SC708: Hierarchical Linear Modeling
Instructor: Natasha Sarkisian
Class notes: HLM Diagnostics

Like OLS, HLM models rely on certain assumptions that have to be satisfied in order for regression coefficients to be unbiased and efficient estimates of the parameters of interest. Therefore, it is important to watch out for possible assumption violations and to take steps to prevent them. We will address the issues of model specification, homoscedasticity, normality of level 1 and level 2 residuals, and linearity.

1. Model specification.

In HLM models, the issue of model specification concerns two main questions: (1) Did we include the right fixed effects? (2) Did we include the right random components? As we discussed, when specifying your model, you should rely heavily on your theory as well as utilize hypothesis testing. But there are some additional steps you can take to prevent model misspecification.

To prevent misspecification of fixed effects:

- Consider including aggregates of level 1 variables. It is always possible that what appears to be an effect of a level 1 variable is, in reality, an effect of its level 2 aggregate. The only way to test is to introduce such an aggregate. So far, we discussed aggregates to the mean, but sometimes, it is also possible to use group-level standard deviations. For example, you can use MEANSES to indicate the average level of SES in the school and SESDEV (within-school standard deviation) to indicate how diverse each school is in terms of SES. Such diversity may have an impact above and beyond the impact of the average level.
- Consider including level 2 predictors of level 1 slopes if you find significant variation in these slopes
- If the proportion of explained variance (R-squared) is substantially reduced when you add a fixed effect, that can be a sign of misspecification.
- Sometimes a fixed effect misspecification (e.g., a nonlinearity) can lead to a misspecification of the random effects (excluded curvilinear effect may show up as a significant variance component for the slope). We will return to the issue of linearity below.

To prevent the misspecification problems in terms of random components:

- Always test whether each of your level 1 slopes varies across level 2 units (i.e., try to estimate each slope as random). However, you have to be careful not to "overtax" your data.
- The number of iterations can be diagnostic – if the data are highly informative, the algorithm will converge rapidly (e.g. in less than 10 iterations). In contrast, if the model has an extensive number of random effects and the data are relatively sparse, hundreds of iterations may be needed. In general, you should be cautious in specifying level-1 coefficients as random – as the number of random effects grows, the number of variances/covariances to be estimated increases even faster (for m random predictors,

there are 1+m(m+1)/2 variance covariance components). As the number of random effects grows, significantly mode information is required to obtain reasonable estimates of variance/covariance components. The maximum depends on a number of factors: the magnitude of the variance components, the degree of intercorrelation among the random effects, the magnitude of sigma squared, and other characteristics of the data.

- If there are high correlations among level-1 coefficients (i.e., slopes for different variables—correlations with the intercept are ok), the model must be simplified. There are a number of ways of dealing with it. You can, for example, use factor analysis to form scales and reduce the number of variables. You can also constrain one or more random effects to be zero (i.e. keep only the fixed effect for that variable), thus eliminating the correlation. This works well if that random effect is negligible.

2. Multicollinearity

Like regular OLS, HLM models can be affected by multicollinearity. There are no tools to check for it in HLM, so you should do some tests before you import your data into HLM. You should check correlations among your independent variables as well as variance inflation factors (VIFs) in another statistical program. E.g. in Stata:

```
. pwcorr  mathach ses female meanses sector size

             |  mathach      ses   female  meanses   sector     size
-------------+------------------------------------------------------
     mathach |   1.0000
         ses |   0.3608   1.0000
      female |  -0.1231  -0.0679   1.0000
     meanses |   0.3437   0.5306  -0.0589   1.0000
      sector |   0.2040   0.1896   0.0065   0.3573   1.0000
        size |  -0.0506  -0.0673  -0.0388  -0.1268  -0.4237   1.0000

. reg  mathach ses female meanses sector size

      Source |       SS       df       MS              Number of obs =    7185
-------------+------------------------------           F(  5,  7179) =  315.35
       Model |  61205.6611     5  12241.1322           Prob > F      =  0.0000
    Residual |  278671.273  7179  38.8175614           R-squared     =  0.1801
-------------+------------------------------           Adj R-squared =  0.1795
       Total |  339876.934  7184  47.3102637           Root MSE      =  6.2304

------------------------------------------------------------------------------
     mathach |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         ses |   2.148034   .1113801    19.29   0.000     1.929697    2.366372
      female |  -1.321295   .1478042    -8.94   0.000    -1.611034   -1.031555
     meanses |   2.889622   .2206451    13.10   0.000     2.457093    3.322151
      sector |   1.503238   .1724585     8.72   0.000     1.165169    1.841308
        size |   .0003457   .0001345     2.57   0.010     .0000821    .0006093
       _cons |   12.32108   .2222729    55.43   0.000     11.88536     12.7568
------------------------------------------------------------------------------

. vif

    Variable |       VIF       1/VIF
-------------+----------------------
     meanses |      1.54    0.648947
         ses |      1.39    0.717093
      sector |      1.38    0.726733
```

```
        size |      1.22    0.818586
      female |      1.01    0.992362
-------------+----------------------
    Mean VIF |      1.31
```

Different researchers advocate for different cutoff points for VIF. Some say that if any one of VIF values is larger than 4, there are some multicollinearity problems associated with that variable. Others use cutoffs of 5 or even 10.

It is also useful to check level 2 separately:

```
. bysort id: egen mathachm=mean(mathach)

. reg  mathach meanses sector size if case==1

      Source |       SS       df       MS              Number of obs =     160
-------------+------------------------------           F(  3,   156) =    8.30
       Model |  1087.63418     3  362.544726           Prob > F      =  0.0000
    Residual |  6814.85021   156  43.6849372           R-squared     =  0.1376
-------------+------------------------------           Adj R-squared =  0.1210
       Total |  7902.48439   159  49.7011597           Root MSE      =  6.6095

------------------------------------------------------------------------------
     mathach |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     meanses |   4.994022   1.355516     3.68   0.000     2.316488    7.671556
      sector |   2.201746   1.253382     1.76   0.081    -.2740427    4.677536
        size |   .0002882    .000934     0.31   0.758    -.0015568    .0021331
       _cons |    11.4176   1.455231     7.85   0.000     8.543103     14.2921
------------------------------------------------------------------------------

. vif

    Variable |       VIF       1/VIF
-------------+----------------------
      sector |      1.38    0.726829
        size |      1.22    0.819769
     meanses |      1.15    0.871633
-------------+----------------------
    Mean VIF |      1.25
```

Once you have your data in HLM and run your models, you should also watch out for potential signs of multicollinearity (e.g., large coefficients in opposite directions, high standard errors).

3. Homoscedasticity.

In HLM, the level-1 error terms should have equal variance across level-2 units (the assumption of homoscedasticity or homogeneity of variance) – e.g., all schools should have variances equal to the other schools in the sample. To test the homogeneity of variance assumption, under Other Settings →Hypotheses testing, select “Test homogeneity of level-1 variance.” Then run the model. The output looks like this:

```
Summary of the model specified (in equation format)
 ---------------------------------------------------

Level-1 Model
```

```
      Y = B0 + B1*(SES) + R

Level-2 Model
      B0 = G00 + G01*(SECTOR) + G02*(MEANSES) + U0
      B1 = G10 + G11*(SECTOR) + G12*(MEANSES)

Iterations stopped due to small change in likelihood function

******* ITERATION 6 *******

 Sigma_squared =     36.76611

 Tau
 INTRCPT1,B0      2.37524

Tau (as correlations)
 INTRCPT1,B0  1.000
 ----------------------------------------------------
  Random level-1 coefficient   Reliability estimate
 ----------------------------------------------------
  INTRCPT1, B0                        0.732
 ----------------------------------------------------
The value of the likelihood function at iteration 6 = -2.325148E+004

 The outcome variable is  MATHACH

 Final estimation of fixed effects:
 ----------------------------------------------------------------------------
                                       Standard             Approx.
    Fixed Effect         Coefficient   Error      T-ratio   d.f.     P-value
 ----------------------------------------------------------------------------
 For       INTRCPT1, B0
    INTRCPT2, G00          12.095250   0.198627   60.894       157    0.000
      SECTOR, G01           1.224401   0.306117    4.000       157    0.000
     MEANSES, G02           5.336698   0.368978   14.463       157    0.000
 For      SES slope, B1
    INTRCPT2, G10           2.935664   0.150690   19.482      7179    0.000
      SECTOR, G11          -1.642102   0.233097   -7.045      7179    0.000
     MEANSES, G12           1.044120   0.291042    3.588      7179    0.001
 ----------------------------------------------------------------------------

 The outcome variable is  MATHACH

 Final estimation of fixed effects
 (with robust standard errors)
 ----------------------------------------------------------------------------
                                       Standard             Approx.
    Fixed Effect         Coefficient   Error      T-ratio   d.f.     P-value
 ----------------------------------------------------------------------------
 For       INTRCPT1, B0
    INTRCPT2, G00          12.095250   0.173679   69.641       157    0.000
      SECTOR, G01           1.224401   0.308507    3.969       157    0.000
     MEANSES, G02           5.336698   0.334617   15.949       157    0.000
 For      SES slope, B1
    INTRCPT2, G10           2.935664   0.147576   19.893      7179    0.000
      SECTOR, G11          -1.642102   0.237223   -6.922      7179    0.000
     MEANSES, G12           1.044120   0.332897    3.136      7179    0.002
 ----------------------------------------------------------------------------

 Final estimation of variance components:
 -----------------------------------------------------------------------------
 Random Effect           Standard      Variance     df    Chi-square  P-value
                         Deviation     Component
```

```
 ----------------------------------------------------------------------------
 INTRCPT1,       U0        1.54118       2.37524   157     604.29895    0.000
  level-1,       R         6.06351      36.76611
 ----------------------------------------------------------------------------

 Statistics for current covariance components model
 --------------------------------------------------
 Deviance                       = 46502.952743
 Number of estimated parameters = 2

 Test of homogeneity of level-1 variance
 ----------------------------------------
 Chi-square statistic         =    244.08638
 Number of degrees of freedom =  159
 P-value                      = 0.000
```

This output indicates that for this model, the assumption was violated, so the variance is heterogeneous. Note that this method relies on fitting separate OLS regressions in each of the groups, so there should be a substantial number of groups with a relatively large number of cases in each group in order for this test to be accurate.

Heterogeneity of variance can be a nuisance, or it can be substantively interesting. When it is a nuisance, the causes can be:

- One or more important level-1 predictors may have been omitted from the model.
- The effects of a level-1 predictor that is random or nonrandomly varying have been erroneously treated as fixed.
- Dependent variable is severely skewed.
- One (or more) of the independent variables has a nonlinear relationship to the dependent variable that we failed to model correctly.
- There are outliers or bad data.

