SOCY7708: Hierarchical Linear Modeling
Instructor: Natasha Sarkisian

Class notes: Longitudinal Data Analysis using HLM

HLM can also be used to model longitudinal data where multiple observations over time are nested within one person. We will start by looking at one very common type of longitudinal analysis that works well with HLM – growth curve models. That is, HLM allows us to examine change over time as diverse across units – here, we assume that the trajectories over time, as well as the effects of any time-varying variables, are not the same across units. We will look at average effect of such variables, the extent to which there is variation around that average, and at level 2 (time-invariant) predictors that may explain that variation (cross-level interactions).

We will use an example that examines how attitudes toward deviant behavior change over time for teenagers. We will use a file called nys.dta. This file contains data for a cohort of adolescents in the National Youth Survey, ages 14 to 18. The dependent variable attit is a 9-item scale assessing attitudes favorable to deviant behavior (property damage, drug and alcohol use, stealing, etc.). The level-1 independent variables include: expo measuring exposure to deviant peers (students were asked how many of their friends engaged in the 9 deviant behaviors) and age (age in years). Level 2 include person-level variables: female, minority, and income.

```
. use http://www.sarkisian.net/socy7708/nys.dta
. reshape long attit expo, i(id) j(age)
(note: j = 14 15 16 17 18)
Data                               wide   ->   long
-----------------------------------------------------------------------------
Number of obs.                      241   ->    1205
Number of variables                  14   ->       7
j variable (5 values)                     ->   age
xij variables:
            attit14 attit15 ... attit18   ->   attit
               expo14 expo15 ... expo18   ->   expo
-----------------------------------------------------------------------------

. egen miss=rowmiss( attit expo)
. tab miss
       miss |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      1,066       88.46       88.46
          2 |        139       11.54      100.00
------------+-----------------------------------
      Total |      1,205      100.00

. drop if miss==2
(139 observations deleted)

. xtset id age, yearly
       panel variable:  id (unbalanced)
        time variable:  age, 14 to 18, but with gaps
                delta:  1 year
```

With panel datasets, data are considered strongly balanced if all the time points are the same and all cases are observed at all time points. Data are considered balanced if the cases have the same number of time values but these are not exactly the same time points. Data are unbalanced if cases are observed at different numbers of time points.

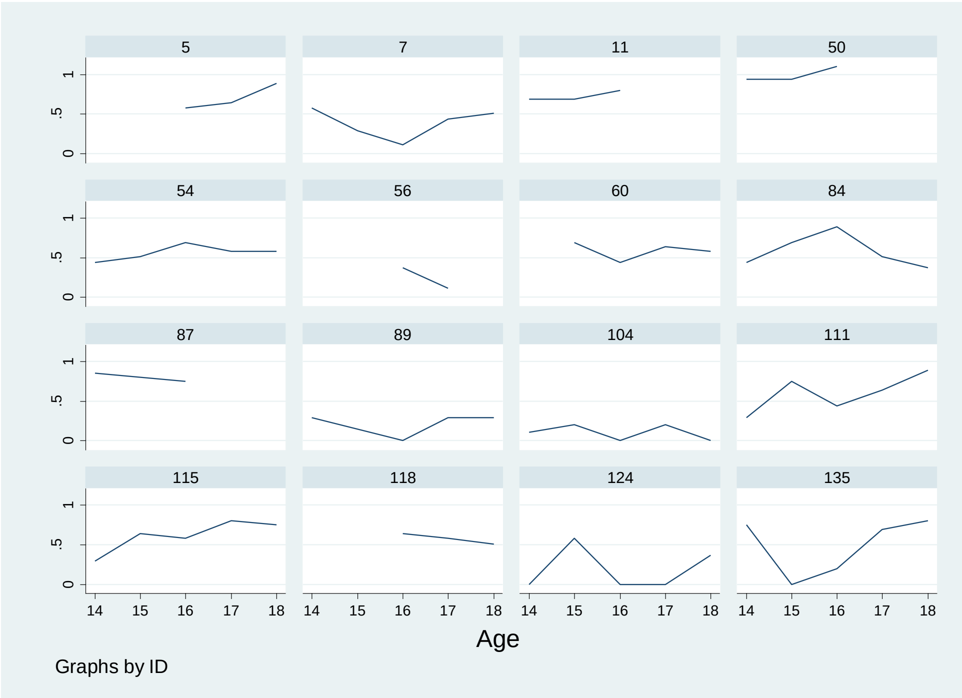
Focusing just on age, we could estimate a mixed effects model with random intercept:

```
. mixed attit age || id:

Mixed-effects ML regression                     Number of obs     =      1,066
Group variable: id                              Number of groups  =        241

                                                Obs per group:
                                                              min =          1
                                                              avg =        4.4
                                                              max =          5

                                                Wald chi2(1)      =      57.94
Log likelihood =  36.668959                     Prob > chi2       =     0.0000
------------------------------------------------------------------------------
       attit |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |    .032384   .0042543     7.61   0.000     .0240456    .0407223
       _cons |  -.0255082   .0693628    -0.37   0.713    -.1614569    .1104404
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
id: Identity                 |
                  var(_cons) |   .0443473   .0048941      .0357215    .0550558
-----------------------------+------------------------------------------------
               var(Residual) |   .0360828     .00178      .0327575    .0397458
------------------------------------------------------------------------------
LR test vs. linear model: chibar2(01) = 397.38        Prob >= chibar2 = 0.0000
```

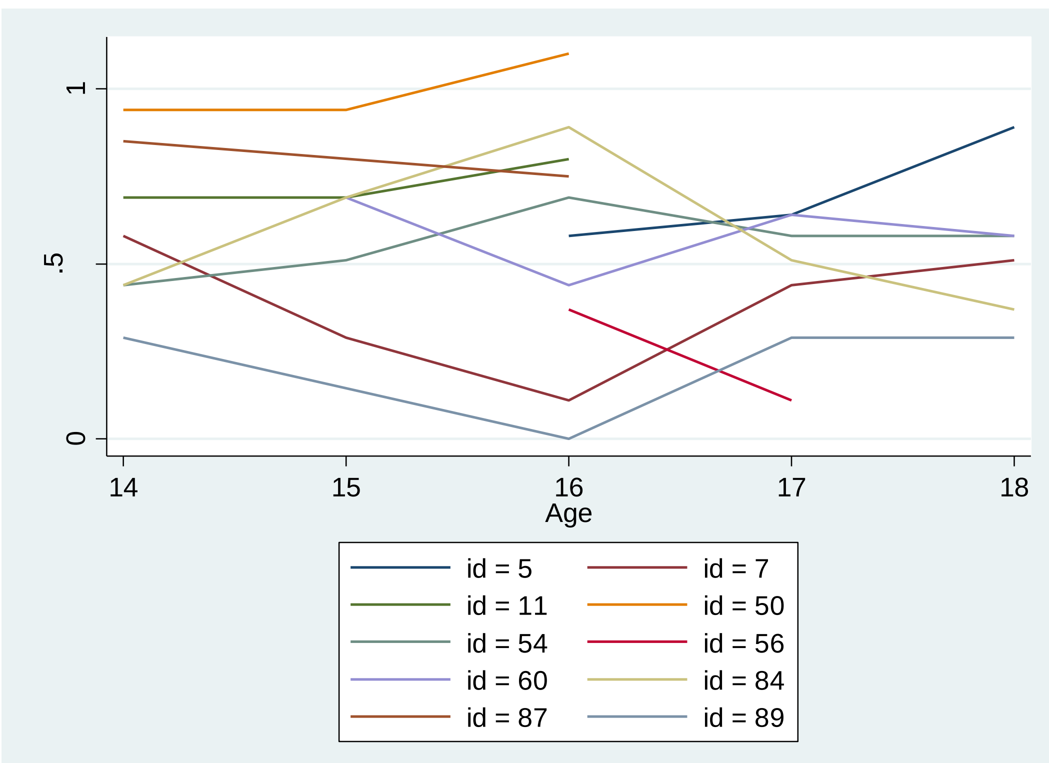
Let's examine time trends graphically:

```
. xtline attit
```

```
. xtline attit if id<100
```




```
. xtline attit if id<100, overlay
```



