

## Using STATA for mixed-effects models (i.e. hierarchical linear model)

The **XTMIXED** function is for Multilevel mixed-effects linear regressions

From the help file for xtmixed:

Remarks on specifying random-effects equations

Mixed models consist of fixed effects and random effects. The fixed effects are specified as regression parameters in a manner similar to most other Stata estimation commands, that is, as a dependent variable followed by a set of regressors. The random-effects portion of the model is specified by first considering the grouping structure of the data. For example, if random effects are to vary according to variable school, then the call to xtmixed would be of the form

```
. xtmixed fixed_portion || school: ... , options
```

The variable lists that make up each equation describe how the random effects enter into the model, either as random intercepts (constant term) or as random coefficients on regressors in the data. One may also specify the variance-covariance structure of the within-equation random effects, according to the four available structures described above. For example,

```
. xtmixed f_p || school: z1, covariance(unstructured) options
```

will fit a model with a random intercept and random slope for variable z1 and treat the variance-covariance structure of these two random effects as unstructured.

If the data are organized by a series of nested groups, for example, classes within schools, then the random-effects structure is specified by a series of equations, each separated by ||. The order of nesting proceeds from left to right. For our example, this would mean that an equation for schools would be specified first, followed by an equation for classes. As an example, consider

```
. xtmixed f_p || school: z1, cov(un) || class: z1 z2 z3, nocons cov(ex) options
```

where variables school and class identify the schools and classes within schools, respectively. This model contains a random intercept and random coefficient on z1 at the school level and has random coefficients on variables z1, z2, and z3 at the class level. The covariance structure for the random effects at the class level is exchangeable, meaning that the random effects share a common variance and common pairwise covariance.

Group variables may be repeated, allowing for more general covariance structures to be constructed as block-diagonal matrices based on the four original structures. Consider

```
. xtmixed f_p || school: z1 z2, nocons cov(id) || school: z3 z4, nocons cov(un) options
```

which specifies four random coefficients at the school level. The variance-covariance matrix of the random effects is the 4 x 4 matrix where the upper 2 x 2 diagonal block is a multiple of the identity matrix and the lower 2 x 2 diagonal block is unstructured. In effect, the coefficients on z1 and z2 are constrained to be independent and share a common variance. The coefficients on z3 and z4 each have a distinct variance and a variance distinct from that of the coefficients on z1 and z2. They are also allowed to be correlated, yet they are independent from the coefficients on z1 and z2.

For mixed models with no nested grouping structure, thinking of the entire estimation data as one group is convenient. Toward this end, xtmixed allows the special group designation \_all. xtmixed also allows the factor variable notation R.varname, which is shorthand for describing the levels of varname as a series of indicator variables. See [XT] xtmixed for more details.

## Using STATA for the Math Achievement data example

### 1. Unconditional Means Model (i.e. no covariates and a random school effect)

```
. xtmixed mathach ||school:
```

Performing EM optimization:

Performing gradient-based optimization:

Iteration 0: log restricted-likelihood = -23558.397

Iteration 1: log restricted-likelihood = -23558.397

Computing standard errors:

Mixed-effects REML regression

Group variable: school

Number of obs = 7185

Number of groups = 160

Obs per group: min = 14

avg = 44.9

max = 67

Wald chi2(0) = .

Prob > chi2 = .

Log restricted-likelihood = -23558.397

mathach	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
_cons	12.63697	.2443937	51.71	0.000	12.15797	13.11598

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
school: Identity				
sd(_cons)	2.934968	.1837849	2.595983	3.318217
sd(Residual)	6.256862	.0527936	6.154239	6.361196

LR test vs. linear regression: chibar2(01) = 986.12 Prob >= chibar2 = 0.0000

**NOTICE that STATA outputs standard deviations of the between and within school math achievement (in comparison to SAS which outputs variance = std<sup>2</sup>)**

### 2. Including a school level predictor

```
. xtmixed mathach meanses ||school:
```

Performing EM optimization:

Performing gradient-based optimization:

Iteration 0: log restricted-likelihood = -23480.642

Iteration 1: log restricted-likelihood = -23480.642

Computing standard errors:

Mixed-effects REML regression

Group variable: school

Number of obs = 7185

Number of groups = 160

Obs per group: min = 14

avg = 44.9

max = 67

Wald chi2(1) = 263.15

Prob > chi2 = 0.0000

Log restricted-likelihood = -23480.642

```

-----
      mathach |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      meanses |      5.863538   .361458    16.22   0.000     5.155094     6.571983
       _cons   |     12.64944   .1492801    84.74   0.000     12.35685     12.94202
-----+-----

Random-effects Parameters |   Estimate  Std. Err.      [95% Conf. Interval]
-----+-----
school: Identity
      sd(_cons) |       1.62441   .1244571     1.39791     1.887609
-----+-----
      sd(Residual) |       6.257562   .0528002     6.154926     6.361909
-----+-----
LR test vs. linear regression: chibar2(01) =   239.95 Prob >= chibar2 = 0.0000

```

### 3. Including student level predictor (Centered around school level mean)

#### a. Only allowing for a random intercept

```

. xtmixed mathach cses ||school:
Log restricted-likelihood = -23361.999          Prob > chi2          =    0.0000
-----+-----
      mathach |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
       cses   |      2.191168   .1086546    20.17   0.000     1.978209     2.404127
       _cons   |     12.64929   .2444897    51.74   0.000     12.1701     13.12848
-----+-----

Random-effects Parameters |   Estimate  Std. Err.      [95% Conf. Interval]
-----+-----
school: Identity
      sd(_cons) |       2.944894   .1832708     2.606733     3.326924
-----+-----
      sd(Residual) |       6.083618   .0513353     5.98383     6.18507
-----+-----
LR test vs. linear regression: chibar2(01) =  1055.94 Prob >= chibar2 = 0.0000

```

#### b. including random intercept and slope

```

. xtmixed mathach cses ||school: cses, variance covar(un)
Log restricted-likelihood = -23357.118          Prob > chi2          =    0.0000
-----+-----
      mathach |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
       cses   |      2.193192   .1282582    17.10   0.000     1.94181     2.444574
       _cons   |     12.64934   .2445134    51.73   0.000     12.1701     13.12858
-----+-----

Random-effects Parameters |   Estimate  Std. Err.      [95% Conf. Interval]
-----+-----
school: Unstructured
      var(cses) |       .6939724   .2807827     .3140105     1.533699
      var(_cons) |       8.681651   1.079627     6.803763    11.07785
      cov(cses,_cons) |       .0507474   .4063922    -.7457667     .8472615
-----+-----
      var(Residual) |       36.7002   .6257441    35.49403    37.94736
-----+-----
LR test vs. linear regression:      chi2(3) =  1065.70  Prob > chi2 = 0.0000
Note: LR test is conservative and provided only for reference

```

#### 4. Including Both Level 1 and Level 2 predictors with cross level interaction (i.e. level 2 variable effecting level 1 slope)

Need to create interaction terms:

```
. gen cses_sector = cses*sector
. gen cses_meanses = cses*meanses
```

```
. xtmixed mathach cses meanses sector cses_sector cses_meanses ||school: cses, variance
covar(un)
```

mathach	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
cses	2.938756	.1551034	18.95	0.000	2.634759	3.242753
meanses	5.339122	.3693012	14.46	0.000	4.615305	6.062939
sector	1.216671	.3063874	3.97	0.000	.6161622	1.817179
cses_sector	-1.642572	.2398074	-6.85	0.000	-2.112586	-1.172558
cses_meanses	1.038844	.2989198	3.48	0.001	.4529717	1.624716
_cons	12.11359	.1988085	60.93	0.000	11.72393	12.50324

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
school: Unstructured				
var(cses)	.1016223	.2134717	.0016555	6.238149
var(_cons)	2.381904	.3717541	1.754178	3.234261
cov(cses,_cons)	.1924953	.204535	-.2083859	.5933765
var(Residual)	36.72101	.6261091	35.51414	37.9689