Let's try and free the slope of SES:

```
Test of homogeneity of level-1 variance
 ----------------------------------------
 Chi-square statistic         =    245.76576
 Number of degrees of freedom =  159
 P-value                      = 0.000
```

Still a problem. We could consider examining issues of normality or linearity, but for now, let's try to think about heterogeneity as substantively interesting and model it using level 1 predictors – to see whether there are some predictors that seem to explain differential level 1 variance:

```
RESULTS FOR HETEROGENEOUS SIGMA-SQUARED
(macro iteration 4)

 Var(R) = Sigma_squared and
 log(Sigma_squared) = alpha0 + alpha1(MINORITY) + alpha2(FEMALE) + alpha3(SES)

 Model for level-1 variance
 -------------------------------------------------------------------
                                    Standard
    Parameter        Coefficient      Error       Z-ratio   P-value
 -------------------------------------------------------------------
 INTRCPT1    ,alpha0      3.67952      0.026787    137.360     0.000
```

```
 MINORITY    ,alpha1    -0.04998      0.038205     -1.308     0.191
   FEMALE    ,alpha2    -0.12185      0.033969     -3.587     0.001
      SES    ,alpha3     0.00129      0.026087      0.050     0.961
 --------------------------------------------------------------------
Summary of Model Fit
 -------------------------------------------------------------------
 Model                                Number of         Deviance
                                      Parameters
 -------------------------------------------------------------------
 1. Homogeneous sigma_squared             10          46494.59261
 2. Heterogeneous sigma_squared           13          46480.44853
 -------------------------------------------------------------------
 Model Comparison                 Chi-square       df    P-value
 -------------------------------------------------------------------
 Model 1 vs Model 2                  14.14407       3     0.003

 Test of homogeneity of level-1 variance
 ----------------------------------------
 Chi-square statistic         =    237.93937
 Number of degrees of freedom =  159
 P-value                      = 0.000
```

Looks like gender explains some heterogeneity – there is lower amount of unexplained variance in math achievement among girls. Further, heterogenous model has significantly lower deviance than the homogenous model. There is still some unexplained heterogeneity left, however. Note that HLM does not allow us to model the residual variance using level 2 (school) characteristics; another multilevel analysis program, MLwiN, does.

But so far, we did not have gender and minority variables in the model itself. So let's add them and see what happens. Let's try to add other level 1 predictors to the model:

LEVEL 1 MODEL

$$MATHACH_{ij} = \beta_{0j} + \beta_{1j}(MINORITY_{ij}) + \beta_{2j}(FEMALE_{ij}) + \beta_{3j}(SES_{ij} - \overline{SES}_{\cdot j}) + r_{ij}$$

LEVEL 2 MODEL

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(SECTOR_j) + \gamma_{02}(MEANSES_j - \overline{MEANSES}_{\cdot}) + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + \gamma_{11}(SECTOR_j) + \gamma_{12}(MEANSES_j - \overline{MEANSES}_{\cdot}) + u_{1j}$$
$$\beta_{2j} = \gamma_{20} + \gamma_{21}(SECTOR_j) + \gamma_{22}(MEANSES_j - \overline{MEANSES}_{\cdot}) + u_{2j}$$
$$\beta_{3j} = \gamma_{30} + \gamma_{31}(SECTOR_j) + \gamma_{32}(MEANSES_j - \overline{MEANSES}_{\cdot})$$

```
 Sigma_squared =     35.33424

 Standard Error of Sigma_squared =      0.60442

 Tau
 INTRCPT1,B0      2.50575      -0.11413      -1.15746
 MINORITY,B1     -0.11413       0.98114       0.18246
   FEMALE,B2     -1.15746       0.18246       0.86444

 Standard Errors of Tau
 INTRCPT1,B0      0.53220       0.50638       0.44703
 MINORITY,B1      0.50638       0.71802       0.47533
   FEMALE,B2      0.44703       0.47533       0.52350

Tau (as correlations)
```

```
 INTRCPT1,B0  1.000 -0.073 -0.786
 MINORITY,B1 -0.073  1.000  0.198
   FEMALE,B2 -0.786  0.198  1.000
 ----------------------------------------------------

  Random level-1 coefficient   Reliability estimate
 ----------------------------------------------------
  INTRCPT1, B0                       0.504
  MINORITY, B1                       0.121
    FEMALE, B2                       0.195
 ----------------------------------------------------
Note: The reliability estimates reported above are based on only 100 of 160
units that had sufficient data for computation.  Fixed effects and variance
components are based on all the data.

The outcome variable is  MATHACH
 Final estimation of fixed effects
 (with robust standard errors)
 ----------------------------------------------------------------------------
                                      Standard             Approx.
    Fixed Effect         Coefficient   Error      T-ratio   d.f.     P-value
 ----------------------------------------------------------------------------
 For       INTRCPT1, B0
    INTRCPT2, G00          13.493279   0.218255    61.824       157    0.000
      SECTOR, G01           1.179093   0.395065     2.985       157    0.004
     MEANSES, G02           4.257586   0.500135     8.513       157    0.000
 For MINORITY slope, B1
    INTRCPT2, G10          -3.800761   0.329691   -11.528       157    0.000
      SECTOR, G11           1.859578   0.462451     4.021       157    0.000
     MEANSES, G12          -0.742032   0.490089    -1.514       157    0.132
 For   FEMALE slope, B2
    INTRCPT2, G20          -1.269669   0.219399    -5.787       157    0.000
      SECTOR, G21           0.017264   0.398333     0.043       157    0.966
     MEANSES, G22          -0.008239   0.435061    -0.019       157    0.985
 For      SES slope, B3
    INTRCPT2, G30           2.478999   0.144731    17.128      7173    0.000
      SECTOR, G31          -1.311409   0.232813    -5.633      7173    0.000
     MEANSES, G32           0.995932   0.315195     3.160      7173    0.002
 ----------------------------------------------------------------------------
 Final estimation of variance components:
 -----------------------------------------------------------------------------
 Random Effect           Standard      Variance     df    Chi-square  P-value
                         Deviation     Component
 -----------------------------------------------------------------------------
 INTRCPT1,       U0        1.58295       2.50575    97     200.38556    0.000
 MINORITY slope, U1        0.99053       0.98114    97     112.50013    0.135
   FEMALE slope, U2        0.92975       0.86444    97     125.27254    0.028
  level-1,       R         5.94426      35.33424
 -----------------------------------------------------------------------------
Note: The chi-square statistics reported above are based on only 100 of 160
units that had sufficient data for computation.  Fixed effects and variance
components are based on all the data.

 Statistics for current covariance components model
 --------------------------------------------------
 Deviance                       = 46223.669899
 Number of estimated parameters = 19

 Test of homogeneity of level-1 variance
 ----------------------------------------
 Chi-square statistic         =    108.42203
 Number of degrees of freedom =   99
 P-value                      = 0.243
```

We no longer detect significant homogeneity using this test. Note that now only 100 out of 160 schools are used for calculating the reliability estimates and the chi-square statistics for variance components – that's because for some schools, we do not have sufficient numbers of boys and girls to reliably calculate the gender slope.

But, if we once again explore if level 1 variance can be explained by our level 1 predictors, we will still observe some relationships:

```
RESULTS FOR HETEROGENEOUS SIGMA-SQUARED
(macro iteration 4)

 Var(R) = Sigma_squared and
 log(Sigma_squared) = alpha0 + alpha1(MINORITY) + alpha2(FEMALE) + alpha3(SES)

 Model for level-1 variance
 --------------------------------------------------------------------
                                     Standard
    Parameter        Coefficient     Error       Z-ratio   P-value
 --------------------------------------------------------------------
 INTRCPT1    ,alpha0     3.65317      0.026946    135.572     0.000
 MINORITY    ,alpha1    -0.13144      0.038616     -3.404     0.001
   FEMALE    ,alpha2    -0.10502      0.034193     -3.071     0.003
      SES    ,alpha3    -0.00702      0.026020     -0.270     0.787
 --------------------------------------------------------------------

Summary of Model Fit
 -------------------------------------------------------------------
 Model                               Number of         Deviance
                                     Parameters
 -------------------------------------------------------------------
 1. Homogeneous sigma_squared            19          46223.67031
 2. Heterogeneous sigma_squared          22          46203.21463
 -------------------------------------------------------------------
 Model Comparison                Chi-square       df    P-value
 -------------------------------------------------------------------
 Model 1 vs Model 2                 20.45568       3     0.000
```

So we observe less unexplained variance among girls and minorities. We might want to explore what explains the higher variance among boys and whites (e.g., we could consider an interaction term of SES with gender and minority status variables – we'd have to create it outside of HLM). If you find a heteroscedasticity problem or a distributional problem (i.e., non-normality) but cannot correct it, you can rely on robust standard errors.

HLM produces two final tables of fixed effects: one with regular standard errors and one with robust standard errors. Robust standard errors are standard errors that are relatively insensitive to misspecification at the levels of the model and the distributional assumptions at each level. If the robust and model-based standard errors differ substantially, that suggests that you have some problem with normality, homoscedasticity, or linearity, and you should further investigate those HLM assumptions. If it is not possible to correct the problem, you can report robust standard errors.

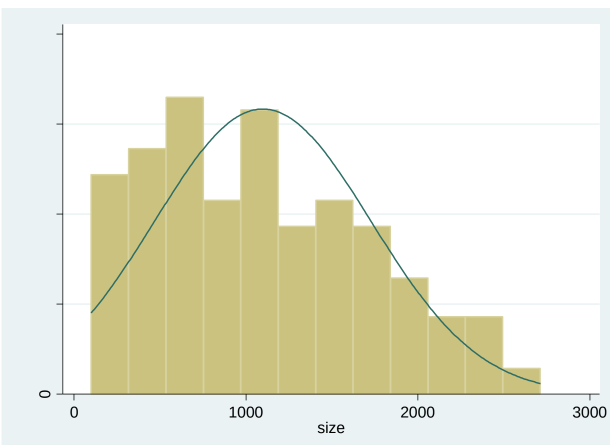
Note, however, that the robust standard errors should be trusted only when the number of higher-level units is moderately large relative to the number of explanatory variables at the higher level.

4. Normality

HLM models assume that the level-1 and level 2 error terms are normally distributed. To make sure this assumption will be met, it is important to do some preliminary data screening before importing data into HLM. It is especially important to ensure that your dependent variable distribution is as close to normal as possible, but independent variables should be checked as well. If substantial deviations from normality are identified, consider fixing them with a transformation. Note that when examining normality of level 2 variables, you should either have a separate level 2 file or you should limit your analysis to one record per higher level unit.

To do the latter, in Stata we could create a within-school id for individuals and then do our examination taking only the first case in each school:

```
. bysort id: gen case=_n

. histogram size if case==1
(bin=12, start=100, width=217.75)
```

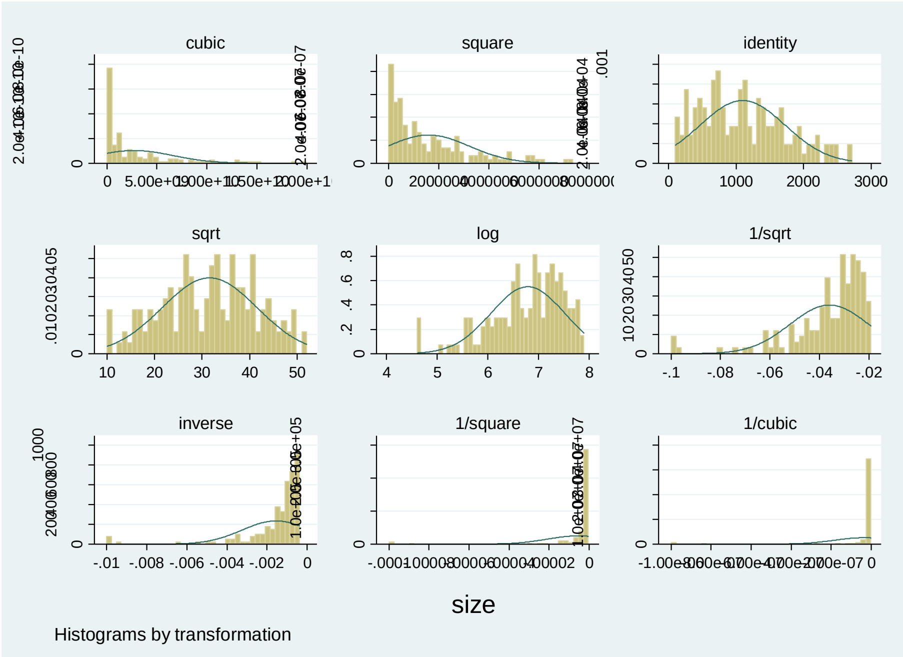


Looks like a right skew; to find a transformation:

```
. ladder size if case==1
```