Very often, in this type of analysis, we are interested in understanding why and how the trajectory over time varies across units (that is why these models are also called growth curve models), so we want to explore that variation in the slope of age.

```
. mixed attit age || id: age, cov(unstructured)

Mixed-effects ML regression                     Number of obs     =      1,066
Group variable: id                              Number of groups  =        241

                                                Obs per group:
                                                              min =          1
                                                              avg =        4.4
                                                              max =          5

                                                Wald chi2(1)      =      36.73
Log likelihood =  57.442108                     Prob > chi2       =     0.0000
------------------------------------------------------------------------------
       attit |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0323534   .0053383     6.06   0.000     .0218905    .0428164
       _cons |  -.0243373   .0870451    -0.28   0.780    -.1949426    .1462679
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
id: Unstructured             |
                    var(age) |   .0031015   .0006365      .0020743    .0046372
                  var(_cons) |   .8692899   .1703095      .5921053    1.276234
              cov(age,_cons) |  -.0505552   .0103397     -.0708206   -.0302899
-----------------------------+------------------------------------------------
               var(Residual) |   .0287285   .0016527      .0256652    .0321575
------------------------------------------------------------------------------
LR test vs. linear model: chi2(3) = 438.92                Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.
```

Note that covariance value indicates how much intercepts and slopes covary: in our example, there is a negative correlation between intercepts and slopes. That is, the higher the intercept, the smaller the slope (i.e. if the starting point in terms of deviant attitudes is higher, then the slope is less steep).

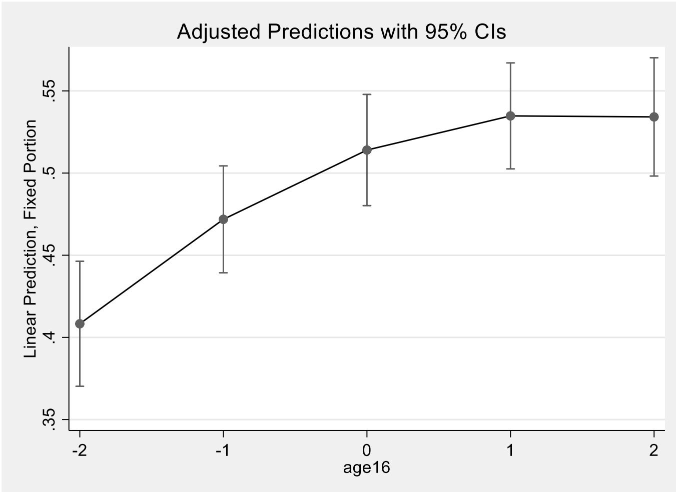
So far we assumed that the time trend is linear but the graph above shows that for many people it is not. Let’s estimate a model with a quadratic trend.

```
. tab age

        Age |      Freq.     Percent        Cum.
------------+-----------------------------------
         14 |        241       20.00       20.00
         15 |        241       20.00       40.00
         16 |        241       20.00       60.00
         17 |        241       20.00       80.00
         18 |        241       20.00      100.00
------------+-----------------------------------
      Total |      1,205      100.00

. gen age16=age-16
```

Note that the intercept will now correspond to value at age 16 rather than at the start of the study.

```
. mixed attit c.age16##c.age16 || id: c.age16##c.age16, cov(unstructured)
Mixed-effects ML regression                     Number of obs     =      1,066
Group variable: id                              Number of groups  =        241

                                                Obs per group:
                                                              min =          1
                                                              avg =        4.4
                                                              max =          5

                                                Wald chi2(2)      =      41.54
Log likelihood =  76.206955                     Prob > chi2       =     0.0000
------------------------------------------------------------------------------
       attit |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       age16 |   .0314681   .0053202     5.91   0.000     .0210407    .0418956
             |
    c.age16#|
     c.age16 |  -.0106942   .0036435    -2.94   0.003    -.0178353   -.0035532
             |
       _cons |   .5140137   .0172699    29.76   0.000     .4801654     .547862
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
id: Unstructured             |
                  var(age16) |   .0036617   .0006295      .0026143    .0051287
            var(age16#age16) |   .0011685   .0003037      .0007021    .0019447
                  var(_cons) |   .0579519    .006591      .0463722    .0724232
      cov(age16,age16#age16) |  -.0003337   .0002893     -.0009008    .0002333
            cov(age16,_cons) |  -.0003278   .0014214     -.0031136    .0024581
      cov(age16#age16,_cons) |   -.004129   .0011176     -.0063195   -.0019385
-----------------------------+------------------------------------------------
               var(Residual) |   .0229085   .0016112      .0199586    .0262943
------------------------------------------------------------------------------
LR test vs. linear model: chi2(6) = 471.08                Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.

. margins, at(age16=(-2(1)2))

Adjusted predictions                            Number of obs     =      1,066

Expression   : Linear prediction, fixed portion, predict()

1._at        : age16           =          -2
2._at        : age16           =          -1
3._at        : age16           =           0
4._at        : age16           =           1
5._at        : age16           =           2
------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         _at |
          1  |   .4083006   .0194029    21.04   0.000     .3702716    .4463297
          2  |   .4718514   .0165959    28.43   0.000     .4393239    .5043788
          3  |   .5140137   .0172699    29.76   0.000     .4801654     .547862
          4  |   .5347876   .0164525    32.50   0.000     .5025413     .567034
          5  |   .5341731   .0183595    29.10   0.000     .4981892     .570157
------------------------------------------------------------------------------
```

```
. marginsplot, x(age16)

  Variables that uniquely identify margins: age16
```

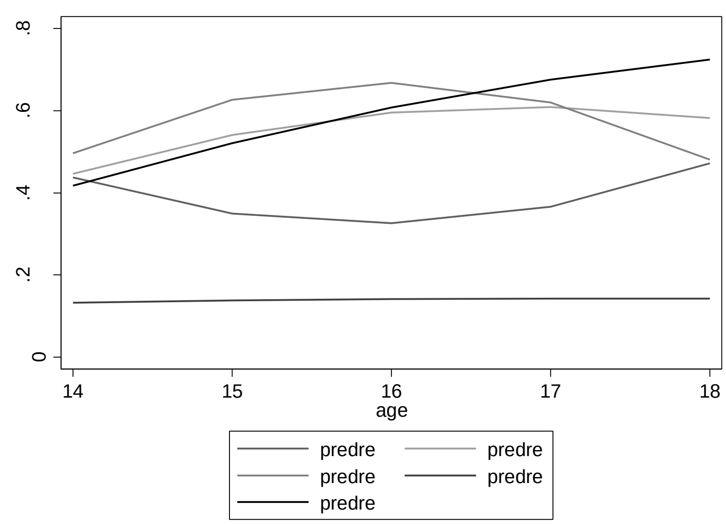


This is identical to calculating:

```
. gen pred= .5140183+.0314627 *age16 -.0106962 *age16sq
```

This is the average trajectory; let's see some of the variation across individuals, however. For that, we will obtain estimates of random effects for all three components of the equation and add them to the average coefficients:

```
. predict re*, reffects

. gen predre=.5140183+re3+(.0314627+re1) *age16 +(-.0106962+re2) *age16*age16

. graph twoway (line predre age if id==7) (line predre age if id==54) (line predre age
if id==84)  (line predre age if id==104)  (line predre age if id==111)
```



Let's compare the linear and quadratic models using LR test and BIC.

```
. est store squared
```

```
. estat ic
Akaike's information criterion and Bayesian information criterion
-----------------------------------------------------------------------------
       Model |          N   ll(null)  ll(model)      df        AIC        BIC
-------------+---------------------------------------------------------------
     squared |      1,066          .   76.20696      10  -132.4139  -82.69722
-----------------------------------------------------------------------------
Note: BIC uses N = number of observations. See [R] BIC note.

. qui mixed attit age16 || id: age16, cov(unstructured)

. est store linear

. estat ic
Akaike's information criterion and Bayesian information criterion
-----------------------------------------------------------------------------
       Model |          N   ll(null)  ll(model)      df        AIC        BIC
-------------+---------------------------------------------------------------
      linear |      1,066          .   57.44211       6  -102.8842   -73.0542
-----------------------------------------------------------------------------
Note: BIC uses N = number of observations. See [R] BIC note.

. lrtest squared linear

Likelihood-ratio test                                 LR chi2(4)  =     37.53
(Assumption: linear nested in squared)                Prob > chi2 =    0.0000

Note: The reported degrees of freedom assumes the null hypothesis is not on
      the boundary of the parameter space.  If this is not true, then the
      reported test is conservative.
```

