47.43509   23.42125
```

Getting estimates of u_i: BLUP

```
. xtreg avgpassing avgclass, re mle i(campus)

Fitting constant-only model:
Iteration 0:   log likelihood = -53584.523
Iteration 1:   log likelihood = -53584.523

Fitting full model:
Iteration 0:   log likelihood = -53674.187
Iteration 1:   log likelihood = -53583.763
Iteration 2:   log likelihood = -53582.969
Iteration 3:   log likelihood = -53582.969

Random-effects ML regression                    Number of obs      =     14,796
Group variable: campus                          Number of groups   =      3,699

Random effects u_i ~ Gaussian                   Obs per group:
                                                             min =          4
                                                             avg =        4.0
                                                             max =          4

                                                LR chi2(1)         =       3.11
Log likelihood  = -53582.969                    Prob > chi2        =     0.0780

------------------------------------------------------------------------------
  avgpassing |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    avgclass |  -.0496391   .0281539    -1.76   0.078    -.1048197    .0055415
       _cons |   76.53576   .5755876   132.97   0.000     75.40763    77.66389
-------------+----------------------------------------------------------------
    /sigma_u |    11.8066   .1481004                      11.51987    12.10047
    /sigma_e |   6.492198   .0436102                      6.407283    6.578237
         rho |   .7678329   .0051631                      .7575916    .7778289
------------------------------------------------------------------------------
LR test of sigma_u=0: chibar2(01) = 1.2e+04           Prob >= chibar2 = 0.000

. gen shrink = _b[/sigma_u]^2 / (_b[/sigma_u]^2 + (_b[/sigma_e]^2)/4)

. gen uhat1s = uhat1*shrink

. summ uhat1s shrink

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      uhat1s |     14,796    1.16e-08    11.38455  -44.10139   21.77522
      shrink |     14,796    .9297209           0   .9297209   .9297209
```

Getting estimates of u_i: BLUP using `xtmixed`

```
. xtmixed avgpassing avgclass || campus: , mle

Performing EM optimization:

Performing gradient-based optimization:

Iteration 0:   log likelihood = -53582.969
Iteration 1:   log likelihood = -53582.969

Computing standard errors:

Mixed-effects ML regression                     Number of obs      =     14,796
Group variable: campus                          Number of groups   =      3,699

                                                Obs per group:
                                                              min =          4
                                                              avg =        4.0
                                                              max =          4

                                                Wald chi2(1)       =       3.13
Log likelihood = -53582.969                     Prob > chi2        =     0.0770

------------------------------------------------------------------------------
  avgpassing |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    avgclass |  -.0496392   .0280727    -1.77   0.077    -.1046606    .0053823
       _cons |   76.53576   .5741313   133.31   0.000     75.41048    77.66103
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
campus: Identity             |
                   sd(_cons) |    11.8066   .1481006      11.51987    12.10047
-----------------------------+------------------------------------------------
                sd(Residual) |   6.492197   .0436102      6.407283    6.578236
------------------------------------------------------------------------------
LR test vs. linear model: chibar2(01) = 11666.05      Prob >= chibar2 = 0.0000

. predict uhat2, reffects

. sum uhat2

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       uhat2 |     14,796   -6.21e-10    11.38455  -44.10139   21.77523
```

Getting estimates of u_i

The shrinkage factor is smaller for groups with fewer observations (T_i). Their $\hat{u}_i$ is "shrunk" more toward the overall mean group effect of 0.

- RE estimates generally smaller than FE estimates in absolute value
- True for both MLE and EB estimates of the RE, but especially the EB
- The rank order of the $\hat{u}_i$ is usually preserved whether one assumes RE or FE

Random vs. fixed effects

When and where random effects are appropriate:

- As a rule, if the FE assumption holds the RE model is inappropriate. See the Texas class size example, where the Hausman test rejected RE.
- RE is appropriate with grouped or clustered data. See the Success for All example: assignment to treatment was random at the school level, so we need not be concerned about correlation between treatment and the error term. However, the errors are not i.i.d.

See Rabe-Hesketh and Skrondal MLM text for more guidance on RE vs. FE decision.

xttest0

The command `xttest0` (following `xtreg`) provides a formal test for the presence of random effects. H_0 in this case is that the variance across panel units is zero, and thus RE is unnecessary.

```
. xttest0

Breusch and Pagan Lagrangian multiplier test for random effects

        wattack[schid,t] = Xb + u[schid] + e[schid,t]

        Estimated results:
                         |       Var     sd = sqrt(Var)
                ---------+-----------------------------
                 wattack |   395.1636       19.87872
                       e |   279.7314       16.72517
                       u |   47.72378        6.90824

        Test:   Var(u) = 0
                             chibar2(01) =  1266.18
                          Prob > chibar2 =   0.0000
```