Transformation	formula	chi2(2)	P(chi2)
cubic	size^3	60.02	0.000
square	size^2	31.36	0.000
identity	size	8.37	0.015
square root	sqrt(size)	7.18	0.028
log	log(size)	16.55	0.000
1/(square root)	1/sqrt(size)	58.10	0.000
inverse	1/size	.	0.000
1/square	1/(size^2)	.	0.000
1/cubic	1/(size^3)	.	0.000

```
. gladder size if case==1
```



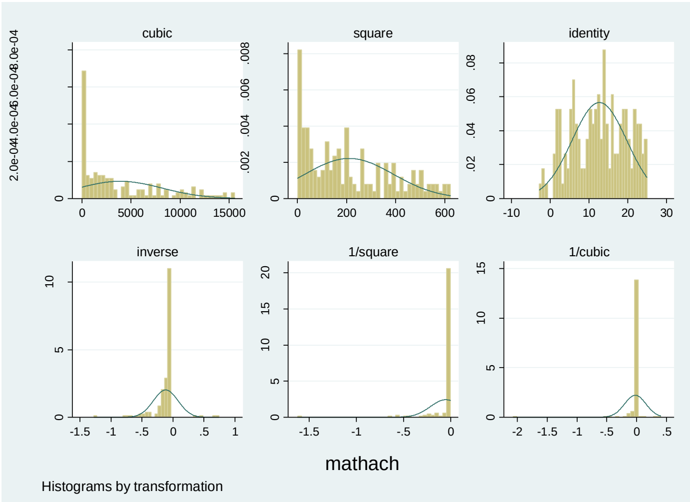
Square root looks the best, so we would generate it and then later on import that transformed variable into HLM:

```
. gen sizesqrt=sqrt(size)
```

If a variable contains zero or negative values, you need to add a constant to it before looking for transformations (such that all values of the variable become larger than zero). For example:

```
. sum mathach

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
     mathach |      7185    12.74785    6.878246     -2.832     24.993

. ladder mathach if case==1

Transformation         formula                  chi2(2)       P(chi2)
------------------------------------------------------------------
cubic                  mathach^3                 19.29         0.000
square                 mathach^2                 16.12         0.000
identity               mathach                   18.12         0.000
square root            sqrt(mathach)                 .             .
log                    log(mathach)                  .             .
1/(square root)        1/sqrt(mathach)               .             .
inverse                1/mathach                 49.83         0.000
1/square               1/(mathach^2)                 .         0.000
1/cubic                1/(mathach^3)                 .         0.000
```

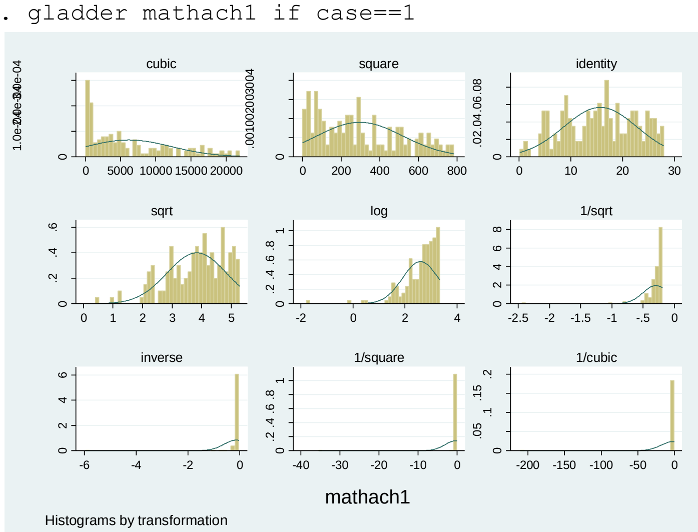
```
. gladder mathach if case==1
```




```
. gen mathach1=mathach+3

. ladder mathach1 if case==1
```

Transformation	formula	chi2(2)	P(chi2)
cubic	mathach1^3	16.03	0.000
square	mathach1^2	18.22	0.000
identity	mathach1	18.12	0.000
square root	sqrt(mathach1)	11.27	0.004
log	log(mathach1)	.	0.000
1/(square root)	1/sqrt(mathach1)	.	0.000
inverse	1/mathach1	.	0.000
1/square	1/(mathach1^2)	.	0.000
1/cubic	1/(mathach1^3)	.	0.000

```
. gladder mathach1 if case==1
```



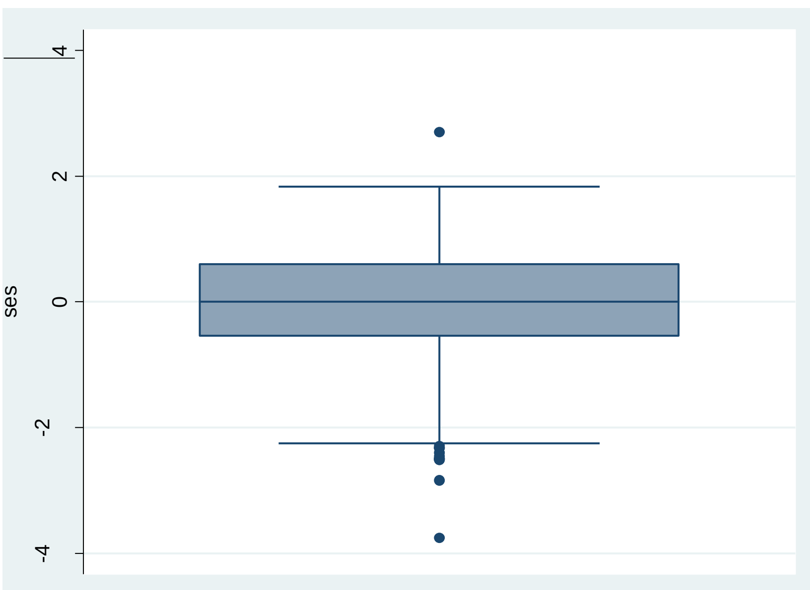
If your sample size is large, everything will be significantly different from normal, so you should either rely on graphical examination (gladder) or randomly select a subsample of your dataset and do this type of analysis for that subsample.

If as variable is negatively skewed, you might have an easier time finding a transformation for it after reversing it. To reverse the variable and yet keep all the values positive, you can subtract it from its maximum value +1; for example:

```
. sum mathach

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
     mathach |      7185    12.74785    6.878246     -2.832     24.993

. gen mathachr=24.993+1-mathach

. sum mathachr

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
    mathachr |      7185    13.24515    6.878246   .9999999     28.825
```

As you are examining normality, pay attention to outliers as well – sometimes, it is useful to top-code or bottom-code outliers in addition to or instead of transforming a variable.

```
. graph box ses
```




```
. sum ses, detail
                             ses
-------------------------------------------------------------
      Percentiles      Smallest
 1%       -1.848         -3.758
 5%       -1.318         -2.838
10%       -1.038         -2.508       Obs                7185
25%        -.538         -2.508       Sum of Wgt.        7185

50%         .002                      Mean           .0001434
                        Largest       Std. Dev.      .7793552
75%         .602          1.732
90%        1.022          1.762       Variance       .6073945
95%        1.212          1.832       Skewness      -.2281447
99%        1.512          2.692       Kurtosis       2.620279
```

```
. gen ses1=ses

. replace ses1=1.9 if ses>1.9 & ses<.
(1 real change made)

. replace ses1=-2.9 if ses<-2.9
(1 real change made)
```

Never top-code or bottom-code more than 5% of the distribution; better yet, do 1% or less. Sometimes transformation might be a better way to bring in outliers so consider both options or a combination of them.

If you do a good job dealing with normality problems and with outliers during preliminary screening, you should not run into problems with multivariate normality. Still, we need to check both level 1 and level 2 residuals for normality. Let's estimate a model, obtain residuals, and inspect them:

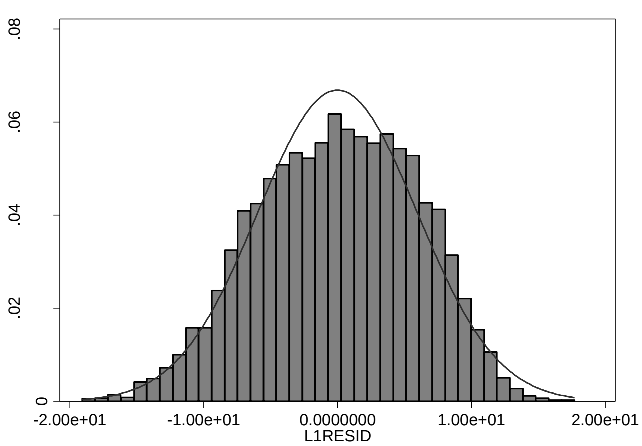
```
  The model specified for the fixed effects was:
 ----------------------------------------------------
   Level-1                  Level-2
   Coefficients             Predictors
 ----------------------   ---------------
        INTRCPT1, B0      INTRCPT2, G00
                            SECTOR, G01
$                          MEANSES, G02
     FEMALE slope, B1     INTRCPT2, G10
                            SECTOR, G11
$                          MEANSES, G12
 *      SES slope, B2     INTRCPT2, G20
                            SECTOR, G21
$                          MEANSES, G22

'*' - This level-1 predictor has been centered around its group mean.
'$' - This level-2 predictor has been centered around its grand mean.

Level-1 Model
      Y = B0 + B1*(FEMALE) + B2*(SES) + R

Level-2 Model
      B0 = G00 + G01*(SECTOR) + G02*(MEANSES) + U0
      B1 = G10 + G11*(SECTOR) + G12*(MEANSES) + U1
      B2 = G20 + G21*(SECTOR) + G22*(MEANSES) + U2

The outcome variable is  MATHACH

 Final estimation of fixed effects
 (with robust standard errors)
 ----------------------------------------------------------------------------
                                    Standard             Approx.
    Fixed Effect         Coefficient   Error      T-ratio   d.f.     P-value
 ----------------------------------------------------------------------------
 For       INTRCPT1, B0
    INTRCPT2, G00         12.728314   0.213807    59.532       157    0.000
      SECTOR, G01          1.182789   0.393223     3.008       157    0.004
     MEANSES, G02          5.206435   0.431487    12.066       157    0.000
 For   FEMALE slope, B1
    INTRCPT2, G10         -1.230407   0.221181    -5.563       157    0.000
      SECTOR, G11          0.075948   0.414157     0.183       157    0.855
     MEANSES, G12         -0.012379   0.419525    -0.030       157    0.977
 For      SES slope, B2
```

```
   INTRCPT2, G20          2.884798   0.145874    19.776       157    0.000
     SECTOR, G21         -1.605447   0.234757    -6.839       157    0.000
    MEANSES, G22          1.043186   0.328828     3.172       157    0.002
----------------------------------------------------------------------------

Final estimation of variance components:
-----------------------------------------------------------------------------
Random Effect           Standard      Variance     df    Chi-square  P-value
                        Deviation     Component
-----------------------------------------------------------------------------
INTRCPT1,       U0        1.72298       2.96867   120     303.16423    0.000
  FEMALE slope, U1        1.00971       1.01951   120     149.42515    0.035
     SES slope, U2        0.35080       0.12306   120     122.48752    0.420
 level-1,       R         6.02785      36.33494
-----------------------------------------------------------------------------
```

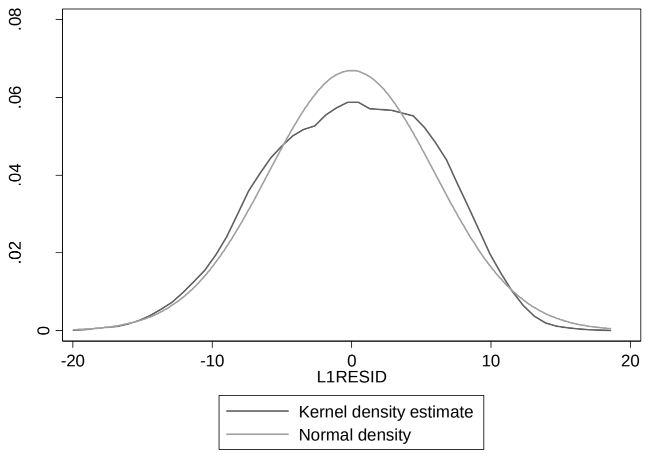
To check the distribution of level 1 error term, we should obtain a level-1 residuals file by clicking on Basic Settings → Level 1 residuals file, and then selecting the variables we want in that file and the type of output file we want (make sure the file extension corresponds to the type of file you selected –HLM does not automatically adjust that). I would advise to include all potentially interesting variables in that file, but you can also merge them later if you have person-level ID (in our case, we don't). Similarly, we obtain level-2 residuals as well.