Both LR test and difference in BIC (almost 10) indicate that the model with age squared offers a better fit. Let's also test if there is statistically significant variance in age and age squared slopes – for that, we estimate a model with the quadratic term but without variance components for either term for age:

```
. mixed attit c.age16##c.age16 || id:
Mixed-effects ML regression                     Number of obs     =      1,066
Group variable: id                              Number of groups  =        241

                                                Obs per group:
                                                              min =          1
                                                              avg =        4.4
                                                              max =          5

                                                Wald chi2(2)      =      67.32
Log likelihood =  41.042571                     Prob > chi2       =     0.0000
------------------------------------------------------------------------------
       attit |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       age16 |   .0322138   .0042341     7.61   0.000     .0239152    .0405125
             |
     c.age16#|
     c.age16 |  -.0103428   .0034889    -2.96   0.003    -.0171809   -.0035046
             |
       _cons |   .5130925   .0163594    31.36   0.000     .4810287    .5451563
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
id: Identity                 |
                  var(_cons) |    .044296   .0048807      .0356925    .0549734
```

```
-----------------------------+------------------------------------------------
               var(Residual) |   .0357293   .0017625      .0324366    .0393562
------------------------------------------------------------------------------
LR test vs. linear model: chibar2(01) = 400.75        Prob >= chibar2 = 0.0000

. lrtest . squared
Likelihood-ratio test                                 LR chi2(5)  =     70.33
(Assumption: . nested in squared)                     Prob > chi2 =    0.0000

Note: The reported degrees of freedom assumes the null hypothesis is not on
      the boundary of the parameter space.  If this is not true, then the
      reported test is conservative.

. estat ic

Akaike's information criterion and Bayesian information criterion
-----------------------------------------------------------------------------
       Model |          N   ll(null)  ll(model)      df        AIC        BIC
-------------+---------------------------------------------------------------
           . |      1,066          .   41.04257       5  -72.08514   -47.2268
-----------------------------------------------------------------------------
Note: BIC uses N = number of observations. See [R] BIC note.
```

The model with random variation in age and age squared slopes is clearly superior. Next, let's add variables that could explain variation in attitudes. We start with time-varying (level 1) variables – here we have expo. But it is possible for effects of this variable to also vary across individuals so we allow for such variation:

```
. mixed attit c.age16##c.age16 expo || id: c.age16##c.age16 expo, cov(unstructured)
Mixed-effects ML regression                     Number of obs     =      1,066
Group variable: id                              Number of groups  =        241

                                                Obs per group:
                                                              min =          1
                                                              avg =        4.4
                                                              max =          5

                                                Wald chi2(3)      =     271.97
Log likelihood =  206.97823                     Prob > chi2       =     0.0000
------------------------------------------------------------------------------
       attit |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       age16 |   .0229368   .0048524     4.73   0.000     .0134264    .0324473
             |
     c.age16#|
     c.age16 |  -.0045777   .0032361    -1.41   0.157    -.0109204     .001765
             |
        expo |   .4394308    .030223    14.54   0.000     .3801949    .4986667
       _cons |   .2521722   .0214879    11.74   0.000     .2100568    .2942877
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
id: Unstructured             |
                  var(age16) |   .0026496   .0005289      .0017917    .0039181
            var(age16#age16) |   .0006709   .0002482      .0003249    .0013854
                   var(expo) |   .0544416   .0170385      .0294803    .1005381
                  var(_cons) |     .04243   .0088457      .0281978    .0638455
      cov(age16,age16#age16) |  -.0003008   .0002386     -.0007686    .0001669
             cov(age16,expo) |   -.002203    .002109     -.0063365    .0019305
            cov(age16,_cons) |    .001878     .00158     -.0012187    .0049748
```

```
       cov(age16#age16,expo) |   .0010966   .0013227     -.0014958    .0036889
      cov(age16#age16,_cons) |  -.0022545    .001151     -.0045104    1.46e-06
             cov(expo,_cons) |  -.0305941   .0112535     -.0526506   -.0085377
-----------------------------+------------------------------------------------
               var(Residual) |   .0198866   .0015101      .0171366     .023078
------------------------------------------------------------------------------
LR test vs. linear model: chi2(10) = 289.94              Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.
```

There is significant variation in slopes of all of these three level 1 variables. Next, we add level 2 (time invariant) variables as predictors of attitudes (but not yet of slopes). We have the following level 2 predictors: female, minority, and income. Since we are now trying to model variance in the constant (intercept), we should make sure that intercept is meaningful by mean-centering continuous predictors.

```
. for var expo income: sum X \ gen Xm=X-r(mean)

->  sum expo
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
        expo |      1066    .5601501    .3106114          0       1.61

->  gen expom=expo-r(mean)
(139 missing values generated)

->  sum income
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      income |      1205    4.091286    2.346617          1         10

->  gen incomem=income-r(mean)

. mixed attit c.age16##c.age16 expom female minority incomem  || id: c.age16##c.age16
expom, cov(unstructured)
Mixed-effects ML regression                     Number of obs     =      1,066
Group variable: id                              Number of groups  =        241

                                                Obs per group:
                                                              min =          1
                                                              avg =        4.4
                                                              max =          5

                                                Wald chi2(6)      =     294.71
Log likelihood =  213.47687                     Prob > chi2       =     0.0000
------------------------------------------------------------------------------
       attit |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       age16 |   .0227999   .0048543     4.70   0.000     .0132857    .0323141
             |
     c.age16#|
     c.age16 |  -.0043997   .0032383    -1.36   0.174    -.0107467    .0019472
             |
       expom |   .4427109   .0298912    14.81   0.000     .3841253    .5012965
      female |  -.0497872   .0217828    -2.29   0.022    -.0924807   -.0070937
    minority |   .0224325   .0275622     0.81   0.416    -.0315885    .0764534
     incomem |   .0141915   .0048076     2.95   0.003     .0047688    .0236142
       _cons |   .5146037   .0176118    29.22   0.000     .4800851    .5491223
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
```

```
-----------------------------+------------------------------------------------
id: Unstructured             |
                 var(age16) |   .0026491   .0005294      .0017906    .0039192
           var(age16#age16) |   .0006767   .0002489      .0003291    .0013914
                 var(expom) |   .0523096   .0165384      .0281489    .0972077
                 var(_cons) |   .0237323   .0036756      .0175189    .0321494
     cov(age16,age16#age16) |  -.0003033   .0002387     -.0007712    .0001645
           cov(age16,expom) |  -.0020604   .0020852     -.0061474    .0020266
           cov(age16,_cons) |   .0009049   .0009612     -.0009789    .0027888
     cov(age16#age16,expom) |   .0011388   .0013086      -.001426    .0037037
     cov(age16#age16,_cons) |  -.0016059   .0007489     -.0030738   -.0001381
           cov(expom,_cons) |   .0018312   .0055434     -.0090338    .0126961
-----------------------------+------------------------------------------------
              var(Residual) |   .0199329   .0015127      .0171781    .0231296
------------------------------------------------------------------------------
LR test vs. linear model: chi2(10) = 270.95             Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.
```

Alternatively, we could consider examining the effects of average exposure over time (aggregated against time points) vs exposure at a particular time point. We could either grand-mean center or group mean center exposure for this. Here, group mean centering allows interpretation of parameter estimates as effects of change over time within-person. [Under grand-mean centering or no centering, the parameter estimates reflect a combination of change over time and differences across individuals.] But when we use a group-centered predictor, we only estimate only change effects (within-person component). Of course, when group mean centering, we should include person level variables alongside group-mean centered predictors. This way, we separate within and between unit effects.

```
. by id: egen expomean=mean(expo)

. gen expochange=expo-expomean
(139 missing values generated)

. sum expomean expochange
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
    expomean |      1,066    .5601501    .2532099          0       1.32
  expochange |      1,066   -1.47e-09    .1799004      -.904       .706

. sum expomean
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
    expomean |      1,066    .5601501    .2532099          0       1.32

. gen expomeanm=expomean-r(mean)

. mixed attit c.age16##c.age16 expochange expomeanm female minority incomem  || id:
c.age16##c.age16 expochange, cov(unstructured)
Mixed-effects ML regression                     Number of obs     =      1,066
Group variable: id                              Number of groups  =        241

                                                Obs per group:
                                                              min =          1
                                                              avg =        4.4
                                                              max =          5

                                                Wald chi2(7)      =     384.26
Log likelihood =  227.43873                     Prob > chi2       =     0.0000
```

```
------------------------------------------------------------------------------
       attit |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       age16 |    .024962   .0049115     5.08   0.000     .0153357    .0345883
             |
     c.age16#|
     c.age16 |  -.0058366   .0032127    -1.82   0.069    -.0121333    .0004601
             |
  expochange |   .3470481   .0372215     9.32   0.000     .2740952     .420001
   expomeanm |   .6187336   .0407646    15.18   0.000     .5388365    .6986307
      female |  -.0433309   .0211836    -2.05   0.041    -.0848499   -.0018119
    minority |   .0118951   .0267071     0.45   0.656    -.0404499    .0642401
     incomem |   .0165267   .0046909     3.52   0.000     .0073327    .0257207
       _cons |    .520923   .0170787    30.50   0.000     .4874494    .5543966
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
id: Unstructured             |
                  var(age16) |   .0027411   .0005396      .0018637    .0040317
            var(age16#age16) |   .0006593   .0002439      .0003193    .0013614
             var(expochange) |   .0688762   .0234179      .0353722     .134115
                  var(_cons) |   .0252743   .0034363       .019362    .0329919
      cov(age16,age16#age16) |  -.0003023   .0002421     -.0007768    .0001722
       cov(age16,expochange) |  -.0025692   .0025629     -.0075924     .002454
            cov(age16,_cons) |   .0012069   .0009569     -.0006687    .0030825
 cov(age16#age16,expochange) |   .0004807   .0015122     -.0024831    .0034446
      cov(age16#age16,_cons) |  -.0017197   .0007289     -.0031483   -.0002911
       cov(expochange,_cons) |    .006767   .0068674     -.0066929    .0202269
-----------------------------+------------------------------------------------
               var(Residual) |   .0191308   .0014619      .0164698    .0222217
------------------------------------------------------------------------------
LR test vs. linear model: chi2(10) = 272.29               Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.
```

Next, we will estimate a model where we will use cross-level interactions to explain variance in slopes across individuals. That is, we will introduce interactions of level 1 predictors with level 2 time-invariant variables and then see what happens to variance of slopes of those level 1 predictors.

```
. mixed attit c.age16##c.age16##c.expomeanm c.age16##c.age16##i.female
c.age16##c.age16##i.minority c.age16##c.age16##c.incomem c.expochange##c.expomeanm
c.expochange##i.female c.expochange##i.minority c.expochange##c.incomem  || id:
c.age16##c.age16 expochange, cov(unstructured)
Mixed-effects ML regression                     Number of obs     =      1,066
Group variable: id                              Number of groups  =        241

                                                Obs per group:
                                                              min =          1
                                                              avg =        4.4
                                                              max =          5

                                                Wald chi2(19)     =     439.44
Log likelihood =  243.59055                     Prob > chi2       =     0.0000
------------------------------------------------------------------------------
       attit |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       age16 |    .023436   .0074107     3.16   0.002     .0089113    .0379608
             |
     c.age16#|
```

```
      c.age16 |  -.0127451   .0047738    -2.67   0.008    -.0221016   -.0033887
              |
    expomeanm |   .6235837     .04901    12.72   0.000     .5275259    .7196416
              |
      c.age16#|
  c.expomeanm |   -.045851   .0188732    -2.43   0.015    -.0828419   -.0088602
              |
      c.age16#|
      c.age16#|
  c.expomeanm |  -.0041202   .0126952    -0.32   0.746    -.0290023    .0207619
              |
        age16 |          0  (omitted)
     1.female |  -.0738721   .0252524    -2.93   0.003     -.123366   -.0243783
              |
       female#|
      c.age16 |
           1  |   .0048259   .0097991     0.49   0.622    -.0143799    .0240318
              |
       female#|
      c.age16#|
      c.age16 |
           1  |   .0148405   .0064114     2.31   0.021     .0022744    .0274066
              |
        age16 |          0  (omitted)
   1.minority |  -.0079024   .0318651    -0.25   0.804    -.0703567     .054552
              |
     minority#|
      c.age16 |
           1  |  -.0073877   .0124986    -0.59   0.554    -.0318845     .017109
              |
     minority#|
      c.age16#|
      c.age16 |
           1  |   .0047483   .0082399     0.58   0.564    -.0114016    .0208982
              |
        age16 |          0  (omitted)
      incomem |   .0102222   .0055722     1.83   0.067     -.000699    .0211435
              |
      c.age16#|
    c.incomem |  -.0024181   .0021411    -1.13   0.259    -.0066145    .0017784
              |
      c.age16#|
      c.age16#|
    c.incomem |   .0022079   .0013923     1.59   0.113    -.0005208    .0049367
              |
   expochange |   .4121574   .0551548     7.47   0.000     .3040559    .5202589
    expomeanm |          0  (omitted)
              |
           c. |
   expochange#|
  c.expomeanm |   .2282904    .156432     1.46   0.144    -.0783106    .5348915
              |
   expochange |          0  (omitted)
              |
       female#|
 c.expochange |
           1  |  -.0151889   .0744531    -0.20   0.838    -.1611144    .1307366
              |
   expochange |          0  (omitted)
              |
     minority#|
 c.expochange |
           1  |  -.3040498   .0908928    -3.35   0.001    -.4821965   -.1259031
```

```
             |
  expochange |          0  (omitted)
     incomem |          0  (omitted)
             |
          c. |
  expochange#|
   c.incomem |  -.0445934   .0168649    -2.64   0.008     -.077648   -.0115387
             |
       _cons |   .5369439   .0186606    28.77   0.000     .5003697     .573518
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
id: Unstructured             |
                  var(age16) |   .0025617   .0005206      .0017201    .0038151
            var(age16#age16) |    .000551   .0002324      .0002411    .0012595
             var(expochange) |   .0545077   .0203652      .0262078    .1133664
                  var(_cons) |   .0247072   .0033559      .0189325    .0322433
      cov(age16,age16#age16) |  -.0003286   .0002313     -.0007819    .0001247
       cov(age16,expochange) |  -.0029839   .0024093     -.0077061    .0017383
            cov(age16,_cons) |   .0012685   .0009229     -.0005403    .0030773
 cov(age16#age16,expochange) |   .0013568   .0014238     -.0014339    .0041475
      cov(age16#age16,_cons) |  -.0014805   .0006986     -.0028496   -.0001113
       cov(expochange,_cons) |   .0017472   .0064933     -.0109795    .0144739
-----------------------------+------------------------------------------------
               var(Residual) |   .0191457   .0014546      .0164969    .0222199
------------------------------------------------------------------------------
LR test vs. linear model: chi2(10) = 278.01               Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.
```