We can now use the statistical software of our choice (e.g., Stata or SPSS) to check for normality of level-1 residuals. We can examine the distribution graphically as well as use formal statistical tests for normality.

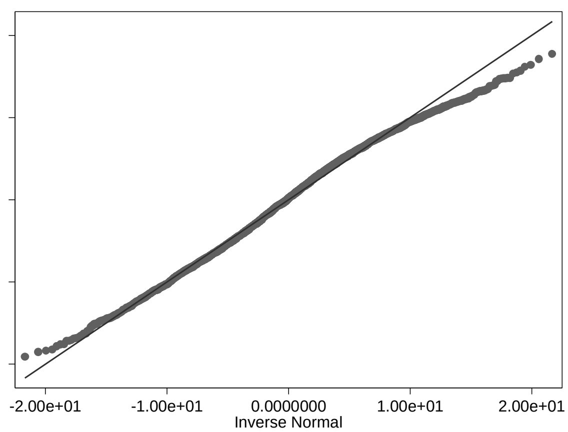
```
. sktest  L1RESID
                   Skewness/Kurtosis tests for Normality
                                                   ------- joint ------
    Variable |  Pr(Skewness)   Pr(Kurtosis)  adj chi2(2)    Prob>chi2
-------------+-------------------------------------------------------
     L1RESID |      0.000         0.000               .       0.0000

. histogram  L1RESID, normal
(bin=38, start=-19.084782, width=.96868813)
```




```
. kdensity  L1RESID, normal
```




```
. qnorm  L1RESID
```



We conclude that they look normal enough. Note that level 1 residuals file also contains predicted values – FITVAL—and predicted values of SIGMA (that is only relevant when we allow sigma to vary, as we did when we tried to model heterogeneity of variance above).

Next, let's test the multivariate normality of level 2 residuals. The level-2 residual file contains a single record per group unit. The first variable in this file contains the unit ID, followed by the number of level-1 units within that level-2 unit (denoted by NJ), and various summary statistics (CHIPCT through MDRSVAR).

MDIST is the Mahalanobis distance measure (i.e., the standardized squared distance of a unit from the center of a multidimensional distribution) for each level 2 group that measures the distance between residual estimates for each group relative to the expected distance based on the model (MDIST variable). CHIPCT contains the expected values for that distance. After MDIST, there are three estimates of the level-1 variability:

- The natural logarithm of the total standard deviation within each unit, LNTOTVAR.
- The natural logarithm of the residual standard deviation within each unit based on its least squares regression, OLSRSVAR. Note that this estimate exists only for those units which have sufficient data to compute level-1 OLS estimates.
- The MSRSVAR, the natural logarithm of the residual standard deviation from the final fitted fixed effects model.

The most useful thing for us, however, is residuals themselves. Here, we get OL (Ordinary Least Squares) residuals – residuals based on separately fitting a regular OLS model for each group, as well as EB residuals (empirical Bayes residuals) that are based on so-called shrinkage estimates of individual schools' regression equations – these are based on both group-specific regression coefficients and the overall coefficients for the entire model. Since these estimates are a weighted average of those components, the regression coefficients for each group are essentially shrunk towards the overall coefficient for the whole sample. (See p. 29-31 in Hox book for a good explanation of EB estimators.)

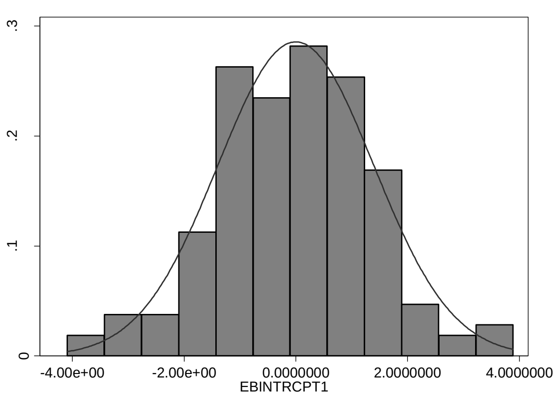
When we graphed level one slopes for each group using HLM graphing functions, we were relying on such shrinkage estimates for each group. So if we are interested in assessing what a predicted slope would be for a given group, we could take the overall coefficient and add the corresponding EB residual for that group. Note that we get one OLS residual variable and one EB residual variable for each intercept or slope that we are modeling as random; here we have three random effects and three residuals variables.

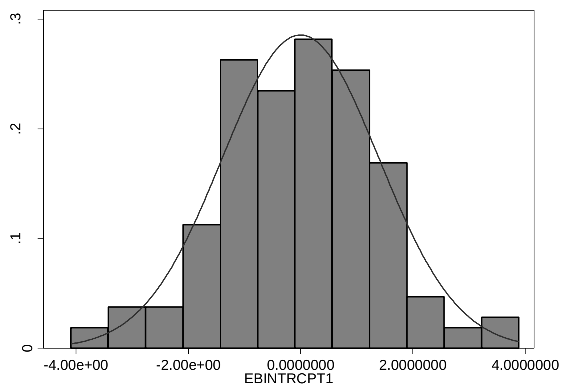
We also get the fitted or predicted values (FV) of the level-1 coefficients based on estimated level-2 models, and the EC coefficients, which are the sum of the fitted values plus the EB residuals. The posterior variances and covariances of the estimates of the intercept and the SES slopes are given next (PV00 to PVC10). Finally, the level-2 predictors used in the analysis plus those additional level-2 predictors that we requested for inclusion in the file are included.

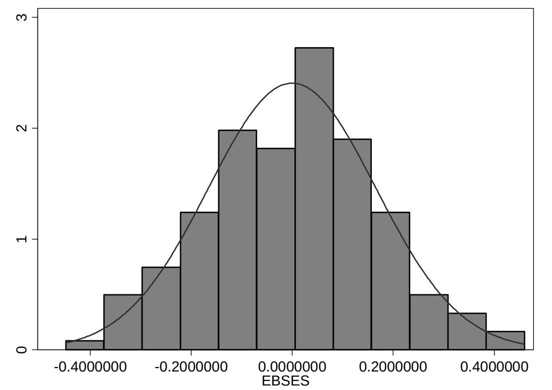
We will most heavily utilize EB residuals. First, we can examine the normality of each set of residuals separately.

. We can examine normality for each of these:

```
. histogram  EBINTRCPT1, normal
(bin=12, start=-4.0914528, width=.6652911)

. histogram   EBFEMALE, normal
(bin=12, start=-1.3546074, width=.2191729)

. histogram   EBSES, normal
(bin=12, start=-.44807637, width=.07564943)
```

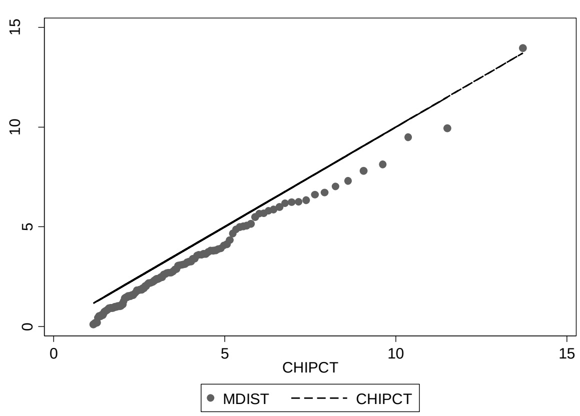




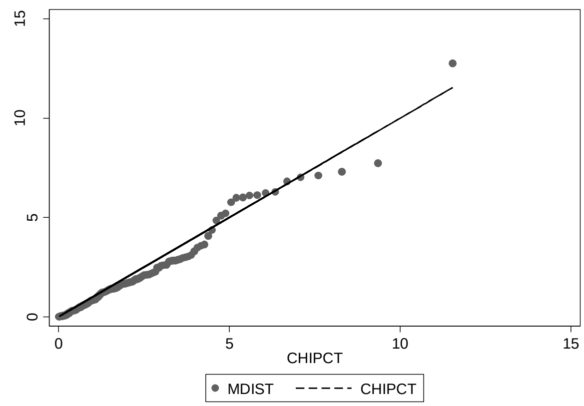


Second, we can assess multivariate normality by examining Mahalanobis distance measure (MDIST variable). Note that the units in the residual file are sorted in ascending order by MDIST. Analogous to univariate normal probability plotting, we can construct a Q-Q plot of MDIST vs. CHIPCT. CHIPCT contains the expected values; if the Q-Q plot resembles a 45 degree line, we have evidence that the random effects are distributed multivariate normal. In addition, the plot will help us detect outlying units (units with large MDIST values well above the 45 degree line).

```
. scatter   MDIST CHIPCT  CHIPCT, s(. i) c(. l)
```

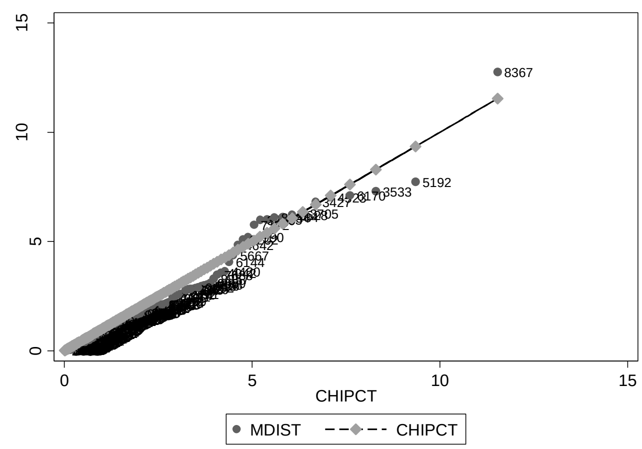


Here, we seem to have some model fit problem – all distance values are below the expected values, with one being above and much higher than others. We know that there is at least one a problem with the model – SES slope variance is not statistically significant, but we included it. If we fix that problem, the graph actually looks better:



If we want to examine which school has the highest MDIST:

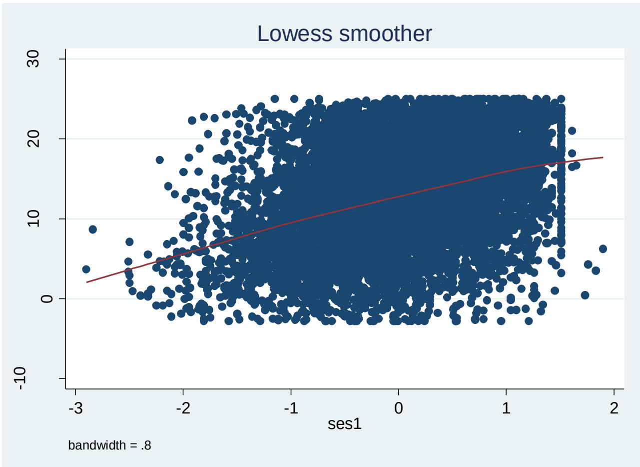
```
. scatter   MDIST CHIPCT  CHIPCT, mlabel(L2ID) c(. l)
```



5. Linearity

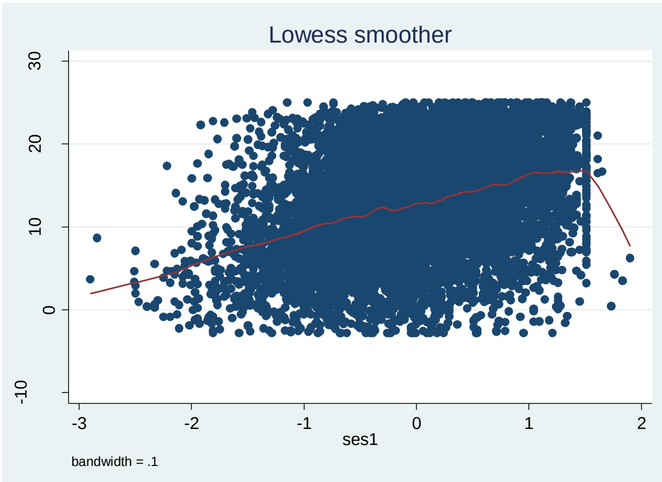
Before you get your data into HLM, it's also a good idea to examine the relationship of each independent variable to the dependent to assess its linearity. A good tool for such an examination is a lowess plot (called LOESS in SPSS) – that is, a scatterplot with locally weighted regression line (based on means or medians) going through it:

```
. lowess mathach ses1
```



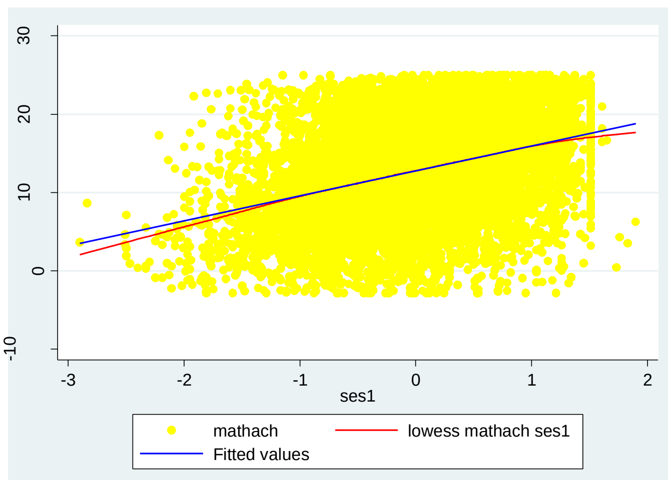
We can change bandwidth to make the curve less smooth (decrease the number) or smoother (increase the number):

```
. lowess mathach ses1, bwidth(.1)
```



We can also add a regression line to see the difference better:

```
. scatter mathach ses1, mcolor(yellow) || lowess mathach ses1, lcolor(red) ||
lfit mathach ses1, lcolor(blue)
```

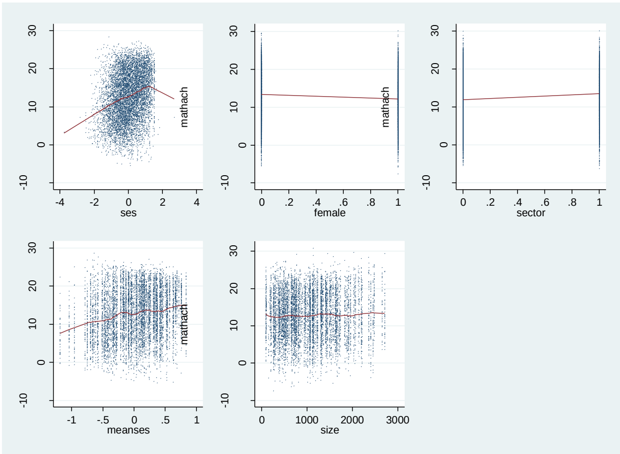


You can do an approximate test for multivariate linearity (based on OLS); in Stata, we could install a user-written mrunning program:

```
. search mrunning
Keyword search
        Keywords:  mrunning
          Search:  (1) Official help files, FAQs, Examples, SJs, and STBs
Search of official help files, FAQs, Examples, SJs, and STBs
SJ-5-3  gr0017  . . . . . . . . . . . . . . A multivariable scatterplot smoother
        (help mrunning, running if installed) . . . . P. Royston and N. J. Cox
        Q3/05   SJ 5(3):405--412
        presents an extension to running for use in a
        multivariable context
```

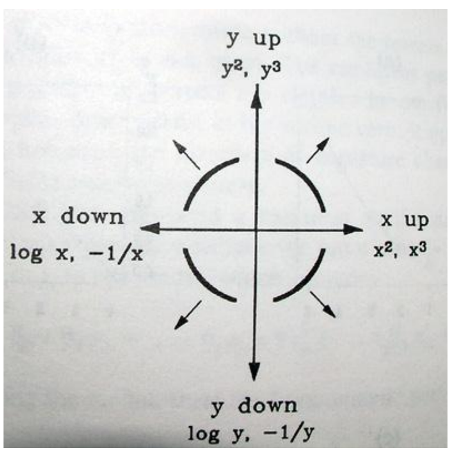
Click on gr0017 to install the program. Now we can use it:

```
. mrunning mathach ses female sector meanses size
```



If the relationship looks nonlinear on lowess plot, consider using transformations to fix it before importing data into HLM. (Note that if the relationship is too complex, sometimes we may choose to break up the corresponding independent variable into a series of dummies instead.)

<u>Monotone nonlinear relationship:</u> Power transformations can be used to linearize relationships if strong monotone nonlinearities are found. The following chart gives suggestions for transformations when the curve looks a certain way:



<u>Nonmonotone relationship:</u> For non-monotone relationships (e.g. parabola or cubic), use polynomial functions of the variable, e.g. ses and ses squared, etc. Note that when including variables that are generated using other variables already in the model (as in this case, or when we want to enter a product of two variables to model an interaction term), we should mean-center the variable outside of HLM (only if it is continuous; don't mean-center dichotomous variables!), and then square and/or cube the mean-centered variable. We will then include the mean-centered variable itself and its transformations into our HLM file and our models. For example, if we are dealing with a second level variable, we would get its mean across 160 level 2 cases by restricting the calculation to one case per level 2 unit:

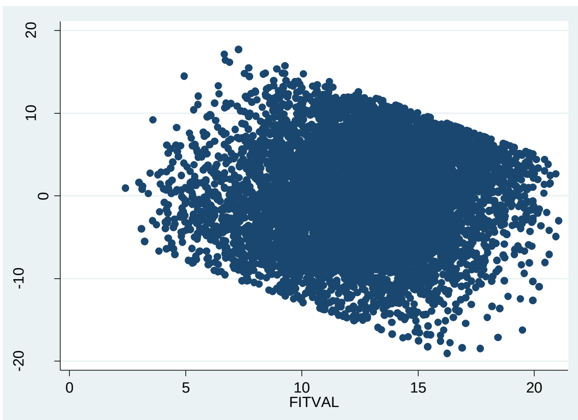
```
. sum size if case==1
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        size |       160    1097.825    629.5064        100       2713
```

```
. gen sizem=size-r(mean)
. gen sizem2=sizem^2
```

Oftentimes, the same transformation that helps with normality also will improve linearity, but that it is not always the case. Overall, linearity is more important to enforce than normality for a given variable, so if you end up with incompatible transformations, opt for the one improving linearity.

Once we estimated our HLM model and obtained residuals, we can inspect them to further assess linearity. First, we can assess the overall pattern by plotting level 1 residuals against predicted values; there should be no discernable pattern:

```
. scatter L1RESID FITVAL
```

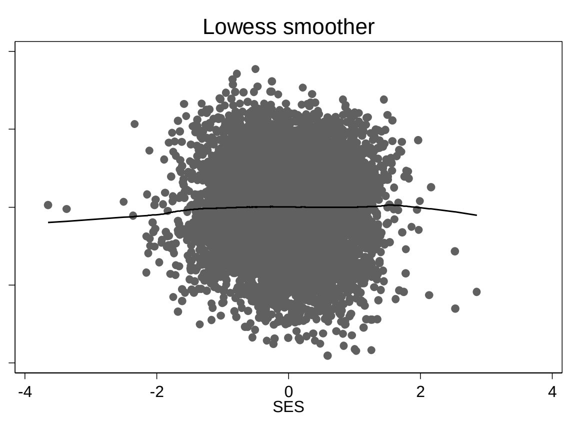


This does not look too good; indicates potential heteroscedasticity or nonlinearity problems.

To test the linearity assumption for continuous predictors, it is useful to plot residuals against each of the continuous dependent variable. To improve our ability to detect a curvilinear relationship, we will include a smoother in our plot using lowess command in Stata (in SPSS, you have to edit your plot to get a LOESS smoother).

In level 1 file:

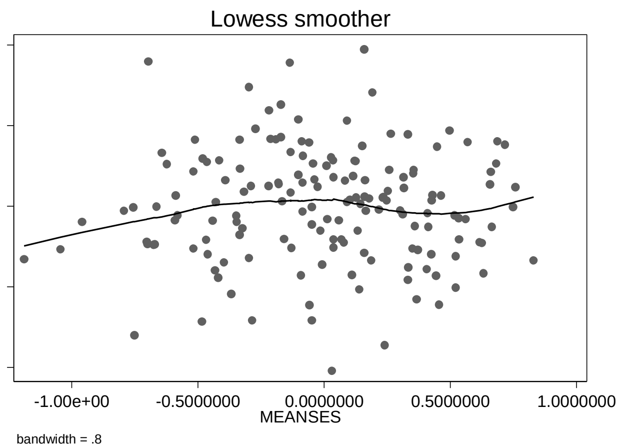
```
. lowess  L1RESID  SES
```

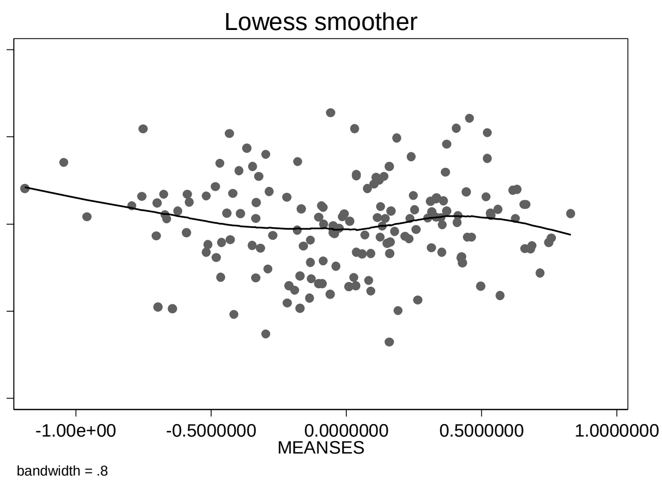


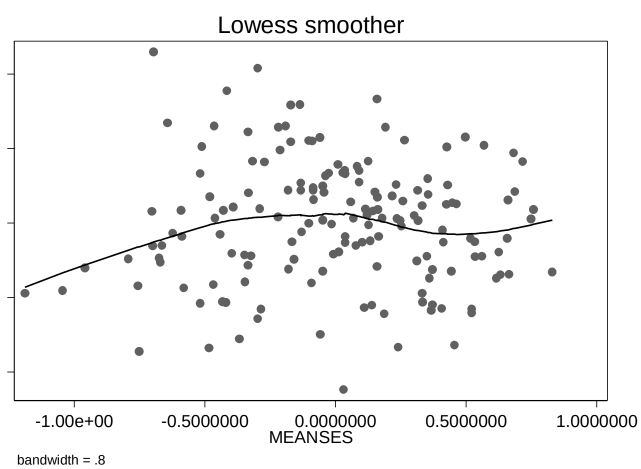
Looks more or less fine, but we do see those outliers we discussed above.

In level 2 file:

```
. lowess  EBINTRCPT1 MEANSES
. lowess   EBFEMALE MEANSES
. lowess    EBSES MEANSES
```

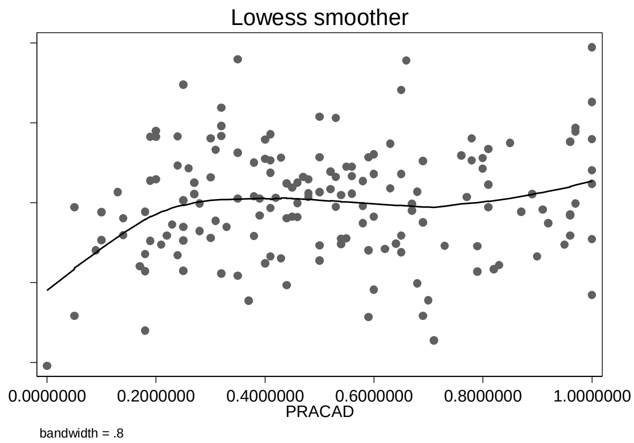


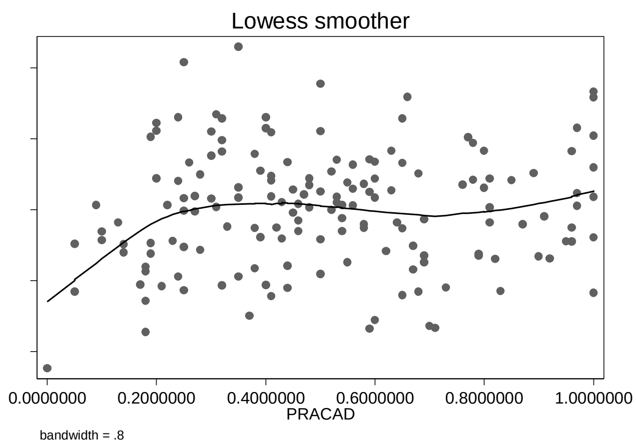


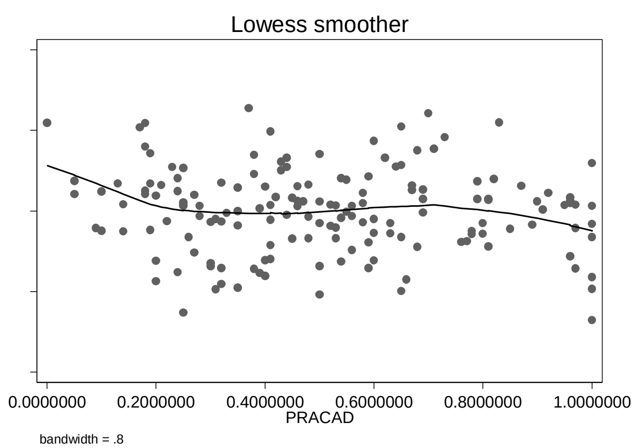


Based on these graphs, we could consider modeling nonlinear relationships with MEANSES (e.g. cubic).

We can also use such plots to search for potential other relationships and examine their shape, e.g. with PRACAD:







Based on graphs for MEANSES, looks like we need a quadratic and cubic term; let's try that. First, create them in Stata:

```
. use "C:\Users\sarkisin\hsb.dta", clear
```

```
. sum meanses
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
     meanses |      7185    .0061385    .4135539     -1.188       .831

. gen meansesm=meanses-r(mean)

. gen meanses2=meansesm^2

. gen meanses3=meansesm^3

. save "C:\Users\sarkisin\hsb_n.dta"
```

Next, we import these data with new level 2 variables into MDM and run the following model:

Summary of the model specified

Level-1 Model

$MATHACH_{ij} = \beta_{0j} + \beta_{1j}*(FEMALE_{ij}) + \beta_{2j}*(SES_{ij}) + r_{ij}$

Level-2 Model

$\beta_{0j} = \gamma_{00} + \gamma_{01}*(SECTOR_j) + \gamma_{02}*(MEANSESM_j) + \gamma_{03}*(MEANSES2_j) + \gamma_{04}*(MEANSES3_j) + u_{0j}$
$\beta_{1j} = \gamma_{10} + \gamma_{11}*(SECTOR_j) + \gamma_{12}*(MEANSESM_j) + \gamma_{13}*(MEANSES2_j) + \gamma_{14}*(MEANSES3_j) + u_{1j}$
$\beta_{2j} = \gamma_{20} + \gamma_{21}*(SECTOR_j) + \gamma_{22}*(MEANSESM_j) + \gamma_{23}*(MEANSES2_j) + \gamma_{24}*(MEANSES3_j) + u_{2j}$

Mixed Model

$MATHACH_{ij} = \gamma_{00} + \gamma_{01}*SECTOR_j + \gamma_{02}*MEANSESM_j + \gamma_{03}*MEANSES2_j$
$+ \gamma_{04}*MEANSES3_j$
$+ \gamma_{10}*FEMALE_{ij} + \gamma_{11}*SECTOR_j*FEMALE_{ij} + \gamma_{12}*MEANSESM_j*FEMALE_{ij} + \gamma_{13}*MEANSES2_j*FEMALE_{ij}$
$+ \gamma_{14}*MEANSES3_j*FEMALE_{ij}$
$+ \gamma_{20}*SES_{ij} + \gamma_{21}*SECTOR_j*SES_{ij} + \gamma_{22}*MEANSESM_j*SES_{ij} + \gamma_{23}*MEANSES2_j*SES_{ij}$
$+ \gamma_{24}*MEANSES3_j*SES_{ij}$
$+ u_{0j} + u_{1j}*FEMALE_{ij} + u_{2j}*SES_{ij} + r_{ij}$

Final Results - Iteration 196
Iterations stopped due to small change in likelihood function

$\sigma^2 = 36.33164$

τ

INTRCPT1,β_0	2.89652	-1.10566	0.15643
FEMALE,β_1	-1.10566	1.07071	-0.10789
SES,β_2	0.15643	-0.10789	0.03481

τ (as correlations)

INTRCPT1, β_0	1.000	-0.628	0.493
FEMALE, β_1	-0.628	1.000	-0.559
SES, β_2	0.493	-0.559	1.000

Random level-1 coefficient	Reliability estimate
INTRCPT1, β_0	0.562
FEMALE, β_1	0.226
SES, β_2	0.018

Note: The reliability estimates reported above are based on only 123 of 160 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.
The value of the log-likelihood function at iteration 196 = -2.321064E+004

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	12.907247	0.279465	46.186	155	<0.001
SECTOR, γ_{01}	1.283642	0.404468	3.174	155	0.002
MEANSESM, γ_{02}	1.797330	0.817768	2.198	155	0.029
MEANSES2, γ_{03}	-0.951531	1.120816	-0.849	155	0.397
MEANSES3, γ_{04}	3.836587	1.690063	2.270	155	0.025
For FEMALE slope, β_1					
INTRCPT2, γ_{10}	-1.233096	0.280534	-4.396	155	<0.001
SECTOR, γ_{11}	-0.001974	0.428020	-0.005	155	0.996
MEANSESM, γ_{12}	0.799645	0.824353	0.970	155	0.334
MEANSES2, γ_{13}	-0.111261	1.131924	-0.098	155	0.922
MEANSES3, γ_{14}	-1.773548	1.617949	-1.096	155	0.275
For SES slope, β_2					
INTRCPT2, γ_{20}	3.054068	0.168069	18.172	155	<0.001
SECTOR, γ_{21}	-1.448019	0.228292	-6.343	155	<0.001
MEANSESM, γ_{22}	0.436659	0.458282	0.953	155	0.342
MEANSES2, γ_{23}	-1.423832	0.651288	-2.186	155	0.030
MEANSES3, γ_{24}	0.830308	0.944632	0.879	155	0.381

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	12.907247	0.247259	52.201	155	<0.001
SECTOR, γ_{01}	1.283642	0.405022	3.169	155	0.002
MEANSESM, γ_{02}	1.797330	0.764002	2.353	155	0.020

MEANSES2, γ_{03}	-0.951531	1.018663	-0.934	155	0.352
MEANSES3, γ_{04}	3.836587	1.262893	3.038	155	0.003
For FEMALE slope, β_1					
INTRCPT2, γ_{10}	-1.233096	0.263774	-4.675	155	<0.001
SECTOR, γ_{11}	-0.001974	0.415286	-0.005	155	0.996
MEANSESM, γ_{12}	0.799645	0.745535	1.073	155	0.285
MEANSES2, γ_{13}	-0.111261	0.877522	-0.127	155	0.899
MEANSES3, γ_{14}	-1.773548	1.238739	-1.432	155	0.154
For SES slope, β_2					
INTRCPT2, γ_{20}	3.054068	0.142724	21.398	155	<0.001
SECTOR, γ_{21}	-1.448019	0.222484	-6.508	155	<0.001
MEANSESM, γ_{22}	0.436659	0.455257	0.959	155	0.339
MEANSES2, γ_{23}	-1.423832	0.570011	-2.498	155	0.014
MEANSES3, γ_{24}	0.830308	0.782073	1.062	155	0.290

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	*d.f.*	χ^2	*p*-value
INTRCPT1, u_0	1.70192	2.89652	118	294.11109	<0.001
FEMALE slope, u_1	1.03475	1.07071	118	147.36019	0.035
SES slope, u_2	0.18658	0.03481	118	115.27717	>0.500
level-1, r	6.02757	36.33164			

Note: The chi-square statistics reported above are based on only 123 of 160 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Statistics for current covariance components model
Deviance = 46421.274684
Number of estimated parameters = 7

It looks like there is a bunch of non-significant coefficients that we could omit; let's do a hypothesis test:

Results of General Linear Hypothesis Testing - Test 1

	Coefficients	Contrast				
For INTRCPT1, β_0						
INTRCPT2, γ_{00}	12.907247	0.0000	0.0000	0.0000	0.0000	0.0000
SECTOR, γ_{01}	1.283642	0.0000	0.0000	0.0000	0.0000	0.0000
MEANSESM, γ_{02}	1.797330	0.0000	0.0000	0.0000	0.0000	0.0000
MEANSES2, γ_{03}	-0.951531	0.0000	0.0000	0.0000	0.0000	0.0000
MEANSES3, γ_{04}	3.836587	0.0000	0.0000	0.0000	0.0000	0.0000
For FEMALE slope, β_1						
INTRCPT2, γ_{10}	-1.233096	0.0000	0.0000	0.0000	0.0000	0.0000
SECTOR, γ_{11}	-0.001974	1.0000	0.0000	0.0000	0.0000	0.0000