Let's simplify the model by omitting non-significant cross-level interactions; we will use LR test and BIC to make sure we do not omit anything important:

```
. est store full

. estat ic
Akaike's information criterion and Bayesian information criterion
-----------------------------------------------------------------------------
       Model |          N   ll(null)  ll(model)      df        AIC        BIC
-------------+---------------------------------------------------------------
        full |      1,066          .   243.5905      31  -425.1811  -271.0594
-----------------------------------------------------------------------------
Note: BIC uses N = number of observations. See [R] BIC note.

. mixed attit c.age16##c.age16##c.expomeanm c.age16##c.age16##i.female
c.expochange##i.minority c.expochange##c.incomem  || id: c.age16##c.age16 expochange,
cov(unstructured)
Mixed-effects ML regression                     Number of obs     =      1,066
Group variable: id                              Number of groups  =        241

                                                Obs per group:
                                                              min =          1
                                                              avg =        4.4
                                                              max =          5
                                                Wald chi2(13)     =     430.09
Log likelihood =  240.81527                     Prob > chi2       =     0.0000
------------------------------------------------------------------------------
       attit |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       age16 |   .0219426   .0064814     3.39   0.001     .0092393    .0346459
             |
     c.age16#|
```

```
     c.age16 |  -.0122447   .0042228    -2.90   0.004    -.0205212   -.0039681
             |
   expomeanm |   .6318958    .049011    12.89   0.000      .535836    .7279556
             |
     c.age16#|
 c.expomeanm |  -.0408274   .0187166    -2.18   0.029    -.0775112   -.0041436
             |
     c.age16#|
     c.age16#|
 c.expomeanm |  -.0091057   .0124199    -0.73   0.463    -.0334482    .0152368
             |
       age16 |          0  (omitted)
    1.female |  -.0738328   .0252351    -2.93   0.003    -.1232926    -.024373
             |
      female#|
     c.age16 |
          1  |   .0039725   .0095473     0.42   0.677    -.0147399    .0226849
             |
      female#|
     c.age16#|
     c.age16 |
          1  |   .0151555   .0062652     2.42   0.016     .0028759    .0274351
             |
  expochange |   .4158301   .0420899     9.88   0.000     .3333353    .4983248
  1.minority |   .0038862   .0267865     0.15   0.885    -.0486143    .0563868
             |
    minority#|
c.expochange |
          1  |  -.3121521   .0882582    -3.54   0.000    -.4851349   -.1391693
             |
  expochange |          0  (omitted)
     incomem |   .0152012   .0047064     3.23   0.001     .0059768    .0244256
             |
          c. |
  expochange#|
   c.incomem |  -.0550901    .016154    -3.41   0.001    -.0867514   -.0234288
             |
       _cons |   .5347902   .0180973    29.55   0.000     .4993201    .5702604
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
id: Unstructured             |
                  var(age16) |   .0026167   .0005263      .0017642     .003881
            var(age16#age16) |   .0005755   .0002339      .0002595    .0012763
             var(expochange) |   .0574034   .0207247      .0282894    .1164801
                  var(_cons) |    .024968   .0033839      .0191435    .0325648
      cov(age16,age16#age16) |  -.0003702   .0002346      -.00083     .0000897
       cov(age16,expochange) |  -.0030557   .0024417     -.0078412    .0017299
            cov(age16,_cons) |   .0013106   .0009329     -.0005178     .003139
 cov(age16#age16,expochange) |   .0015322   .0014623     -.0013339    .0043982
      cov(age16#age16,_cons) |  -.0015776   .0007069     -.0029631    -.000192
       cov(expochange,_cons) |   .0017044   .0066061     -.0112434    .0146521
-----------------------------+------------------------------------------------
               var(Residual) |   .0191298   .0014506      .0164878     .022195
------------------------------------------------------------------------------
LR test vs. linear model: chi2(10) = 278.16            Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.

. est store reduced
. lrtest reduced full
Likelihood-ratio test                                 LR chi2(6)  =      5.55
```

```
(Assumption: . nested in full)                         Prob > chi2 =    0.4754

. estat ic
Akaike's information criterion and Bayesian information criterion
-----------------------------------------------------------------------------
       Model |          N   ll(null)  ll(model)      df        AIC        BIC
-------------+---------------------------------------------------------------
           . |      1,066          .   240.8153      25  -431.6305  -307.3388
-----------------------------------------------------------------------------
Note: BIC uses N = number of observations. See [R] BIC note.
```

No significant difference in model fit indicated by LR test, and BIC is substantially smaller in the reduced model; therefore, we can use the reduced model.

<u>Longitudinal HLM with More Complex Level 1 Variance Structures</u>

So far, we did not take into account possible correlations of residuals across time points. Examples of covariance structures for a dataset with six time points (t1-t6):

Independent = basic ME (one parameters for level 1 residual variance, σ^2):

	t_1	t_2	t_3	t_4	t_5	t_6
t_1	—	0	0	0	0	0
t_2	0	—	0	0	0	0
t_3	0	0	—	0	0	0
t_4	0	0	0	—	0	0
t_5	0	0	0	0	—	0
t_6	0	0	0	0	0	—

Exchangeable (2 parameters for residuals: variance σ^2 and covariance):

	t_1	t_2	t_3	t_4	t_5	t_6
t_1	—	ρ	ρ	ρ	ρ	ρ
t_2	ρ	—	ρ	ρ	ρ	ρ
t_3	ρ	ρ	—	ρ	ρ	ρ
t_4	ρ	ρ	ρ	—	ρ	ρ
t_5	ρ	ρ	ρ	ρ	—	ρ
t_6	ρ	ρ	ρ	ρ	ρ	—

First-order autoregressive (two parameters for level 1 residual variance, σ^2 and ρ):

	t_1	t_2	t_3	t_4	t_5	t_6
t_1	—	ρ	0	0	0	0
t_2	ρ	—	ρ	0	0	0
t_3	0	ρ	—	ρ	0	0
t_4	0	0	ρ	—	ρ	0
t_5	0	0	0	ρ	—	ρ
t_6	0	0	0	0	ρ	—

Unrestricted (six separate σ^2 parameters for residual variance and 15 covariances):

	t_1	t_2	t_3	t_4	t_5	t_6
t_1	—	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5
t_2	ρ_1	—	ρ_6	ρ_7	ρ_8	ρ_9
t_3	ρ_2	ρ_6	—	ρ_{10}	ρ_{11}	ρ_{12}
t_4	ρ_3	ρ_7	ρ_{10}	—	ρ_{13}	ρ_{14}
t_5	ρ_4	ρ_8	ρ_{11}	ρ_{13}	—	ρ_{15}
t_6	ρ_5	ρ_9	ρ_{12}	ρ_{14}	ρ_{15}	—

Exponential (two parameters for level 1 residual variance, σ^2 and ρ):

	t_1	t_2	t_3	t_4	t_5	t_6
t_1	—	ρ^1	ρ^2	ρ^3	ρ^4	ρ^5
t_2	ρ^1	—	ρ^1	ρ^2	ρ^3	ρ^4
t_3	ρ^2	ρ^1	—	ρ^1	ρ^2	ρ^3
t_4	ρ^3	ρ^2	ρ^1	—	ρ^1	ρ^2
t_5	ρ^4	ρ^3	ρ^2	ρ^1	—	ρ^1
t_6	ρ^5	ρ^4	ρ^3	ρ^2	ρ^1	—

We will model some of these error structures for our final example from NYS data.
Exchangeable residuals:

```
. mixed attit c.age16##c.age16##c.expomeanm c.age16##c.age16##i.female
c.expochange##i.minority c.expochange##c.incomem  || id: c.age16##c.age16 expochange,
cov(unstructured) residuals(exchangeable)

Mixed-effects ML regression                     Number of obs     =      1,066
Group variable: id                              Number of groups  =        241

                                                Obs per group:
                                                              min =          1
                                                              avg =        4.4
                                                              max =          5

                                                Wald chi2(13)     =     430.09
Log likelihood =  240.81527                     Prob > chi2       =     0.0000
------------------------------------------------------------------------------
       attit |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       age16 |   .0219426   .0064814     3.39   0.001     .0092393    .0346459
             |
     c.age16#|
     c.age16 |  -.0122447   .0042228    -2.90   0.004    -.0205212   -.0039681
             |
   expomeanm |   .6318958    .049011    12.89   0.000      .535836    .7279557
             |
     c.age16#|
 c.expomeanm |  -.0408274   .0187166    -2.18   0.029    -.0775112   -.0041436
             |
     c.age16#|
     c.age16#|
 c.expomeanm |  -.0091057   .0124199    -0.73   0.463    -.0334482    .0152368
             |
       age16 |          0  (omitted)
    1.female |  -.0738328   .0252351    -2.93   0.003    -.1232926   -.0243729
             |
      female#|
     c.age16 |
          1  |   .0039725   .0095473     0.42   0.677    -.0147399    .0226849
             |
      female#|
     c.age16#|
     c.age16 |
          1  |   .0151555   .0062652     2.42   0.016     .0028759    .0274351
             |
  expochange |     .41583   .0420899     9.88   0.000     .3333353    .4983247
  1.minority |   .0038862   .0267865     0.15   0.885    -.0486143    .0563868
             |
    minority#|
c.expochange |
          1  |  -.3121521   .0882581    -3.54   0.000    -.4851349   -.1391693
             |
  expochange |          0  (omitted)
     incomem |   .0152012   .0047064     3.23   0.001     .0059767    .0244256
             |
          c. |
  expochange#|
   c.incomem |  -.0550901    .016154    -3.41   0.001    -.0867514   -.0234288
             |
       _cons |   .5347902   .0180973    29.55   0.000     .4993201    .5702604
------------------------------------------------------------------------------
```

```
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
id: Unstructured             |
                 var(age16) |   .0026167   .0005263      .0017642     .003881
           var(age16#age16) |   .0005755   .0002339      .0002595    .0012763
            var(expochange) |   .0574034   .0207247      .0282894    .1164801
                 var(_cons) |   .0144426   .8755494      3.61e-54    5.78e+49
     cov(age16,age16#age16) |  -.0003702   .0002346       -.00083    .0000897
      cov(age16,expochange) |  -.0030557   .0024417     -.0078412    .0017299
           cov(age16,_cons) |   .0013106   .0009329     -.0005178     .003139
 cov(age16#age16,expochange) |   .0015322   .0014623     -.0013339    .0043982
     cov(age16#age16,_cons) |  -.0015776   .0007069     -.0029631    -.000192
      cov(expochange,_cons) |   .0017044   .0066061     -.0112434    .0146521
-----------------------------+------------------------------------------------
Residual: Exchangeable       |
                     var(e) |   .0296553   .8755483      2.19e-27    4.01e+23
                     cov(e) |   .0105255   .8755479     -1.705517    1.726568
------------------------------------------------------------------------------
LR test vs. linear model: chi2(11) = 278.16              Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.

. lrtest reduced exch
Likelihood-ratio test                                LR chi2(1)  =     -0.00
(Assumption: reduced nested in exch)                 Prob > chi2 =    1.0000

. estat ic
Akaike's information criterion and Bayesian information criterion
-----------------------------------------------------------------------------
       Model |          N   ll(null)  ll(model)      df        AIC        BIC
-------------+---------------------------------------------------------------
        exch |      1,066          .   240.8153      26  -429.6305  -300.3672
-----------------------------------------------------------------------------
Note: BIC uses N = number of observations. See [R] BIC note.
```

Doesn't look like the model with exchangeable residuals provides any advantage (likely because random effects can account for this type of residual covariance anyways). Model with unstructured residuals is too complex and doesn't converge. Now AR1 model:

```
. mixed attit c.age16##c.age16##c.expomeanm c.age16##c.age16##i.female
c.expochange##i.minority c.expochange##c.incomem  || id: c.age16##c.age16 expochange,
cov(unstructured) residuals(ar1, t(age))
Mixed-effects ML regression                     Number of obs     =      1,066
Group variable: id                              Number of groups  =        241

                                                Obs per group:
                                                              min =          1
                                                              avg =        4.4
                                                              max =          5

                                                Wald chi2(13)     =     430.20
Log likelihood =  240.87777                     Prob > chi2       =     0.0000
------------------------------------------------------------------------------
       attit |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       age16 |   .0221019   .0064781     3.41   0.001      .009405    .0347987
             |
     c.age16#|
     c.age16 |  -.0122122   .0042226    -2.89   0.004    -.0204883    -.003936
             |
```

```
   expomeanm |   .6316503   .0489905    12.89   0.000     .5356308    .7276699
             |
     c.age16#|
 c.expomeanm |  -.0408138   .0187053    -2.18   0.029    -.0774754   -.0041521
             |
     c.age16#|
     c.age16#|
 c.expomeanm |  -.0090311   .0124208    -0.73   0.467    -.0333754    .0153133
             |
       age16 |          0  (omitted)
    1.female |  -.0738404   .0252218    -2.93   0.003    -.1232741   -.0244067
             |
      female#|
     c.age16 |
          1  |   .0039593   .0095424     0.41   0.678    -.0147435     .022662
             |
      female#|
     c.age16#|
     c.age16 |
          1  |   .0151638   .0062656     2.42   0.016     .0028835    .0274441
             |
  expochange |   .4163352   .0420688     9.90   0.000     .3338818    .4987886
  1.minority |   .0036617   .0267985     0.14   0.891    -.0488623    .0561858
             |
    minority#|
c.expochange |
          1  |  -.3145651    .088194    -3.57   0.000    -.4874222   -.1417081
             |
  expochange |          0  (omitted)
     incomem |   .0151699   .0047086     3.22   0.001     .0059413    .0243985
             |
          c. |
  expochange#|
   c.incomem |  -.0552256   .0161464    -3.42   0.001     -.086872   -.0235791
             |
       _cons |   .5347964   .0180898    29.56   0.000     .4993411    .5702518
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
id: Unstructured             |
                  var(age16) |   .0024375   .0007501      .0013335    .0044554
            var(age16#age16) |   .0005043   .0003135      .0001491    .0017053
             var(expochange) |   .0568497   .0207923      .0277592    .1164256
                 var(_cons) |   .0239302   .0046245      .0163852    .0349494
      cov(age16,age16#age16) |  -.0003641   .0002346      -.000824    .0000957
       cov(age16,expochange) |  -.0029671   .0024495     -.0077679    .0018338
            cov(age16,_cons) |   .0012863   .0009337     -.0005438    .0031164
 cov(age16#age16,expochange) |   .0015119   .0014592      -.001348    .0043718
      cov(age16#age16,_cons) |  -.0013656   .0009448     -.0032174    .0004861
       cov(expochange,_cons) |   .0019403   .0066299     -.0110541    .0149346
-----------------------------+------------------------------------------------
Residual: AR(1)              |
                         rho |     .04265   .1225802     -.1954666     .276018
                      var(e) |   .0201736   .0035154      .0143369    .0283866
------------------------------------------------------------------------------
LR test vs. linear model: chi2(11) = 278.29              Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.

. est store ar1

. estat ic
```

```
Akaike's information criterion and Bayesian information criterion
-----------------------------------------------------------------------------
       Model |          N   ll(null)  ll(model)      df         AIC        BIC
-------------+---------------------------------------------------------------
         ar1 |      1,066          .   240.8778      26   -429.7555  -300.4922
-----------------------------------------------------------------------------
Note: BIC uses N = number of observations. See [R] BIC note.

. lrtest reduced ar1
Likelihood-ratio test                                 LR chi2(1)  =      0.13
(Assumption: reduced nested in ar1)                   Prob > chi2 =    0.7237
```

You can try introducing even more complexity – for example, if you suspect that there is second-order autocorrelation and estimate AR2 model:

```
. mixed attit c.age16##c.age16##c.expomeanm c.age16##c.age16##i.female
c.expochange##i.minority c.expochange##c.incomem  || id: c.age16##c.age16 expochange,
cov(unstructured) residuals(ar2, t(age))

Mixed-effects ML regression                     Number of obs     =      1,066
Group variable: id                              Number of groups  =        241

                                                Obs per group:
                                                              min =          1
                                                              avg =        4.4
                                                              max =          5

                                                Wald chi2(13)     =     429.58
Log likelihood =   241.3358                     Prob > chi2       =     0.0000
------------------------------------------------------------------------------
       attit |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       age16 |   .0224681   .0064888     3.46   0.001     .0097503    .0351858
             |
     c.age16#|
     c.age16 |  -.0122816   .0042095    -2.92   0.004    -.0205321   -.0040311
             |
   expomeanm |    .632239   .0489395    12.92   0.000     .5363194    .7281586
             |
     c.age16#|
 c.expomeanm |  -.0404713   .0187303    -2.16   0.031    -.0771821   -.0037605
             |
     c.age16#|
     c.age16#|
 c.expomeanm |  -.0093271   .0123895    -0.75   0.452    -.0336101    .0149558
             |
       age16 |          0  (omitted)
    1.female |  -.0739367   .0252005    -2.93   0.003    -.1233288   -.0245445
             |
      female#|
     c.age16 |
          1  |    .004258   .0095605     0.45   0.656    -.0144803    .0229962
             |
      female#|
     c.age16#|
     c.age16 |
          1  |   .0151775   .0062521     2.43   0.015     .0029236    .0274314
             |
  expochange |   .4149024   .0420531     9.87   0.000     .3324797     .497325
  1.minority |    .004255   .0267561     0.16   0.874     -.048186     .056696
             |
    minority#|
```

```
c.expochange |
          1  |  -.3200006   .0881899    -3.63   0.000    -.4928497   -.1471516
             |
  expochange |          0  (omitted)
     incomem |   .0151398   .0047024     3.22   0.001     .0059233    .0243562
             |
          c. |
  expochange#|
   c.incomem |  -.0542243   .0161455    -3.36   0.001    -.0858688   -.0225797
             |
       _cons |   .5348592   .0180685    29.60   0.000     .4994455    .5702729
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
id: Unstructured             |
                 var(age16)  |   .0009802   .0020999      .0000147     .065287
           var(age16#age16)  |   .0003188   .0003423      .0000389    .0026144
            var(expochange)  |   .0568038   .0208301      .0276847    .1165506
                 var(_cons)  |   .0095053   .0291655      .0000232    3.888144
       cov(age16,age16#age16) |  -.000366    .000234     -.0008247    .0000928
       cov(age16,expochange) |  -.0029069   .0024651     -.0077384    .0019246
            cov(age16,_cons) |   .0012498   .0009353     -.0005833    .0030829
 cov(age16#age16,expochange) |   .0014084   .0014574     -.0014481    .0042648
      cov(age16#age16,_cons) |  -.0003384   .0015672     -.0034102    .0027333
       cov(expochange,_cons) |    .002175   .0066269     -.0108134    .0151634
-----------------------------+------------------------------------------------
Residual: AR(2)              |
                       phi1  |   .3372661   .3164079     -.2828819    .9574142
                       phi2  |   .1730083    .161653     -.1438258    .4898423
                     var(e)  |   .0346567   .0291324      .0066722     .180014
------------------------------------------------------------------------------
LR test vs. linear model: chi2(12) = 279.20              Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.

. est store ar2

. estat ic
Akaike's information criterion and Bayesian information criterion
-----------------------------------------------------------------------------
       Model |          N   ll(null)  ll(model)      df        AIC        BIC
-------------+---------------------------------------------------------------
         ar2 |      1,066          .   241.3358      27  -428.6716  -294.4366
-----------------------------------------------------------------------------
Note: BIC uses N = number of observations. See [R] BIC note.

. lrtest reduced ar2
Likelihood-ratio test                                 LR chi2(2)  =      1.04
(Assumption: reduced nested in ar2)                   Prob > chi2 =    0.5942
```

Exponential model:

```
. mixed attit c.age16##c.age16##c.expomeanm c.age16##c.age16##i.female
c.expochange##i.minority c.expochange##c.incomem  || id: c.age16##c.age16 expochange,
cov(unstructured) residuals(exponential, t(age))

Mixed-effects ML regression                     Number of obs     =      1,066
Group variable: id                              Number of groups  =        241

                                                Obs per group:
```

```
                                                     min =          1
                                                     avg =        4.4
                                                     max =          5

                                                Wald chi2(13)     =     430.20
Log likelihood =  240.87777                     Prob > chi2       =     0.0000
------------------------------------------------------------------------------
       attit |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       age16 |   .0221019   .0064781     3.41   0.001      .009405    .0347987
             |
     c.age16#|
     c.age16 |  -.0122122   .0042226    -2.89   0.004    -.0204883    -.003936
             |
   expomeanm |   .6316504   .0489905    12.89   0.000     .5356307      .72767
             |
     c.age16#|
 c.expomeanm |  -.0408138   .0187053    -2.18   0.029    -.0774754   -.0041521
             |
     c.age16#|
     c.age16#|
 c.expomeanm |  -.0090311   .0124208    -0.73   0.467    -.0333755    .0153133
             |
       age16 |          0  (omitted)
    1.female |  -.0738404   .0252218    -2.93   0.003    -.1232742   -.0244066
             |
      female#|
     c.age16 |
          1  |   .0039593   .0095424     0.41   0.678    -.0147435     .022662
             |
      female#|
     c.age16#|
     c.age16 |
          1  |   .0151638   .0062656     2.42   0.016     .0028835    .0274441
             |
  expochange |   .4163351   .0420688     9.90   0.000     .3338817    .4987885
  1.minority |   .0036618   .0267985     0.14   0.891    -.0488624    .0561859
             |
    minority#|
c.expochange |
          1  |  -.3145651    .088194    -3.57   0.000    -.4874222    -.141708
             |
  expochange |          0  (omitted)
     incomem |   .0151699   .0047086     3.22   0.001     .0059413    .0243985
             |
          c. |
  expochange#|
   c.incomem |  -.0552255   .0161464    -3.42   0.001     -.086872   -.0235791
             |
       _cons |   .5347964   .0180898    29.56   0.000     .4993411    .5702518
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
id: Unstructured             |
                  var(age16) |   .0024375     .00075      .0013336    .0044552
            var(age16#age16) |   .0005043   .0003135      .0001492    .0017052
             var(expochange) |   .0568498   .0207923      .0277593    .1164256
                  var(_cons) |   .0239303   .0046242      .0163857    .0349486
      cov(age16,age16#age16) |  -.0003641   .0002346      -.000824    .0000957
       cov(age16,expochange) |  -.0029671   .0024495     -.0077679    .0018338
            cov(age16,_cons) |   .0012863   .0009338     -.0005438    .0031164
 cov(age16#age16,expochange) |   .0015119   .0014592      -.001348    .0043718
```

```
      cov(age16#age16,_cons) |  -.0013656   .0009447     -.0032173     .000486
       cov(expochange,_cons) |   .0019403   .0066299     -.0110541    .0149346
-----------------------------+------------------------------------------------
Residual: Exponential        |
                         rho |   .0426496   .1225629      .0001241    .9411521
                      var(e) |   .0201736    .003515      .0143374    .0283853
------------------------------------------------------------------------------
LR test vs. linear model: chi2(11) = 278.29             Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.

. est store exp

. lrtest reduced exp
Likelihood-ratio test                                 LR chi2(1)  =      0.13
(Assumption: reduced nested in exp)                   Prob > chi2 =    0.7237

. estat ic
Akaike's information criterion and Bayesian information criterion
-----------------------------------------------------------------------------
       Model |          N   ll(null)  ll(model)      df        AIC        BIC
-------------+---------------------------------------------------------------
         exp |      1,066          .   240.8778      26  -429.7555  -300.4922
-----------------------------------------------------------------------------
Note: BIC uses N = number of observations. See [R] BIC note.
```

All in all, we would select the basic homogenous model with random slopes because it is more parsimonious. Note that fixed effects and significance tests did not change much from model to model – in many cases, they are pretty robust to variance specifications. Still, it is a good idea to compare these.

```
. est table reduced exch ar1 ar2 exp, star stats(N) b(%5.3f)

--------------------------------------------------------------------------------
    Variable |   reduced         exch          ar1           ar2           exp
-------------+------------------------------------------------------------------
attit        |
       age16 |    0.022***      0.022***      0.022***      0.022***      0.022***
             |
     c.age16#|
     c.age16 |   -0.012**      -0.012**      -0.012**      -0.012**      -0.012**
             |
   expomeanm |    0.632***      0.632***      0.632***      0.632***      0.632***
             |
     c.age16#|
 c.expomeanm |   -0.041*       -0.041*       -0.041*       -0.040*       -0.041*
             |
     c.age16#|
     c.age16#|
 c.expomeanm |   -0.009        -0.009        -0.009        -0.009        -0.009
             |
       age16 |  (omitted)
             |
      female |
         1   |   -0.074**      -0.074**      -0.074**      -0.074**      -0.074**
             |
     female#|
     c.age16 |
         1   |    0.004         0.004         0.004         0.004         0.004
             |
     female#|
     c.age16#|
     c.age16 |
         1   |    0.015*        0.015*        0.015*        0.015*        0.015*
             |
       age16 |                                                      (omitted)
```

age16				(omitted)	
age16			(omitted)		
age16		(omitted)			
expochange	0.416***	0.416***	0.416***	0.415***	0.416***
minority					
1	0.004	0.004	0.004	0.004	0.004
minority# c.expochange					
1	-0.312***	-0.312***	-0.315***	-0.320***	-0.315***
expochange					(omitted)
expochange				(omitted)	
expochange			(omitted)		
expochange		(omitted)			
expochange	(omitted)				
incomem	0.015**	0.015**	0.015**	0.015**	0.015**
c. expochange# c.incomem	-0.055***	-0.055***	-0.055***	-0.054***	-0.055***
_cons	0.535***	0.535***	0.535***	0.535***	0.535***
lns1_1_1					
_cons	-2.973***	-2.973***	-3.008***	-3.464**	-3.008***
lns1_1_2					
_cons	-3.730***	-3.730***	-3.796***	-4.025***	-3.796***
lns1_1_3					
_cons	-1.429***	-1.429***	-1.434***	-1.434***	-1.434***
lns1_1_4					
_cons	-1.845***	-2.119	-1.866***	-2.328	-1.866***
atr1_1_1_2					
_cons	-0.311	-0.311	-0.341	-0.783	-0.341
atr1_1_1_3					
_cons	-0.255	-0.255	-0.258	-0.411	-0.258
atr1_1_1_4					
_cons	0.164	0.217	0.170	0.435	0.170
atr1_1_2_3					
_cons	0.273	0.273	0.290	0.344	0.290
atr1_1_2_4					
_cons	-0.443**	-0.614	-0.415*	-0.197	-0.415*
atr1_1_3_4					
_cons	0.045	0.059	0.053	0.094	0.053
lnsig_e					
_cons	-1.978***	-1.759	-1.952***	-1.681***	-1.952***
r_atr1					
_cons		0.371	0.043		
r_phi1_1					
_cons				0.337	
r_phi1_2					
_cons				0.173	
r_logitr1					
_cons					-3.111
Statistics					

```
          N |      1066           1066           1066           1066           1066
-----------------------------------------------------------------------------------------
                                                   legend: * p<0.05; ** p<0.01; *** p<0.001
```

Compare resulting covariance structures (we will do for ID==7 because it has all 5 time points; these predicted covariance structures – expressed as correlations -- are based on both RE and level 1 residual structures). We use estat wcorrelation, which displays the overall correlation matrix for a given cluster calculated on the basis of the design of the random effects and their assumed covariance and the correlation structure of the residuals. This allows for a comparison of different multilevel models in terms of the ultimate within-cluster correlation matrix that each model implies.

```
. est restore reduced
(results reduced are active now)

. estat wcorrelation, at(id==7)

Standard deviations and correlations for id = 7:

Standard deviations:

         obs |      1        2        3        4        5
-------------+---------------------------------------------
          sd |  0.252    0.212    0.215    0.215    0.225

Correlations:

         obs |      1        2        3        4        5
-------------+---------------------------------------------
           1 |  1.000
           2 |  0.533    1.000
           3 |  0.180    0.420    1.000
           4 |  0.176    0.407    0.553    1.000
           5 |  0.158    0.319    0.452    0.549    1.000

. est restore exch
(results exch are active now)

. estat wcorrelation, at(id==7)

Standard deviations and correlations for id = 7:

Standard deviations:

         obs |      1        2        3        4        5
-------------+---------------------------------------------
          sd |  0.252    0.212    0.215    0.215    0.225

Correlations:

         obs |      1        2        3        4        5
-------------+---------------------------------------------
           1 |  1.000
           2 |  0.533    1.000
           3 |  0.180    0.420    1.000
           4 |  0.176    0.407    0.553    1.000
           5 |  0.158    0.319    0.452    0.549    1.000

. est restore ar1
(results ar1 are active now)
```

```
. estat wcorrelation, at(id==7)

Standard deviations and correlations for id = 7:

Standard deviations:

       age16 |     -2      -1       0       1       2
-------------+---------------------------------------
          sd |  0.252   0.212   0.215   0.216   0.224

Correlations:

       age16 |     -2      -1       0       1       2
-------------+---------------------------------------
          -2 |  1.000
          -1 |  0.538   1.000
           0 |  0.180   0.421   1.000
           1 |  0.180   0.396   0.552   1.000
           2 |  0.164   0.321   0.447   0.552   1.000

. est restore ar2
(results ar2 are active now)

. estat wcorrelation, at(id==7)

Standard deviations and correlations for id = 7:

Standard deviations:

       age16 |     -2      -1       0       1       2
-------------+---------------------------------------
          sd |  0.251   0.214   0.215   0.217   0.223

Correlations:

       age16 |     -2      -1       0       1       2
-------------+---------------------------------------
          -2 |  1.000
          -1 |  0.545   1.000
           0 |  0.193   0.417   1.000
           1 |  0.163   0.384   0.545   1.000
           2 |  0.176   0.298   0.458   0.558   1.000

. est restore exp
(results exp are active now)

. estat wcorrelation, at(id==7)

Standard deviations and correlations for id = 7:

Standard deviations:

         age |     14      15      16      17      18
-------------+---------------------------------------
          sd |  0.252   0.212   0.215   0.216   0.224

Correlations:

         age |     14      15      16      17      18
-------------+---------------------------------------
          14 |  1.000
          15 |  0.538   1.000
          16 |  0.180   0.421   1.000
```

```
 17 |  0.180   0.396   0.552   1.000
 18 |  0.164   0.321   0.447   0.552   1.000
```

Note that the residuals option of mixed command allows to examine a range of residuals structures – for example, it can allows us to estimate heterogenous residuals to deal with heteroscedasticity issues, etc. Here are all the suboptions to be used within residuals option:

```
independent, the default, specifies that all residuals be independent
            and identically distributed Gaussian with one common variance.
            When combined with by(varname), independence is still assumed,
            but you estimate a distinct variance for each level of varname.
            Unlike with the structures described below, varname does not need
            to be constant within groups.

        exchangeable estimates two parameters, one common within-group
            variance and one common pairwise covariance.  When combined with
            by(varname), these two parameters are distinctly estimated for
            each level of varname.  Because you are modeling a within-group
            covariance, varname must be constant within lowest-level groups.

        ar # assumes that within-group errors have an autoregressive (AR)
            structure of order #; ar 1 is the default.  The t(varname) option
            is required, where varname is an integer-valued time variable
            used to order the observations within groups and to determine the
            lags between successive observations.  Any nonconsecutive time
            values will be treated as gaps.  For this structure, # + 1
            parameters are estimated (# AR coefficients and one overall error
            variance).  restype ar may be combined with by(varname), but
            varname must be constant within groups.

        ma # assumes that within-group errors have a moving average (MA)
            structure of order #; ma 1 is the default.  The t(varname) option
            is required, where varname is an integer-valued time variable
            used to order the observations within groups and to determine the
            lags between successive observations.  Any nonconsecutive time
            values will be treated as gaps.  For this structure, # + 1
            parameters are estimated (# MA coefficients and one overall error
            variance).  restype ma may be combined with by(varname), but
            varname must be constant within groups.

        unstructured is the most general structure; it estimates distinct
            variances for each within-group error and distinct covariances
            for each within-group error pair.  The t(varname) option is
            required, where varname is a nonnegative-integer-valued variable
            that identifies the observations within each group.  The groups
            may be unbalanced in that not all levels of t() need to be
            observed within every group, but you may not have repeated t()
            values within any particular group.  When you have p levels of
            t(), then p(p+1)/2 parameters are estimated.  restype
            unstructured may be combined with by(varname), but varname must
            be constant within groups.

        banded # is a special case of unstructured that restricts estimation
            to the covariances within the first # off-diagonals and sets the
            covariances outside this band to 0.  The t(varname) option is
            required, where varname is a nonnegative-integer-valued variable
            that identifies the observations within each group.  # is an
            integer between 0 and p-1, where p is the number of levels of
            t().  By default, # is p-1; that is, all elements of the
            covariance matrix are estimated.  When # is 0, only the diagonal
            elements of the covariance matrix are estimated.  restype banded
            may be combined with by(varname), but varname must be constant
```

within groups.

toeplitz # assumes that within-group errors have Toeplitz structure of order #, for which correlations are constant with respect to time lags less than or equal to # and are 0 for lags greater than #. The t(varname) option is required, where varname is an integer-valued time variable used to order the observations within groups and to determine the lags between successive observations. # is an integer between 1 and the maximum observed lag (the default). Any nonconsecutive time values will be treated as gaps. For this structure, # + 1 parameters are estimated (# correlations and one overall error variance). restype toeplitz may be combined with by(varname), but varname must be constant within groups.

exponential is a generalization of the AR covariance model that allows for unequally spaced and noninteger time values. The t(varname) option is required, where varname is real-valued. For the exponential covariance model, the correlation between two errors is the parameter rho, raised to a power equal to the absolute value of the difference between the t() values for those errors. For this structure, two parameters are estimated (the correlation parameter rho and one overall error variance). restype exponential may be combined with by(varname), but varname must be constant within groups.

residual_options are by(varname) and t(varname).

by(varname) is for use within the residuals() option and specifies that a set of distinct residual-error parameters be estimated for each level of varname. In other words, you use by() to model heteroskedasticity.

t(varname) is for use within the residuals() option to specify a time variable for the ar, ma, toeplitz, and exponential structures, or to identify the observations when restype is unstructured or banded.