MEANSESM, γ_{12}	0.799645	0.0000	1.0000	0.0000	0.0000	0.0000
MEANSES2, γ_{13}	-0.111261	0.0000	0.0000	1.0000	0.0000	0.0000
MEANSES3, γ_{14}	-1.773548	0.0000	0.0000	0.0000	1.0000	0.0000
For SES slope, β_2						
INTRCPT2, γ_{20}	3.054068	0.0000	0.0000	0.0000	0.0000	0.0000
SECTOR, γ_{21}	-1.448019	0.0000	0.0000	0.0000	0.0000	0.0000
MEANSESM, γ_{22}	0.436659	0.0000	0.0000	0.0000	0.0000	0.0000
MEANSES2, γ_{23}	-1.423832	0.0000	0.0000	0.0000	0.0000	0.0000
MEANSES3, γ_{24}	0.830308	0.0000	0.0000	0.0000	0.0000	1.0000
Estimate		-0.0020	0.7996	-0.1113	-1.7735	0.8303
Standard error of estimate		0.4153	0.7455	0.8775	1.2387	0.7821

χ^2 statistic = 5.187475
Degrees of freedom = 5
p-value = 0.393820

We can safely omit these, as well as fix SES variance. Here's the resulting final model:

Summary of the model specified

Level-1 Model

$$MATHACH_{ij} = \beta_{0j} + \beta_{1j}*(FEMALE_{ij}) + \beta_{2j}*(SES_{ij}) + r_{ij}$$

Level-2 Model

$$\beta_{0j} = \gamma_{00} + \gamma_{01}*(SECTOR_j) + \gamma_{02}*(MEANSESM_j) + \gamma_{03}*(MEANSES2_j) + \gamma_{04}*(MEANSES3_j) + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + u_{1j}$$
$$\beta_{2j} = \gamma_{20} + \gamma_{21}*(SECTOR_j) + \gamma_{22}*(MEANSESM_j) + \gamma_{23}*(MEANSES2_j)$$

Mixed Model

$$MATHACH_{ij} = \gamma_{00} + \gamma_{01}*SECTOR_j + \gamma_{02}*MEANSESM_j + \gamma_{03}*MEANSES2_j$$
$$+ \gamma_{04}*MEANSES3_j$$
$$+ \gamma_{10}*FEMALE_{ij}$$
$$+ \gamma_{20}*SES_{ij} + \gamma_{21}*SECTOR_j*SES_{ij} + \gamma_{22}*MEANSESM_j*SES_{ij} + \gamma_{23}*MEANSES2_j*SES_{ij}$$
$$+ u_{0j} + u_{1j}*FEMALE_{ij} + r_{ij}$$

Final Results - Iteration 41

Iterations stopped due to small change in likelihood function

σ^2 = 36.35048

τ

INTRCPT1,β_0	2.86530	-1.06216
FEMALE,β_1	-1.06216	0.98579

τ (as correlations)

INTRCPT1,β_0	1.000	-0.632
FEMALE,β_1	-0.632	1.000

Random level-1 coefficient	Reliability estimate
INTRCPT1,β_0	0.597
FEMALE,β_1	0.216

Note: The reliability estimates reported above are based on only 123 of 160 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.
The value of the log-likelihood function at iteration 41 = -2.321437E+004

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	12.860968	0.241018	53.361	155	<0.001
SECTOR, γ_{01}	1.267077	0.290825	4.357	155	<0.001
MEANSESM, γ_{02}	2.454522	0.593980	4.132	155	<0.001
MEANSES2, γ_{03}	-0.807528	0.813216	-0.993	155	0.322
MEANSES3, γ_{04}	2.438583	1.275980	1.911	155	0.058
For FEMALE slope, β_1					
INTRCPT2, γ_{10}	-1.210167	0.181840	-6.655	159	<0.001
For SES slope, β_2					
INTRCPT2, γ_{20}	3.083014	0.163651	18.839	6861	<0.001
SECTOR, γ_{21}	-1.442005	0.225767	-6.387	6861	<0.001
MEANSESM, γ_{22}	0.734397	0.304820	2.409	6861	0.016
MEANSES2, γ_{23}	-1.715493	0.572301	-2.998	6861	0.003

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	12.860968	0.220636	58.290	155	<0.001
SECTOR, γ_{01}	1.267077	0.305094	4.153	155	<0.001
MEANSESM, γ_{02}	2.454522	0.549808	4.464	155	<0.001

MEANSES2, γ_{03}	-0.807528	0.761298	-1.061	155	0.290
MEANSES3, γ_{04}	2.438583	0.957991	2.546	155	0.012
For FEMALE slope, β_1					
INTRCPT2, γ_{10}	-1.210167	0.180657	-6.699	159	<0.001
For SES slope, β_2					
INTRCPT2, γ_{20}	3.083014	0.146072	21.106	6861	<0.001
SECTOR, γ_{21}	-1.442005	0.222935	-6.468	6861	<0.001
MEANSESM, γ_{22}	0.734397	0.320445	2.292	6861	0.022
MEANSES2, γ_{23}	-1.715493	0.489462	-3.505	6861	<0.001

Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	*d.f.*	χ^2	*p*-value
INTRCPT1, u_0	1.69272	2.86530	118	292.16127	<0.001
FEMALE slope, u_1	0.99287	0.98579	122	152.94966	0.030
level-1, r	6.02914	36.35048			

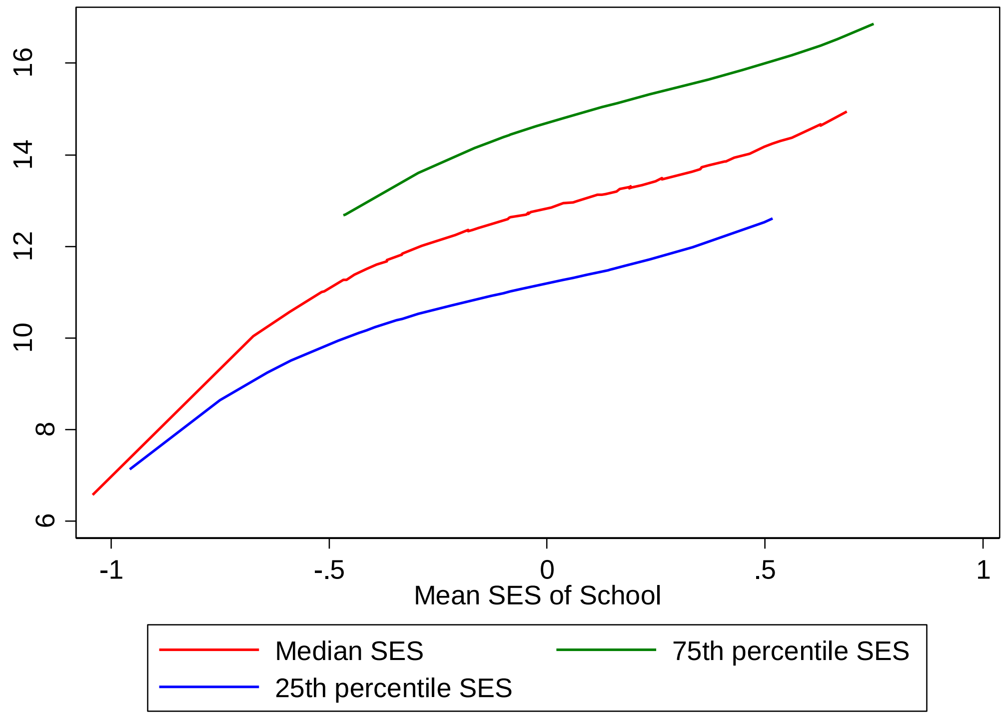
Note: The chi-square statistics reported above are based on only 123 of 160 units that had sufficient data for computation. Fixed effects and variance components are based on all the data.

Statistics for current covariance components model
Deviance = 46428.732167
Number of estimated parameters = 4

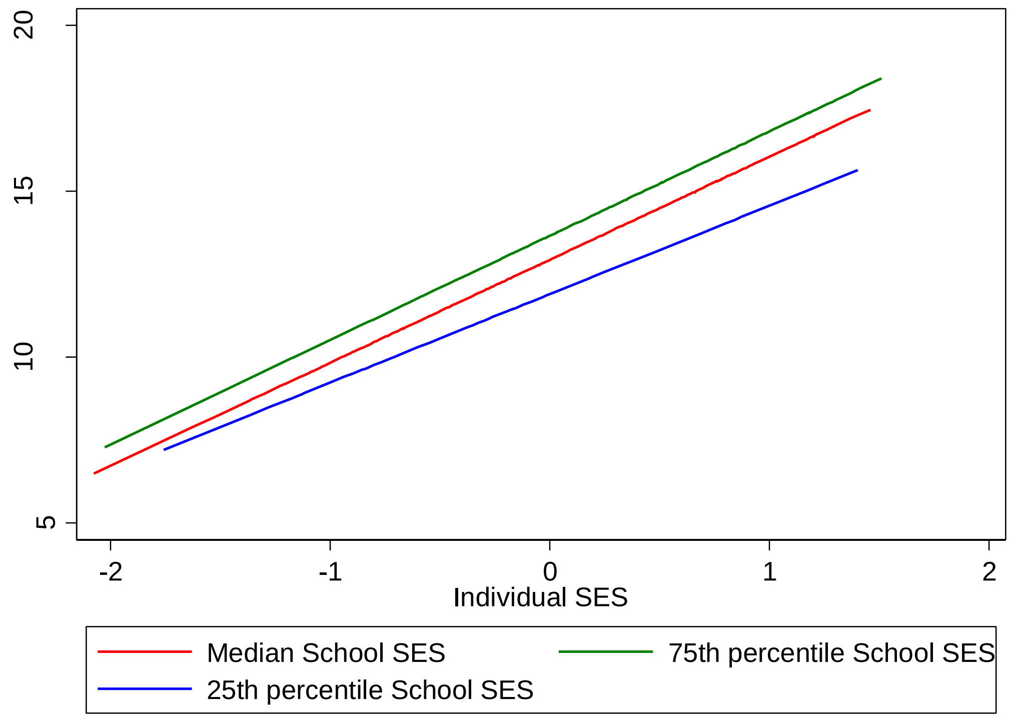
Generating graphs based on that in Stata:

```
. gen pred1=12.860968+1.267077*0 + 2.454522*meansesm -0.807528*meanses2 +
2.438583*meanses3 -1.210167*0 + 3.083014*ses -1.442005*0*ses +
0.734397*meansesm*ses -1.715493*meanses2*ses

. sum ses, det

                              ses
-------------------------------------------------------------
      Percentiles      Smallest
 1%       -1.848         -3.758
 5%       -1.318         -2.838
10%       -1.038         -2.508       Obs                7185
25%        -.538         -2.508       Sum of Wgt.        7185

50%         .002                      Mean           .0001434
                        Largest       Std. Dev.      .7793552
75%         .602          1.732
90%        1.022          1.762       Variance       .6073945
95%        1.212          1.832       Skewness      -.2281447
99%        1.512          2.692       Kurtosis       2.620279
```

```
. graph twoway (line pred1 meanses if ses<.01 & ses>-.01, sort lcolor(red))
(line pred1 meanses if ses<.603 & ses>.601, sort lcolor(green))  (line pred1
meanses if ses<-.537 & ses>-.539, sort lcolor(blue)), legend(label(1 "Median
SES") label(2 "75th percentile SES") label(3 "25th percentile SES"))
ytitle(Predicted math achievement) xtitle(Mean SES of School)
```



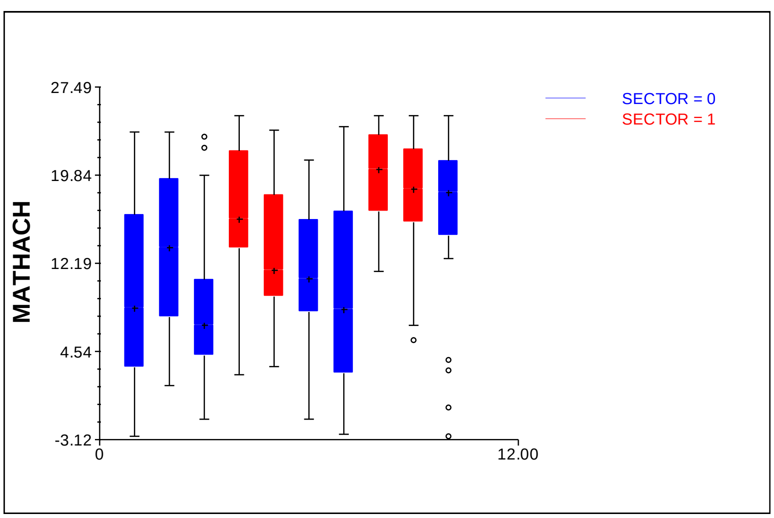

```
. sum meanses, det

                          meanses
-------------------------------------------------------------
      Percentiles      Smallest
 1%       -1.043         -1.188
 5%        -.696         -1.188
10%        -.579         -1.188       Obs                7185
25%        -.317         -1.188       Sum of Wgt.        7185

50%         .038                      Mean           .0061385
                        Largest       Std. Dev.      .4135539
75%         .333           .831
90%         .523           .831       Variance       .1710268
95%         .661           .831       Skewness      -.2681775
99%         .759           .831       Kurtosis       2.520962

. graph twoway (line pred1 ses if meanses<.04 & meanses>.03, sort
lcolor(red))  (line pred1 ses if meanses<.34 & meanses>.32, sort
lcolor(green))  (line pred1 ses if meanses<-.316 & meanses>-.318, sort
lcolor(blue)), legend(label(1 "Median School SES") label(2 "75th percentile
School SES") label(3 "25th percentile School SES")) ytitle(Predicted math
achievement) xtitle(Individual SES)
```



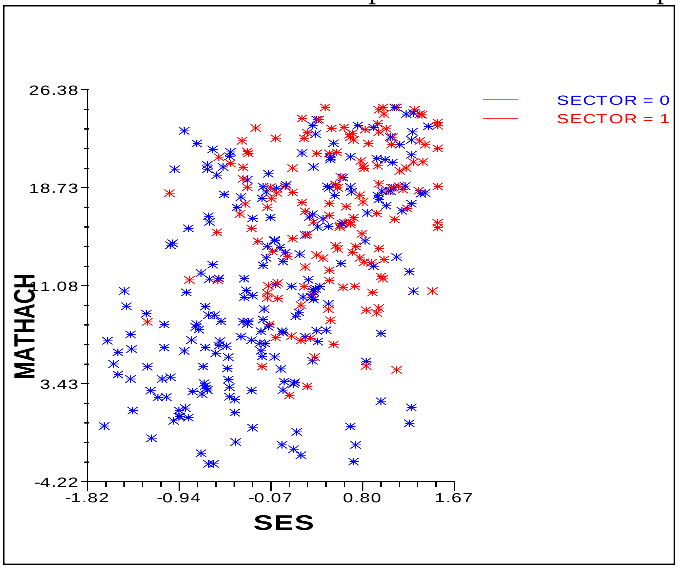
Graphs in HLM

Using graphs to examine the data

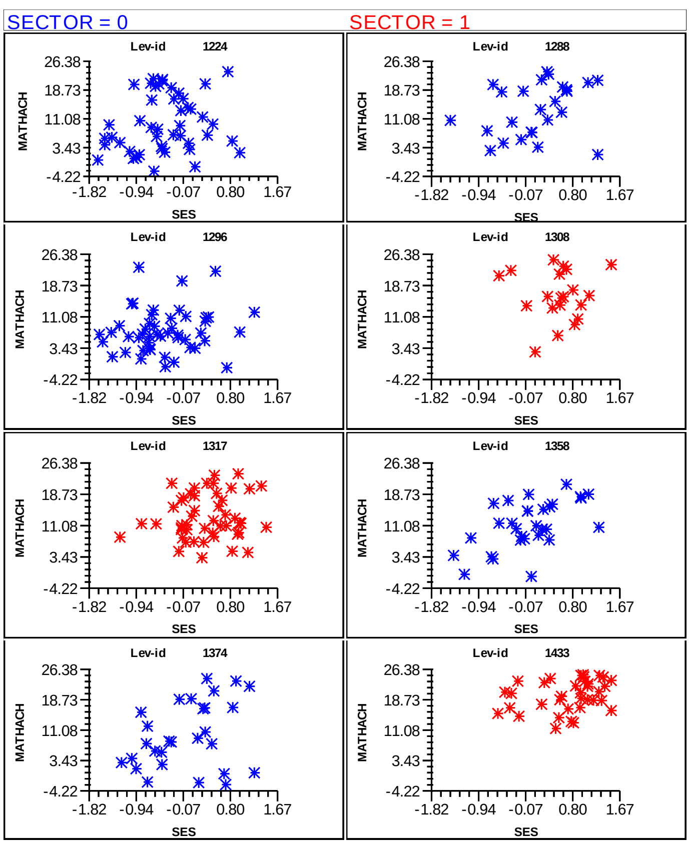
HLM has some limited graphing capabilities allowing you to examine the data before starting to build models. You can examine your data by creating boxplots for a variable, e.g., your dependent variable, by group – to see group differences and group-level outliers. You can also mark these groups according to one level-2 variable:



You can also create a scatterplot for the whole sample by values of a level-2 variable:



Or you can create scatterplots separating groups and colorcoding them by a level-2 variable:



Graphing Equations

Graphs can also be used to better illustrate and understand the models you estimate—but if you have complexity like in the model we just did (with squared and cubed terms), the graphs can produce strange results; Stata is more reliable. But for simpler models, rhese can greatly assist in interpreting the findings. E.g. for a model with SES and SECTOR:

Summary of the model specified

Level-1 Model

$MATHACH_{ij} = \beta_{0j} + \beta_{1j}*(SES_{ij}) + r_{ij}$

Level-2 Model

$\beta_{0j} = \gamma_{00} + \gamma_{01}*(SECTOR_j) + u_{0j}$
$\beta_{1j} = \gamma_{10} + \gamma_{11}*(SECTOR_j) + u_{1j}$

Mixed Model

$MATHACH_{ij} = \gamma_{00} + \gamma_{01}*SECTOR_j$
$+ \gamma_{10}*SES_{ij} + \gamma_{11}*SECTOR_j*SES_{ij}$
$+ u_{0j} + u_{1j}*SES_{ij} + r_{ij}$

Final Results - Iteration 198

Iterations stopped due to small change in likelihood function

$\sigma^2 = 36.76311$

τ

INTRCPT1,β_0	3.83295	0.54112
SES,β_1	0.54112	0.12988

τ (as correlations)

INTRCPT1,β_0	1.000	0.767
SES,β_1	0.767	1.000

Random level-1 coefficient	Reliability estimate
INTRCPT1,β_0	0.759
SES,β_1	0.064

The value of the log-likelihood function at iteration 198 = -2.328373E+004

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	11.750237	0.232241	50.595	158	<0.001
SECTOR, γ_{01}	2.128611	0.346651	6.141	158	<0.001

For SES slope, β_1					
INTRCPT2, γ_{10}	2.958798	0.145460	20.341	158	<0.001
SECTOR, γ_{11}	-1.313096	0.219062	-5.994	158	<0.001

Final estimation of fixed effects
(with robust standard errors)

Fixed Effect	Coefficient	Standard error	*t*-ratio	Approx. *d.f.*	*p*-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	11.750237	0.218675	53.734	158	<0.001
SECTOR, γ_{01}	2.128611	0.355697	5.984	158	<0.001
For SES slope, β_1					
INTRCPT2, γ_{10}	2.958798	0.144092	20.534	158	<0.001
SECTOR, γ_{11}	-1.313096	0.214271	-6.128	158	<0.001

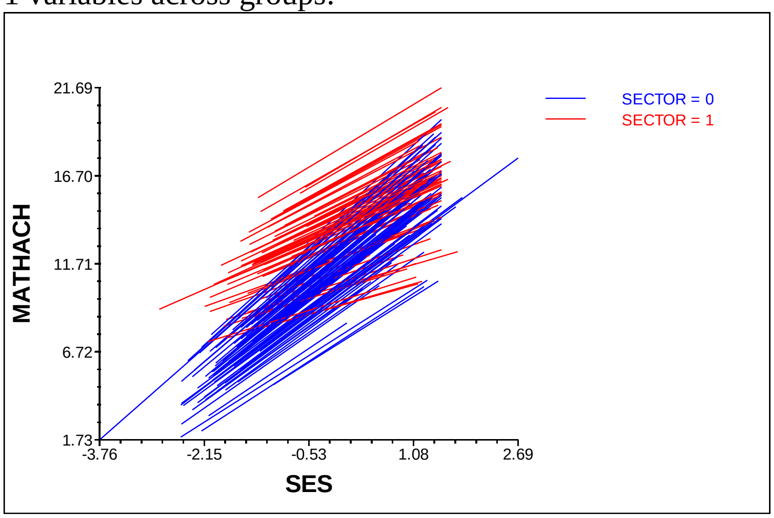
Final estimation of variance components

Random Effect	Standard Deviation	Variance Component	*d.f.*	χ^2	*p*-value
INTRCPT1, u_0	1.95779	3.83295	158	756.04081	<0.001
SES slope, u_1	0.36039	0.12988	158	178.09113	0.131
level-1, r	6.06326	36.76311			

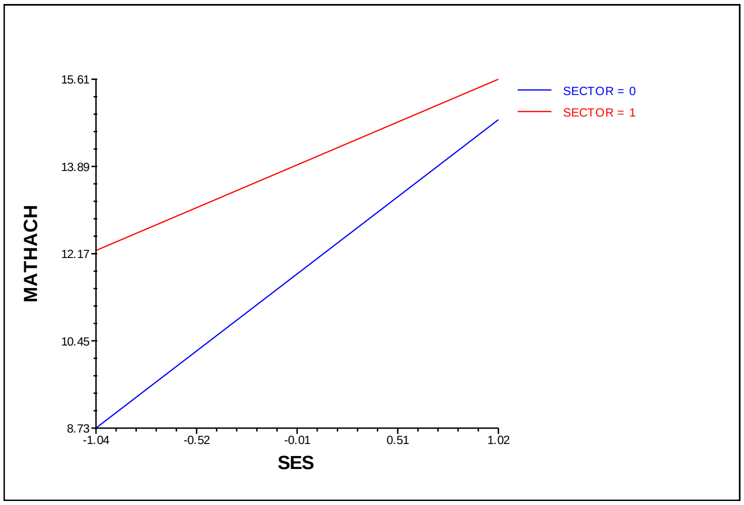
Statistics for current covariance components model

Deviance = 46567.464830
Number of estimated parameters = 4

Let's do Graph Equations → Level 1 equation graphing. Here you can examine slopes for level-1 variables across groups:



Or you can graph the relationships based on the fixed effects in your last model using Graph Equations → Model graphs.   Here, you have a range of options. For example, you can look at how level 1 slopes vary depending on values of level 2 variables (if you have a cross-level interaction in your model). You need to select a level-1 variable as you X, and level-2 variable as Z focus:



Or you can examine how predicted values vary by level of both level-1and level-2 variables by selecting level-2 variable as your X, and level-1 as your Z-focus: