|  |
| --- |
| [[Home]](http://faculty.chass.ncsu.edu/garson/PA765/index.htm)  [[Syllabus]](http://faculty.chass.ncsu.edu/garson/PA765/pa765syl.htm)  [[Statnotes]](http://faculty.chass.ncsu.edu/garson/PA765/statnote.htm)  [[Links]](http://faculty.chass.ncsu.edu/garson/PA765/links.htm)  [[Lab]](http://hcl.chass.ncsu.edu/ssl/ssl.htm)  [[Instructor]](http://faculty.chass.ncsu.edu/garson/PA765/garson.htm)  [[Home]](http://faculty.chass.ncsu.edu/garson/PA765/index.htm) |

**Linear Mixed Models**

**Including hierarchical linear, multilevel, random effects, and repeated measures and growth models**

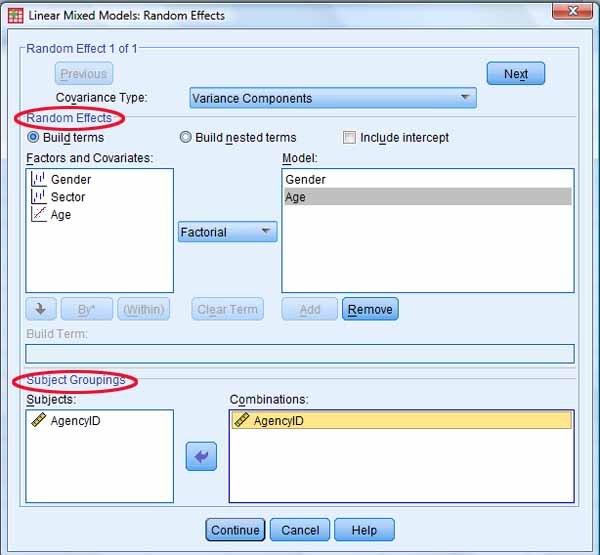
|  |  |
| --- | --- |
| **Overview**  Linear mixed models (LMM) handle data where observations are not independent. That is, LMM correctly models correlated errors, whereas procedures in the general linear model family (GLM, which includes t-tests, analysis of variance, correlation, regression, and factor analysis) usually do not. LMM is a further generalization of GLM to better support analysis of a continuous dependent for:   * + - 1. *Random effects*: where the set of values of a categorical predictor variable are seen not as the complete set but rather as a random sample of all values (ex., the variable "product" has values representing only 5 of a possible 42 brands). Through random effects models, the researcher can make inferences over a wider population in LMM than possible with GLM.          * *GLM vs. LMM for random effects.* It is true that GLM has been adapted to handle random effects models also, but problematically so and therefore LMM is preferred. In estimating model parameters, when there are random effects it is necessary to adjust for the covariance structure of the data. The adjustment made by GLM assumes uncorrelated error for random as well as fixed effects (that is, it assumes data independence) and this is often an unrealistic assumption. LMM handles correlated error and also has the advantage of using maximum likelihood (ML) and restricted maximum likelihood (REML) estimation rather than expected mean squares as does GLM. The result is that GLM produces optimum estimates only for balanced designs (where the groups formed by the factors are equal in size), whereas ML and REML yield asymptotically efficient estimators even for unbalanced designs. Also, ML and REML estimates are normal for large samples (they display asymptotic normality), allowing significant testing of model covariance parameters, something difficult to do in GLM.       2. *Hierarchical effects*: where predictor variables are measured at more than one level (ex., reading achievement scores at the student level and teacher-student ratios at the school level; or sentencing lengths of offenders, gender of judges, and budgets of judicial districts). *Cross-classified effects* are a special case handled by LMM, where higher levels are cross-cutting rather than strictly hierarchical (ex., sudents nested within the cross-classification of schools and neighborhoods).       3. *Repeated measures*: where observations are correlated rather than independent (ex., before-after studies, time series data, matched-pairs designs)   LMM is required whenever the OLS regression assumption of independent error is violated, as it often is whenever data cluster by some grouping variable (ex., scores nested within schools) or by some repeated measure (ex., yearly scores nested by student id).  There are many varieties of LMM, involving diverse labels: random intercept models, random coefficients models, hierarchical linear models, variance components models, covariance components models, and multilevel models, to name a few. While most multilevel modeling is univariate (one dependent variable), multivariate multilevel modeling for two or more dependent variables is available also.  Note that multi-level mixed models are based on a multi-level theory which specifies expected *direct effects* of variables on each other within any one level, and which specifies *cross-level interaction effects* between variables located at different levels. That is, the researcher must postulate mediating mechanisms which cause variables at one level to influence variables at another level (ex., school-level funding may positively affect individual-level student performance by way of recruiting superior teachers, made possible by superior financial incentives). Multilevel modeling tests multi-level theories statistically, simultaneously modeling variables at different levels without necessary recourse to aggregation or disaggregation as in earlier regression approaches. (It should be noted, though, that in practice some variables may represent aggregated scores.)  See the separate sections on [generalized linear mixed modeling](http://faculty.chass.ncsu.edu/garson/PA765/glmm.htm), which is analogous but handles dependent variables which are not normally distributed (ex., binary, count, multinomial, Poisson, gamma, etc.) See also [variance components analysis](http://faculty.chass.ncsu.edu/garson/PA765/variancecomponents.htm) (VARCOMP in SPSS), a subset of LMM capabilities.  See further discussion and additional models in G. David Garson, ed., (2012). *Hierarchical Linear Modeling: Guide and Applications.* Thousand Oaks, CA: Sage Publications. | **Contents**  [Key concepts in all LMM models](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#concepts)  [Estimation settings](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#estimation)  [Covariance structure settings](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#covariance)  [Software differences](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#software)  [The null model and ICC](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#null)  [Two-level random intercept models](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#2ri)  [Two-level random coefficients models](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#2rc)  [Three-level random coefficients models](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#3rc)  [Longitudinal, growth, and repeated measures models](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#longitudinal)  [Multivariate models](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#multivariate)  [Cross-classified models](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#cc)  [Generalized linear mixed models (GLMM)](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#GLMM)  [LMM vs. GLM and VC](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#compareglmandvc)  [Other topics](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#other)  [Assumptions](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#assumptions)  [FAQ](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#faq)  [Bibliography](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#bibliography) |

**Key Concepts and Terms in Linear Mixed Models**

* + **Why model with LMM rather than OLS regression, logistic regression, or GLM?** 
    - *Why not OLS regression?* Compared to OLS regression, linear mixed models correctly computes estimates and standard errors even when observations cluster under higher entities (ex., grouping of employees by agency, students by school, etc.). Clustering of observations within groups leads to correlated error, biased estimates of parameter standard errors (ex., standard errors of regression coefficients), and possible substantive mistakes when interpreting the importance of one or another predictor variable. In most cases, the sampling unit has a random effect. For example, in a study of the federal bureaucracy, "agency" might be the sampling unit and error terms typically would cluster by agency, violating OLS assumptions. Unlike OLS regression, mixed models treat the b coefficients as random effects drawn from a normal distribution of possible b's, whereas OLS regression treats the b parameters as if they were fixed constants. The misestimation of standard errors in OLS regression inflates Type I error (false positives), whereas mixed models handle this potential problem. In addition, LMM can handle a random sampling variable like "agencies" even when there are too many agencies to make into dummy variables in OLS regression and still expect reliable coefficients. The shortcomings of [OLS](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#whylmm) and [GLM](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#glmvslmm) compared to LMM is further discussed in the FAQ section.
    - *Why not logistic regression?* Logistic regression also does not provide for random effects variables, nor does it support near-continuous dependents (ex., test scores) with too large a number of values. Binning such variables, as is sometimes done, loses information and attenuates correlation.
    - *Why not GLM?* GLM does support random effects but it still assumes independent observations. That assumption is often unrealistic because it is common for the sampling unit (ex., cities, schools, agencies) to display intraclass correlation - that is, individual-level observations from the same upper level group will not be independent but rather will be more similar due to such factors as shared group history and group selection processes. LMM does not assume such data independence. This is discussed further in the [assumptions section](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#independence).
    - *Does it really make a difference?*. The application of mixed models such as hierarchical linear models can lead to substantially different conclusions compared to conventional regression analysis. Raudenbush and Bryck (2002: 9-10), citing their 1988 research on the increase over time of math scores among students in grades 1 through 3, wrote that with hierarchical linear modeling, "The results were startling - 83% of the variance in growth rates was between schools. In contrast, only about 14% of the variance in initial status was between schools, which is consistent with results typically encountered in cross-sectional studies of school effects. This analysis identified substantial differences among schools that conventional models would not have detected because such analyses do not allow for the partitioning of learning-rate variance into within- and between-school components."
  + **Fixed versus random effects.**
    - *Fixed effects* are categorical or covariate predictor variables which have a fixed regression slope for each higher level group, though they may well have different intercepts for each grouping level. Fixed effects in relation to the dependent variable are usually the primary variables of interest in a research study. OLS regression models are fixed effects models.
      1. *Cross-level interaction terms as fixed effects*. In a two-level model with both level 1 and level 2 covariates, the interaction of level 1 and level 2 fixed effects are normally modeled as fixed effects. For instance, given student\_SES at level 1 and school\_mean\_ses at level two, entering the interaction term student\_SES\*school\_mean\_SES as a fixed effect is equivalent to modeling the slope of student\_SES as a function of school\_mean\_SES. That is, including such an interaction term in the model is a way of assessing level 2 effects on level 1 variables. This is illustrated in the later section on [full random coefficient models](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#fullrc). Interactions can also be visualized by plotting separate scatterplots of one of the interacting variables with the dependent, getting a separate plot for each value of the other interacting variable.

Interaction effects are effects of "moderator variables." For instance, school budget at level 2 may moderate the effect of socio-economic status (SES) on test performance at the base (student) level. The budget\*SES interaction would test the significance of budget as a moderator variable. The theory, presumably, would be that school budget compensates for some of the educational resources implicit in SES at the individual level. The regression coefficients connecting the moderator variables at level 2 to the regression slopes and intercepts at the individual level are assumed not to vary across groups (schools) and hence these are *fixed coefficients*, in contrast to the *random coefficients* at the base level.

* *Random effects* are categorical or continuous variables whose slopes are seen as a random effect of the level 2 (or higher) grouping variable. In a two-level model, only level 1 variables can be random effects (a level 2 covariate has no higher level for which it can be a random effect). (Note this is not the same as variables based on random sampling of a population: most such variables are fixed factors.)
  + - 1. *When an effect should be seen as random.* An effect should be seen as random if its values (and hence residual error in predictions) cluster by a higher level group. For example, student test scores may cluster by class or school. Random factors affect the covariance structure of the model and if treated as fixed effects, lead to erroneous estimates. Sampling variables (ex., state, where individuals sampled within a sample of states; subject, where a sample of subjects provides repeated measures over time) are random factors as is any grouping variable where the clustering of effects creates correlated error. In a multilevel study, random factors may be grouping factors (ex., individual-level voting data as a random effect of the random factor city, which is the grouping variable, so that voter attributes are level 1 fixed factors and city is a level 2 random factor).
      2. *Random or fixed?* In summary, a given lower level effect may be modeled as a random or as a fixed effect. For instance, SES (socioeconomic status) in a study of employee performance scores grouped by agency might be modeled as a random effect if it is thought its regression coefficient varied randomly by agency, but if the regression coefficient is assumed to be constant across agencies, SES might be modeled as a fixed effect.
      3. *Random intercepts v. random coefficients*. Designating a level 1 variable (ex., employee motivation) as a random effect means the researcher assumes its coefficient varies randomly across level 2 (ex., AgencyID) groups. "Random intercepts" models model only the level 1 intercept of the dependent variable and do not treat level 1 predictors as random effects. "Random coefficients" models model the slope of at least one level 1 predictor as well as the intercept and thus treat that or those predictors as random effects.
      4. *Random factors vs. random effects*. It should be noted that there is ambiguity in the LMM literature over the terms "random factors" and "random effects". In calculation of random coefficients and/or random intercepts, three types of variables are referenced: (1) the level 2 or higher categorical grouping variable, which is the basis of clustering of level 1 individuals or units (ex., employees may cluster on performance score with agency groups); (2) higher level predictor variables also though to influence the intercept and/or slopes at lower levels; and (3) the level 1 or lower level variable whose value is assumed to be a random function of the variability in level 2 or higher regressions by group. The higher level grouping variable or higher level predictors are "random factors" and the lower level variables being influence are the "random effects." As illustrated below, SPSS labels random effects as such in the interface screen below, which is accessed by clicking on the "Random" button on the main LMM dialog screen in SPSS. SPSS calls the random factor associated with grouping a "subject variable".



* + - 1. *Randomness of random factors*. The random factors which define grouping at level 2 or higher are usually any categorical variable whose coding levels are conceived as a sample of possible levels. If the groups defined by this random factor are not random but are skewed in some way, results will also be biased, as would be the case under any statistical procedure. If the groups defined by the random factor are not a random sample but instead are an enumeration of all possible groups, then significance testing is inappropriate. For example, for a study of student test scores grouped by classroom, if all classrooms are included, then any classroom effect on test scores, no matter how small, is a true one: significance only applies when there is a possibility that a different sample of classrooms would yield different results.
    - **Data in LMM models**
      1. *Hierarchical data* involve measurement at multiple levels such as individual and group as, for example, a study of certain variables studied in terms of individual students' opinions, their classes, and their schools. In fact, much early work on multilevel modeling focused on educational settings. In general, hierarchical data are obtained by measurement of units grouped at different levels, such as a study of children nested within families; employees nested within agencies; soldiers nested within platoons, divisions, and armies; or subjects nested within studies.
      2. *Observations*. In univariate multilevel modeling, there is always one dependent variable and it is always at level 1. There may or may not be additional level 1 predictors. There is always a grouping variable which defines level 2. There may or may not be additional level 2 predictors. If there is a second grouping variable, it may be one of two types (1) it can be a level 3 id variable if data are hierarchical (ex., district\_id for level 3, school\_id for level 2, and student\_id for level 1); or (2) it is possible for two or more variables to define the level 2 groupings (ex., race and gender, such that Hispanic males are one group, Hispanic females another group, etc.).
      3. *Multistage sampling*. Hierarchical data are often obtained by multistage sampling. For instance, one might sample schools within school districts, then sample students within sampled schools. In LMM, the higher-level sampling units (ex., schools, districts) are treated as random effects grouping variables.
      4. *Cross-classified data* also involve multilevel data but there can be overlap rather than strict hierarchy at higher levels. This is discussed below in the section on [types of models](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#types).
      5. *Variable entry*. All statistical packages of course require the researcher to specify the id variable(s) defining grouping at each level, the level 1 dependent variable(s), and the predictor variables at each level, if any. These basics are discussed in the FAQ sections for variable entry using [SPSS](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#spssentry) or [HLM](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#hlmentry) software.

*Data format*

* + - 1. *Multiple variable (MV) format*. This is the usual format for between-groups research designs, such as most survey research. There is one dataset with one row per individual observation. There are id variables for level 1 (the individual level, such as employee\_id and any additional level 1 predictors) and for each hierarchical grouping level (ex., for level 2 = agency\_id). If there are additional predictors at level 2, the same row has columns for these as well. This means that all employees from the same agency\_id will have the same values on these additional level 2 predictors. It there is a level 3 (ex., level 3 = state\_id), the same pattern applies. *Warning:* In SPSS MV files, all individuals (or other level 1 units) must have the same value for any level 2 predictors. Failure to do so, due to missing values or other reasons, will cause erroneous estimates, so such values must be adjusted to be the same.
      2. *Multiple datasets*. HLM software, among others, at one time assumed the researcher had created a dataset for each level. It now allows all levels to be entered in a single file. Merging multiple dataset format into MV format is discussed in the [FAQ section](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#singlefile).
      3. *Multiple record (MR) format*. This is the usual format for within-group research designs, such as before-after studies, panel studies, and other repeated measures research. The MR format is similar to the MV format except that there is a variable fot the time period of observation. Any given individual will be represented by as many rows as there are time periods observed. Time becomes level 1 (ex., measures over time nested within students at level 2 nested within schools at level 3).

*Repeated measures data setup*

* + - 1. The data format for repeated measures in LMM is one row per person per measurement time ("time as case format"). Thus if each person is measured at three times, there may be three data rows for person 1, then three rows for person 2, etc. A major advantage of LMM repeated measures over GLM repeated measures is that if a given subject was not measured at a given time, that data row may be simply omitted from the data set. That is, the values of the "time" variable may be different for different subjects. Each data row thus contains an id variable for the person, a time variable (0, 1, 2 in this example), and whatever effects are being modeled. If the researcher's raw data is of the one row per person type, with repeated measures being different columns for each person, then the data must be converted to rows per person per time format. This conversion can be done with the SPSS Data Wizard, under Data, Restructure in the menu. Restructuring is illustrated in SPSS, Inc. (2005: 1-4).

*Standardization*

* + - 1. Raw data are almost always entered in unstandardized form. By default, unstandardized regression coefficients are reported. Entering standardized data will change effect sizes, significance levels, and estimates of variance components in the model. While some software packages (ex., MPlus, LISREL) report standardized coefficients, there are multiple methods of standardization and interpretation of standardized coefficients can be difficult.

*Centering*.

* + - 1. It is customary to center data prior to running LMM or HLM. Centering means subtracting the mean, so means become zero. Two main types of centering are group mean centering and grand mean centering.
      2. *Grand mean centering* is by far the usual type, based on subtracting the grand mean from all raw data values. Grand mean centering often improves the interpretability of coefficients and may reduce multicollinearity. After grand mean centering, "0" has a meaning (ex., 0 income is mean income, whereas before centering, 0 income might be out of the range of actual observations). The phrase "controlling for other variables in the model" after centering becomes equivalent to "holding other variables in the model at their mean." **Grand mean centering is preferred over group mean centering unless there is theoretical justification for the latter**. **Note, however, that binary variables typically are not centered. For the binary variable sex, coded 0=male and 1=female, the phrase "holding sex constant" is equivalent to "holding sex at 0" which is equivalent to "for males."**
      3. *Group mean centering* changes the meaning of coefficients in complex ways which make coefficients hard to interpret, as different mean values are subtracted from different sets of raw scores. As a result, with group mean centering it is not possible to recalculate output back to raw score interpretations. In essence, one is dealing with a different variable after group mean centering. Grand mean centered income, for instance, will yield different slopes but the same deviance and residual errors as uncentered raw data. Group mean centered income does not. Group mean centered income is no longer simple income but rather measures income deviation from group means. The researcher must examine his or her theoretical model and decide if that is really what was wanted for the "income" variable. As noted by Kreft, de Leeuw, and Aiken (1995), the choice of centering must be made on a theoretical rather than statistical basis, and "centering around the group mean amounts to fitting a different model from that obtained by centering around the grand mean or by using raw scores" (p. 1). Most LMM/HLM software packages support various types of automatic centering. Centering considerations are further discussed in Burton (1993) and Hoffman & Gavin (1998).

*Rescaling*

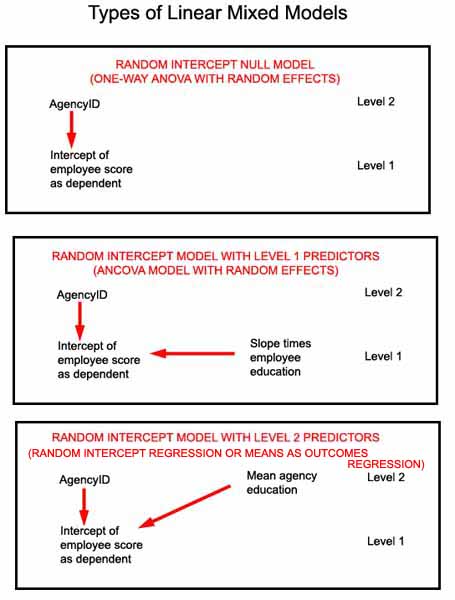
* + - 1. Likewise, some variables may need to be rescaled prior to analysis so that 0 is meaningful and within the range of observations. In longitudingal studies in particular, year or other time variables should be rescaled so that the start year or time is 0. Without rescaling year in this manner, for instance, means that controlling year is controlling year at the traditional year of Christ's birth!

*Missing data*

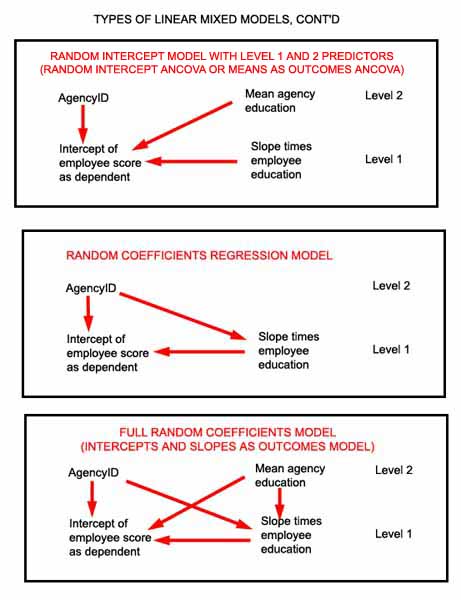
* + - 1. Like most software packages, SPSS defaults to listwise deletion of cases that contain missing values. This biases parameter estimates and is acceptable only if the researcher can demonstrate that missing data are random (see section on [missing data](http://faculty.chass.ncsu.edu/garson/PA765/missing.htm) handling), which is rarely the case. By a common rule of thumb, if missing cases are fewer than 5% of total cases, the researcher may proceed. Otherwise missing data should be imputed using maximum likelihood (ML) or full information maximum likelihood (FIML) estimation methods. In SPSS, this is implemented in the "Missing Data" module.
      2. HLM software handles missing data differently according to which of its several modules (see [below](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#hlmmodules)) is being used. For ordinary two- and three-level linear mixed models (HLM2 and HLM3 modules), observations are deleted in listwise fashion. The researcher may opt for listwise deletion when making HLM's .mdm format data fle, in which case deletion is based on all level 1 variables. Alternatively, listwise deletion will be performed when an analysis is run, in which case it is based only on variables in the model (Other Settings, Estimation Settings menu using the Level-1 Deletion Variables option). For cross-classified models (HCM2, HCM3, and HLMHCM), listwise deletion is performed when the .mdm file is created. Multivariate models (HMLM and HMLM2) assume that the data has no missing values. For all modules, HLM assumes all higher level data have no missing values. If a missing value is encountered in a higher level group, observations for the entire group are deleted. If ASCII rather than statistical package (ex., SPSS .sav files) input is employed, blanks or missing data codes would be treated as valid data and therefore must be removed prior to input.
    - **Types of LMM models**
      1. *LMM nomenclature*. The diverse labels for subtypes of linear mixed models may lead to confusion about what are, in fact, members of the same family of statistical procedures. In some disciplines LMM models are called "random effects" or "mixed effects" models. In economics, the term "random coefficient regression models" is common. In education, the label "hierarchical linear modeling" is commonly used. In sociology, "multilevel modeling" is common. And in statistics, the term "covariance components models" is often used. All these terms are closely related, albeit emphasizing different aspects of linear mixed models.
      2. *Fixed effects models* are models with only fixed factors and optional covariates as predictors. Most models in analysis of variance, regression, and GLM are fixed effects models, which are by far the most common type in social science. An example would be a study of job satisfaction by gender, controlling for salary level. Job satisfaction would be the dependent variable, gender the fixed factor, and salary the covariate, treated as fixed. Both GLM and LMM can model fixed effects models with very similar if not identical estimates and similar but not identical output tables. However, where LMM is clearly superior to GLM lies in handling other types of models discussed below.
      3. *Random effects models* are models with one or more random factors and optional covariates as predictors. If there are covariates, they are treated as fixed effect variables. An example would be a study of job satisfaction at level 1 by city at level 2, controlling for salary level at level 1. Job satisfaction would be the dependent variable, city the random factor (assuming only a random sample of cities was studied), and salary the covariate. The level 1 intercept of job satisfaction may be modeled as a random effect of city at level 2. Likewise, the level 1 slope of salary might be modeled as a random effect of city. If only the intercept is modeled, it is a *random intercept model*. If slopes are modeled also, it is a *random coefficients model*.
      4. *Mixed models* in their full form have both fixed and random factors as well as optional covariates as predictors. Predictors could be at any level. An example would be a study of job satisfaction by gender by city, controlling for salary level. Job satisfaction would be the dependent variable at level 1, gender the fixed factor at level 1, city the random factor t level 2, and salary the covariate at level 1.
      5. *Hierarchical linear models (HLM)* are a type of mixed model with hierarchical data - that is, where data exist at more than one level (ex., student-level data and school-level data). In explaining a dependent variable, HLM models focus on differences between groups (ex., schools) in relation to differences within groups (ex., among students within schools). While it is possible to construct one-level models in LMM, most use of LMM can be seen as one or another form of HLM.
      6. *Homogenous vs. heterogenous variance models*. One of the assumptions of linear mixed modeling is that the variance of the level 1 random error term should have equal variance across groups. Most LMM models are of this type. However, it is possible to create heterogenous variance models. This is discussed in the ["Assumptions" section](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#homogenous).
      7. ***Repeated measures models***. In a repeated measures study, the measures at multiple times are level 1 and the unit of observation (usually individuals) is level 2. When "repeated measures" is selected in statistical software, LMM models the within-subjects (a.k.a., within-groups or residual) variance. Within-subjects variance is the variance among repeated measures at multiple times for given individuals, on the average) and level 2 is between-subjects, with the individual subjects being a grouping variable for the tests/observations. One regression is created for each subject, generating a large number of intercepts, where the true intercept is estimated as a random function of the intercepts of all the regressions. Random slopes may be generated in the same way if there are level 1 predictor variables for the level 1 dependent being studied. In repeated measures studies, the researcher must specify the covariance structure of the R matrix, which corresponds to the residual variance component, as discussed in the section on covariance structure types.
      8. *Growth models* are a common type of repeated measures model in which individuals (or other units of analysis) are measured at three or more time periods, with one data row per time period. The individual id variable is the subject grouping variable and time is modeled as a fixed and/or random effect on some measurement about the individual (ex., a test score). For example, employees might be measured at three times, obtaining three performance scores. There might also be factors or covariates (ex., a seniority status variable or a baseline performance score). The time variable baseline is coded 0 and subsequent measurement times are 1, 2, 3, etc. If the time codes represent equal metric intervals, the time variable is a covariate (rather than a factor) and linear growth (or quadratic growth, or other functions of time) may be analyzed.
         * *Linear vs. quadratic growth models*. If the time codes 0, 1, 2 are not arbitrary but reflect a real scale, such as 1=elapse of one month, 2=elapse of two months, etc., then time can be modeled as a covariate predictor. It is possible for one of the covariates to be time-squared, assuming the time variable for measurement represents equal intervals. Both time and time-squared would be modeled when investigating quadratic growth. To do this, time must be a covariate, not a factor.
    - *Types of models*
      1. *Random intercept models* are models where only the intercept of the level 1 dependent is modeled as an effect of the level 2 grouping variable and possibly other level 1 covariates and possibly other level 2 random effect predictors.
         * *The null model (one-way ANOVA with random effects)*, sometimes called the unconditional model, predicts the level 1 intercept of the dependent as a random effect of the level 2 grouping variable, with no other predictors at level 1 or 2. The same logic is applied if there are three or more levels. For instance, differences in mean performance scores (the intercepts) may be analyzed in terms of the between-groups effect of agency at level 2.

Null or unconditional models are often used as a comparison baseline. A model is "conditional" by the presence of a level 1 predictor and/or a level 2 predictor. Since the researcher almost always employs predictor variables and is not simply interested in the null model, most mixed models are conditional. In fact, the central point of much LMM modeling is to assess the difference between the researcher's conditional model and the unconditional model without predictors. The likelihood ratio test (model chi-square difference test) can be used to assess the difference in fit.

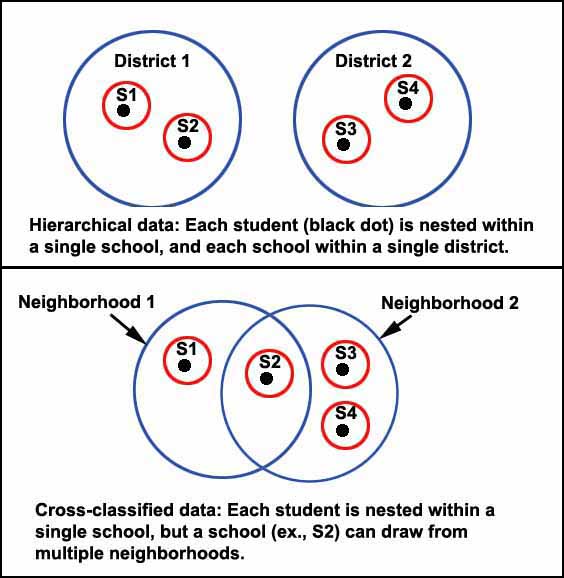
* + - * + *One-way ANCOVA with random effects models*. It is also possible to have a level 1 covariate but still predict the level 1 intercept (and not the slope of the level 1 covariate) as a random effect of the level 2 grouping variable with no other level 2 predictors. For instance, differences in mean performance scores (the intercepts) may be analyzed as predicted by salary at level 1, predicting only the level 1 intercept in terms of the between-groups effect of agency.
        + *Random intercept regression models* are also called "means-as-outcomes regression models". This variant of the random intercept model is to predict the level 1 intercept on the basis of the level 2 grouping variable and also on the basis of one or more level 2 random effect predictors. For instance, differences in mean performance scores (the intercepts) may be analyzed, predicting the level 1 intercept in terms of the between-groups effect of agency and the level 2 random effect variable EquipmentBrand (a factor representing a sample of some of many brands of equipment, where different agencies used different brands).
        + *Random intercept ANCOVA models* are also called "means-as-outcomes ANCOVA models." This is simply a random intercept regression model in which there is also a level 1 covariate treated as a fixed effect (slope not predicted by level 2). Some authors would label this another type of "random intercept regression model."



* + - 1. *Random coefficients models (RC)*, also called multi-level regression models, are a type of mixed model with hierarchical data. The level 1 dependent is predicted by at least one level 1 covariate. The slope of this covariate and usually also the intercept are predicted by the random effect of the grouping variable at level 2. That is, each group at the higher level (ex., school level) is assumed to have different regression slopes as well as different intercepts for purposes of predicting a level 1 dependent variable. While this could be visualized by using OLS regression and superimposing the n regression lines for the n schools, LMM incorporates this variability of regression lines into a single analysis.

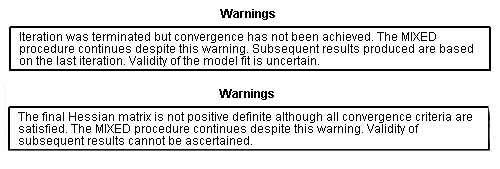


* + - * + *Random coefficients regression models* are a type of RC model in which the level 1 model is a typical regression model in which a dependent is predicted by one or more level 1 covariates. The level 1 slope(s) as well as the level 1 intercept is predicted on the basis of the level 2 grouping variable as a random effect, but there are no other level 2 predictors. For instance, differences in mean performance scores (the intercepts at level 1) may be analyzed as predicted by salary at level 1, predicting both the level 1 intercept and the slope of salary in terms of the between-groups effect of agency at level 2. Random coefficients regression models are unconditional at level 2.
        + *Intercepts-and-slopes-as-outcomes models* are a type of RC model in which the level 1 slopes and intercepts are modeled not only by the level 2 grouping variable as a random factor but also by one or more other level 2 random effect variables. For instance, differences in mean performance scores at level 1 may be analyzed, predicting the level 1 intercept and the slope of the level 1 predictor salary in terms of the between-groups effect of agency and the level 2 random effect variable EquipmentBrand. Intercepts-and-slopes-as-outcomes models are conditional at level 2.
        + *Nonrandomly varying slopes models*. It is possible that a level 2 random effect predictor so fully explains the slopes of the level 2 predictor that the random effect of the level 2 grouping variable is not significantly different from zero. For instance, differences in mean performance scores at level 1 may be analyzed, predicting the level 1 intercept and the slope of the level 1 predictor salary in terms of the level 2 random effect variable EquipmentBrand (and not also the between-groups effect of agency). This type of model may also be considered conditional at level 2.
      1. *Types of HLM models*. HLM 7 software operates through several modules, each designed for a type of HLM model.
         * HLM2. For two-level linear and non-linear models with one dependent (outcome) variable.
         * HLM3 and HLM4. For three-level and four-level linear and non-linear models with one dependent (outcome) variable..
         * HGLM. For generalized linear models for distributions other than normal and link functions other than identity, handling binary, count, multinomial, and ordinal outcome variables with Bernoulli, binomial, Poisson, multinomial, and ordinal models.
         * HMLM. For multivariate normal models with more than one outcome variable, including when the level 1 covariance structure is homogenous, heterogenous, loglinear, or AR(1) (first order autoregressive). Handles incomplete data. Level 2 variables are treated as random effects.
         * HMLM2. For two-level HMLM models where level 1 is nested within level 2.
         * HCM2. For models where level 1 units are *cross-classified* by two level 2 units. In standard HLM the levels are in one-to-one relationships. For instance, each student is nested within one school; each school is nested within a single district. Cross-classified models handle many-to-many relationships such as the student-schools-neighborhoods example illustrated below, where a school can draw from many neighborhoods and a neighborhood can contain many schools.
         * HCM3. For three-level cross-classified models.
         * HLMHCM. For two- and three-level hierarchical linear mode3ls with cross-classified random effects (ex., repeated test scores nested within students who are cross-classified by schools and neighborhoods).



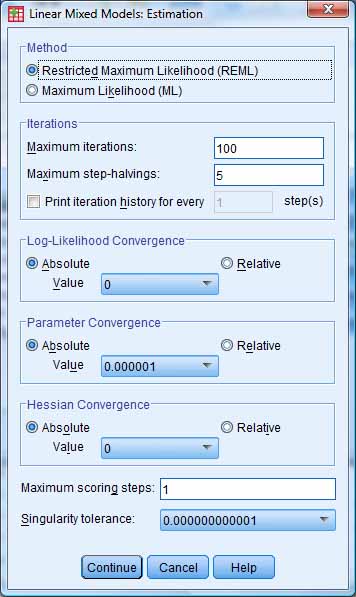
**Estimation Settings in Linear Mixed Models**

* + **Estimation algorithms**. The default algorithm used to compute coefficients for the predictor variables in SPSS is restricted maximum likelihood (REML), but by clicking the Estimation button in SPSS the researcher may choose maximum likelihood (ML) instead. The Estimation button also allows the researcher to set the number of iterations and various convergence options though normally the defaults suffice. LMM estimation, whether by REML or ML, contrasts with the ANOVA methods utilized in GLM.
    - ***ML*.** Maximum likelihood estimation represents an iterative algorithm which finds the parameter estimates (regression coefficients) which maximize the probability of arriving at estimates of the dependent variable equal to the observed values. ML estimates ignore the degrees of freedom used up by fixed effects in mixed models, leading to underestimation of variance components. However, ML may nonetheless be preferred when comparing two models with different parameterizations of the same effect (ex., simple variable vs. quadratically transformed version of the variable), because ML is invariant to different parameterizations of a fixed effect but REML will treat different parameterizations as different models and compute different likelihood ratios. **ML should be used if the research purpose is to compare fixed effects (regression coefficients) for a set of nested models**.
    - ***REML*.** Restricted maximum likelihood estimation has better bias characteristics (Diggle, 1988), handles high correlations more effectively, and is less sensitive to outliers than ML, **but cannot be used for model comparison of fixed effects, as noted** [**below**](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#2lltests) **in the section on likelihood ratio tests**. REML is "restricted" in the sense that the likelihood function includes only variance components, not regression coefficients as does ML. REML estimates will be the same as ML estimates for large samples, but when the number of level 2 groups is small, REML is less biased than ML. For small samples, ML is biased toward lower parameter estimates. REML should be used if the research purpose is to compare random effects (variance components) for a set of nested models.
  + **Convergence** . Occasionally the maximum likelihood estimation algorithm, which is iterative, fails to converge on a solution. Alternatively, convergence may be achieved but the algorithm cannot invert the Hessian, which is a matrix used to compute standard errors for the covariance parameters. If variances equal to zero or negative variances appear on the diagonal of the Hessian it cannot be inverted to compute the needed standard errors. One of the following warnings may appear in SPSS:



If either of these warnings appear, there are three recourses for the researcher:

1. *Stop and throw out the model.* This strategy is the technically correct one, since without convergence, parameter estimates may be suboptimal, and without an invertible Hessian, the computed standard errors may be erroneous.
2. *Proceed on an exploratory basis*. **Convergence and non-positive definite Hessian warnings are similar to other violations of the assumptions of a procedure.** The calculated results cannot be interpreted to mean exactly what they are supposed to mean. The researcher does not know how different the printed estimates and standard errors are from the ones which should have been output. However, in most cases the direction and general magnitude (weak, moderate, strong) of effects will not differ and in many cases differences will prove to be minor, justifying proceeding with analysis on an exploratory basis.
3. *Re-specify the model and/or adjust estimation parameters*. This recourse should be taken before considering either of the first two. Several remedial actions may be possible, some done through adjusting settings in the dialog for the "Estimation" button in the SPSS linear mixed models module. Changes from default estimation settings should be reported.



One or more of the following steps may eliminate the foregoing warnings:

MODEL SPECIFICATION AND DESIGN CHANGES

1. Estimation problems may reflect multicollinearity. Check for and remove any redundant variables. Running a regression to check for variables with high variance inflation factor (VIF) coefficients is one method of doing this.
2. Specifying a simpler covariance type (see below) in the "Random" dialog may avoid warnings. For instance, the "Unstructured" type is the most complex, having the most parameters, and one may consider instead specifying a "Diagonal" or other type. In HLM software, use the "Diagonalize tau" option to replace the default unstructured type with the diagonal type.
3. If the estimate for a slope/intercept covariance for a random effect approaches zero, make the slope variable a fixed effect only. In general, lack of convergence can result from trying to estimate random coefficients which are close to or equal to zero.
4. Consider a simpler model with fewer variables.
5. Increase the sample size if small. There cannot be fewer observations than parameters

ESTIMATION PARAMETER CHANGES

1. Estimation settings must be changed with caution. See the fuller discussion of the change options below in Garson, ed. (2012: ch. 2).
2. Under the Estimation button, increase the maximum iterations above the default (100). Specify a positive integer.
3. Under the Estimation button, increase step-halvings above the default (5). Specify a positive integer.
4. Under the Estimation button, increase the "Parameter Convergence" value to be larger than the default (.000001). Specify a positive value.
5. By default, the values for "Log-likelihood Convergence" and for "Hessian Convergence" are set to 0, which means these convergence criteria are not used. However, setting either to a non-negative value chosen from the drop-down list of alternatives will invoke them. By default the setting value is used as an "absolute" convergence criterion, but one can check the "relative" radio button in the "Estimation" dialog to cause convergence to be assumed if the statistic is less than the product of the value specified and the absolute value of the log-likelihood.
6. Under the Estimation button, increase the singularity tolerance value (a drop-down list of choices is provided).
7. Under the Estimation button, increase the number of scoring steps above the default (1).

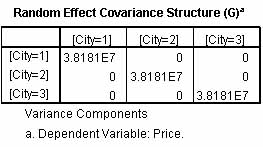
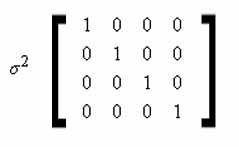
**Covariance Structure Settings in Linear Mixed Models**

* + **Covariance Structure Type**
    - *What covariance structure type is*. This refers to the assumed structure of the variance-covariance matrix, in which the rows and columns are the modeled random effects, Rows and columns are the levels of the grouping variable (ex., city=1, city=2, city=3). The diagonal cells are variances and the off-diagonal cells are covariances. There is an assumed G covariance structure for random effects and separately an R matrix for repeated measures. Depending on the research design, the researcher may need to specity the assumed covariance structure type for the random effects G matrix and/or for the repeated measures R matrix.
      1. *Mathematical notation.* Following notation of Verbeke and Molenberghs, (2000), a linear mixed model has the form:  
         Yi = Xiβ + Zibi + εi

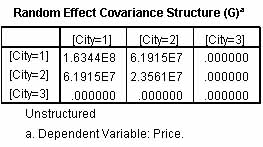
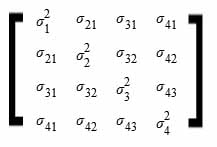
Let....  
i = the number of subjects  
pi = the number of observations per subject  
q = the number of fixed effects  
r = the number of random effects  
Then...  
Yi is a pi-by-1 vector of response values (predictions) for the i subjects  
Xi is a pi-by-q matrix of values of covariates which are fixed effects  
β is a q-by-1 vector of fixed effect (regression) coefficients  
Zi is a pi-by-r matrix of values of covariates which are random effects  
bi is a r-by-1 vector of random effect coefficients  
εi is a pi vector of error terms (residuals) for each of i subjects with p observations each

* + - 1. *The G matrix* is the variance-covariance matrix for bi, representing estimated random effects. There will be one G matrix for each random effect variable (not including the intercept). G is the matrix whose type is set when the researcher specifies the covariance type for random effects. If desired, it is possible to specify a different covariance structure for each random effect. G is a block diagonal matrix. Rows and columns represent the groups defined by categories of the random effect factors (for ex., for "vote", 0 = did not vote, 1 = voted, the G matrix would be a two-by-two matrix with columns vote=1 and vote=2 within groups formed by the subject (grouping) variable (ex., city). SPSS, for instance, will list the first row and column as "[vote=0] | city". Covariate random effects have a one-by-one (single cell) G matrix. In any G matrix for a factor which is a random effect, the upper-left cell contains the variance component for the first group as all the diagonal entries are also variance components for their respective factor levels. The next cell down in the first column contains the estimated covariance of the given factor's random effect in the first group with that in the second group, and so on. For a covariate random effect, the single cell is the variance component for that effect (ex., for income).
      2. *The R matrix* is the variance-covariance matrix for εi, representing estimated residual error, which is within-subjects variation when level 1 is repeated measures and level 2 is subjects. It is this matrix whose type is set when the researcher specifies the covariance type for repeated measures. R is a block matrix, with rows and columns representing the levels of the repeated variable (typically, time). Variances are on the diagonal and covariances in the off-diagonal cells. Thus the upper-left cell in the matrix contains the estimated error variance for the first time period. The next cell down in the first column contains the estimated covariance of error in the first time period with error in the second time period, and so on.
    - ***What specifying a covariance structure type does.* When the researcher specifies a covariance structure type, two things are done: (1) the solution is constrained to have a structure of the specified type; and (2) the algorithm reaches better estimates because the selected type corresponds (hopefully) to the actual structure of the data. Linear mixed modeling uses an iterative algorithm to estimate coefficients. The estimate will be more reliable if the algorithm uses as a starting point the proper covariance type assumption. Assuming a too-simple covariance structure will increase Type I errors, while assuming a too-complex structure will increase Type II errors when interpreting variance components in the "Estimates of Covariance Parameters" output table discussed below**.
    - *Two types of covariance structure assumptions*. If there are random effects or repeated measures, the researcher specifies the covariance structure type(s). A given model may require either or both of two types of covariance structure assumptions.

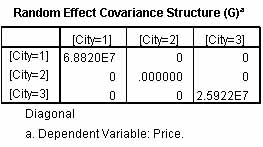
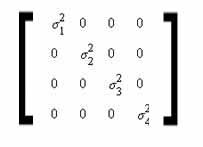
1. *Repeated covariance type*. This type is specified on the opening SPSS LMM screen, if the researcher has entered a repeated measures variable. The default type for repeated measures research is "Diagonal," also called "Simple." This specifies what is called the repeated measures type "R" variance/covariance matrix. It is this matrix which is the basis for creating the within-subjects correlation matrix. Note that if a random effects covariance type is declared because the researcher is undertaking a repeated measures study, the researcher will still declare random effects covariance type for random effects variables, if any. The covariance structure specified for random effects need not be the same as specified for repeated measures.
2. *Random effects covariance type.* This type is specified on the SPSS LMM dialog screen associated with the "Random" button, where random effect variables are declared. If there are one or more random effects, the researcher must specify the type of covariance structure to assume. For random effects, the default is "variance components" structure, discussed below. This specifies what is called the random effects type "G" variance/covariance matrix. This matrix is the basis for estimating between-groups effects.
   * + - * *Multiple random effects*. Ordinarily there will be just one random effects model and just one covariance type specified for it. However, under the Random button in SPSS, it is possible to have multiple random effects models, each with its own specified covariance structure, which may be the same or different from the others. If there are multiple random effects models, they are assumed to be independent of one another (terms within the model for any random effect may be correlated, however). If the same random effects variable is listed in more than one random effects model, it must refer to a different combination of Subject variables. When there are multiple random effects models, a separate covariance structure is estimated for each.
   * **Examples of covariance structure types**.
     + *Types*. For the following types of covariance structure, the figures reflect the example, discussed [below](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#basicmixedmodels) in the section on basic mixed models, of city as a random effect in a city tax study of the relation of appraisal value to selling price. Figures are the "G matrix," reflecting the City random effect covariance structure.
       1. ***Variance Components***, also called "simple structure" or "the independence model" (because residual variance is independent of effect variances). This is the default for random intercept models in both SPSS and SAS linear mixed models. It is not available for repeated measures. **Each group (ex., city) has the same variance of residuals as each other group [contrast diagonal structure (below), where this may be unique by group], and the covariance of residuals (error terms) is assumed to be 0 (ex., residuals are not correlated between any pair of cities).** **If the assumptions of a variance components model can be met, it means that effects are additive: the ratio of effect covariance estimates to residual covariance estimates is the ratio of the importance of between-subject effects to within-subject effects in accounting for the variance of the dependent variable.**



* + - 1. ***Unstructured*.** This is a completely general covariance matrix in which each variance and covariance estimate is computed separately. **Unstructured is a common assumption for random coefficients models (regression models where slopes as well as intercepts are modeled by LMM). In unstructured models, the intercept, slope, and intercept\*slope covariance are all modeled. This choice is appropriate when the researcher has no basis for knowing what the covariance structure is. Unstructured models are the assumption in GLM MANOVA. Slope variances and covariances are estimated from the data in an unconstrained manner which allows them to correlate (in a variance components model, in contrast, slopes of random effects, including intercepts, are assumed uncorrelated).** This type is complex, requiring the computation of many parameters - that is, it is unparsimonious and so BIC or other goodness of fit measures taking parsimony into account will penalize unstructured models. One strategy is to choose the unstructured covariance type initially, but to request output of a matrix of covariances of residuals to discern possible patterns which would warrant re-running the model with a simpler covariance type assumption. Sometimes dropping outliers may make such a pattern more discernable. When slope coefficients or slope-intercept covariances are very small, an unstructured solution may not be possible and an error message will be generated. In such situations, it is common to try a diagonal covariance structure, discussed below.



* + - 1. *Diagonal (Simple)*. This is the default for repeated measures covariance structure type but it may be used in other contexts (and other covariance structure types may be used for repeated measures). Diagonal covariance structure has heterogenous variances (unlike variance components structure discussed below) and zero correlation between groups. For instance, if the repeated measures groups are test scores in three time periods, the diagonal structure assumption is that there is no correlation of the residuals of the Time 1 test scores with the residuals at Time 2 when predicting the dependent variable. **Each element (ex., time period) may have a unique variance for its residuals, but the covariance between time periods is zero. This implies that the measures for an individual in one time period are independent of (uncorrelated with; hence the diagonal model may be called the independence model) measures for that individual in any other time period. This may well be an unrealistic assumption. Alternative common structures that might be assumed for repeated measures are AD(1) and AR(1) or simply Unstructured, discussed below.** While diagonal structure is the default for repeated measures, it may also be specified for random effects, as illustrated for cities as a random effect in the figure below.



* + - 1. *Others*. Many other types of covariance structure are possible, discussed [below](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#cvs) in the FAQ section.
  + **Selecting the best covariance structure assumption**.
    - *Goodness of fit statistics* (-2LL, AIC, AICC, BIC - discussed in the section on [structural equation modeling](http://faculty.chass.ncsu.edu/garson/PA765/structur.htm#infotheory), for instance) can be used to select the best covariance structure type to assume. Simulation studies by Ehlers (2004) suggest BIC is preferred generally (compared to AIC, BIC penalizes for lack of parsimony), but AICC is better for sample sizes < 20 and when there is only one group. Usually, however, this is a moot question because the criteria will agree in most cases. The researcher runs the model under the desired covariance structure assumptions (ex., under variance components, unstructured, and AR(1), or others listed below), then looks in the "Information Criteria" table in the output to take note of the BIC, AICC, or other fit statistic for each run of the model. The model with the lowest value on BIC (or other chosen fit criterion) is the one with the best-fitting covariance structure assumption. Note that if one or more fixed effects is dropped from the model, BIC (or other fit criteria) must be re-estimated and re-evaluated.
      1. *Likelihood ratio tests*. It is also possible to run a likelihood ratio test on the difference between two models, one under a given covariance structure assumption and another model under a different assumption. Likelihood ratio tests are described in the [section on structural equation modeling](http://faculty.chass.ncsu.edu/garson/PA765/structur.htm#diff).
      2. *Data-driven vs. theory-driven selection*. Note, however, that both the information criteria or the LR test methods are data-driven and thus prone to over-fitting, particularly is cross-validation is not part of the design. Covariance structure assumptions should start with theoretical selection based on covariance expectations described [below](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#cvs).
    - *Parsimony*. Between alternative equally well-fitting models with differing covariance structure assumptions, the simpler, more constrained structure is preferred on parsimony grounds. **The unstructured covariance type is the least constrained and least parsimonious. The variance components type is the simplest and most constrained, with variances constrained to be equal and covariances constrained to be 0.**

**Software Differences in Linear Mixed Models**

* + **Software packages for LMM**.
    - *Packages*. Although popularized by the early introduction of HLM software authored by Raudenbush & Bryk, quite a number of software packages now exist for implementing LMM models, including HLM 7, SAS, SPSS, Stata, MPlus, and R.
    - *HLM 7*. HLM software is one of the leading statistical packages for hierarchical linear modeling. Though differences between software packages' capabilities have diminished over time, HLM 7 offers a number of appealing features and capabilities. Among these are what many consider to be the more intuitive model specification environment, the ease in creating three- and four-level models, the wide choice of estimation options, integrated likelihood ratio hypothesis testing, integrated generalized linear modeling, graphics options, and the ability to handle heterogenous hierarchical linear models (where the dependent is thought to have different variances for different levels of some grouping variable such as gender or race, for instance). HLM 7 is distributed by . [Scientific Software International](http://www.ssicentral.com) (SSI). A free student edition of HLM 7 is available at <http://www.ssicentral.com/hlm/downloads/HLM7StudentSetup.exe>. The student edition is full-featured, including examples, but is limited in the size and complexity of models (though it will work all example files).
    - *SPSS* implements linear mixed models in its MIXED module (menu choice: Analyze, Mixed models, Linear). It implements generalized linear mixed models in its GENLINMIXED module (menu choice: Analyze, Mixed models, Generalized linear).
    - *SAS* implements linear mixed models in its PROC MIXED module. It implements generalized linear mixed models in PROC GLIMMIX.
  + **Why different packages may yield different results**.

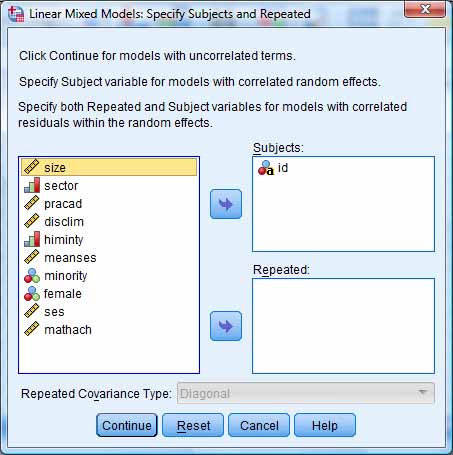
Although alternative packages usually yield identical coefficients and almost always lead to the same substantive conclusions, they may differ, even significantly. There will be differences, usually only after the first few decimal places, for reasons listed below. However, sometimes the differences are appreciable, affecting substantive conclusions about significance and even direction of coefficients. See the fuller discussion in Garson (2012: ch. 2). Among the reasons for dfferent results are these:

* + - 1. *Different estimation algorithms*. HLM 7 uses expectation maximization (EM) with Fisher scoring every fifth iteration. This is different from what SAS and SPSS use. The practical result is that deviance (-2LL) may differ between packages, with the algorithm used by HLM tending toward more optimistic (lower) estimates of model fit as well as being slower but more likely to converge. Variance components of model random effects may also differ between packages.
      2. *Different starting values* Software packages differ in the starting values each employs prior to applying their respective estimation algorithms.
      3. *Different correlation constraints*. HLM 7 constrains the correlation of random effects to be no greater than absolute 0.997 at maximum, whereas SPSS allows correlation greater than absolute 1.0.
      4. *Different covariance structure constraints*. SPSS and SAS support a variety of covariance types as constraints on the solution. In particular, the variance component (VC) type is the default in SPSS and SAS, which requires covariances among random effects to be zero, whereas by default HLM 7 estimates all coefficients in the variance-covariance matrix, akin to the unstructured covariance type in SPSS and SAS. The HLM 7 default can be overridden by checking the "Diagonalize tau" radio button in its "Estimation" window, causing all covariances to be zero as in diagonal and variance components models. For unstructured solutions, convergence under the EM estimation algorithm in HLM 7 is far less apt to be a problem than in the algorithms used by SPSS or SAS, where Hessian and convergence warnings are not uncommon when an unstructured solution is requested. How covariances are constrained sometimes leads to significant differences in estimates. Also, HLM 7 easily supports the modeling of heterogenous residual variance, which in turn yields different deviance values for what is otherwise the same model.
      5. *Diffent algorithms for signifance of fixed effects*. Software packages differ in the algorithms used to calculate degrees of freedom used in the denominator of the formula for significance tests of fixed effects.
      6. *Different tests for the significance of random effects*. HLM uses chi-square tests to test the significance of variance components, while SPSS uses Wald tests, resulting in slightly different significance findings between the two programs.
      7. *Differences in parameterization of variables.* If one has a binary variable (such as "sector" in the school examples above), HLM will treat it as a covariate. SPSS results will be similar if "sector" is entered as a covariate, but if entered as a factor, signs will be reversed. That is, for factors, SPSS predicts the lower value (sector=0) and makes the higher value (sector=1) the reference. For covariates, the highest value (sector=1) is predicted. Similar differences will exist for multinomial variables entered as factors in SPSS.

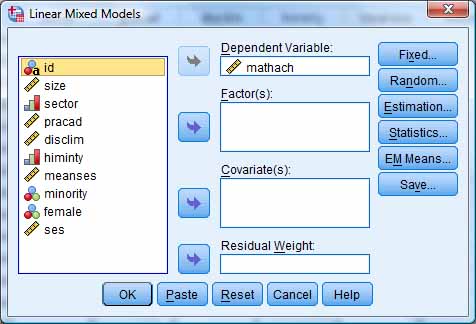
Note that in more complex models, the seemingly minor differences between packages are more likely to be compounded into significantly different reults.

**The Null Model and Intraclass Correlation**

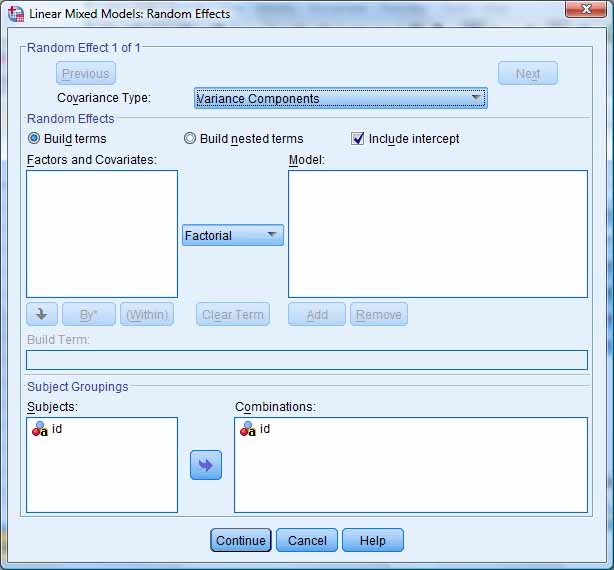
* + **Overview**.
    - *Description.* Also called the *intercept-only model*, the *unconditional model*, or the *one-way ANOVA model with random effects*. The null model, which is a type of random intercept model, predicts the dependent based on an intercept term and an error term at level 1 (ex., the student level where the dependent is test scores). There are no predictors, either factors or covariates, at either level 1 (ex., student level) or level 2 (ex., school level). However, the level 1 intercept is predicted as a function of the level two grouping variable (ex., school). Specifically, the level 1 intercept is predicted on the basis of the level 2 intercept (representing mean test score among schools in this example) plus a level 2 error term (representing unmeasured factors unique to each school). Null models are used for three primary purposes:
      1. To compute the intraclass correlation coefficient (ICC), which is used to test the appropriateness of going ahead with linear mixed modeling.
      2. To compute baseline [deviance](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#2lltests), which often is labeled "-2LL." Deviance is used in measuring significant improvements with models which do include fixed and random predictors other than the grouping variable.
      3. To compute the [reliability](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#reliab) of differences in outcomes between groups. Findings for groups with fewer level 1 observations (ex., schools with fewer students) will be less reliable. By a common rule of thumb, reliability should be above .80 in confirmatory research (though some use .90).
      4. *Why LMM?*. One could approach such an example with a GLM procedure such as fixed-effects analysis of variance, testing differences in mean scores across agencies. However, fixed-effects ANOVA assumes independent observations, whereas in fact it is possible, even likely, that value of the dependent variable are biased up or down depending on the nature of the level 2 grouping units (ex., schools). This can be handled by declaring the level 2 unit (school) to be a random factor in LMM.
      5. *ICC as a test of HLM appropriateness*. ICC is a test of whether hierarchical linear modeling is needed. HLM is not appropriate when ICC is non-significant A significant ICC means there are significant intercept differences between the groups which form the higher levels. If ICC approaches zero, there is no between-groups effect. And if there is no between-groups effect, there is no need to model individual-level regression parameters as random effects of a higher or grouping level. That is, the lower the ICC, the less difference hierarchical linear modeling or linear mixed modeling will make in predicting the dependent (ex., test score) compared to traditional regression techniques.
         * *Warning*. Note that unlike random intercept models, random coefficents models, discussed [below](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#rcneeded), require more than ICC to test appropriateness. That is, if the ICC test for the null model, which is a form of random intercept model, shows mixed modeling is required, it will be required for both random intercept and random coefficients models in which the null model is nested. However, if the ICC test suggests mixed modeling is not required, the researcher can only conclude it is not required for random intercepts models; the test is not sufficient to rule out other higher level effects on level 1 slopes. Typically this is not an issue as when data are nested or cross-classified, there is a higher level effect on the intercepts.
  + **The null model and ICC in SPSS software**.
    - *Description.* The null model is a model without predictors at any level, but with data grouped within at least one higher level. For example, in two-level mixed models, the base layer (level 1) is individuals (ex., students) who are clustered within the groups formed by a second, higher level (ex., level 2 = schools). Intraclass correlation (ICC) tests whether there is a grouping effect on the dependent variable, in which case LMM is required in lieu of ordinary regression methods.
      1. *Example*. Consider the example of student math scores grouped by school. There is a group effect, also called the "between-groups" effect, if the grouping variable (id, which is the id for school, not student) has an effect on the intercept of math score over and beyond whatever other test-level (individual-level) variables are used to predict test score (ex., student age, student IQ, student socioeconomic status, etc.). (Data used are from sample files supplied with HLM 7 software, to maintain parallel treatment).
      2. *Requesting output in SPSS*. The "intraclass correlation" (ICC) is the between-groups effect divided by the total effect for the null model. These effects will be the variance components shown in the SPSS variance components table. Although SPSS does not compute ICC directly, this is easily accomplished manually. The researcher enters a null model - one with just a level 1 dependent such as MATHACH scores and a level 2 grouping variable such as "id" (school id). The id variable is entered as a "Subjects" variable in the initial linear mixed models dialog, as shown below:



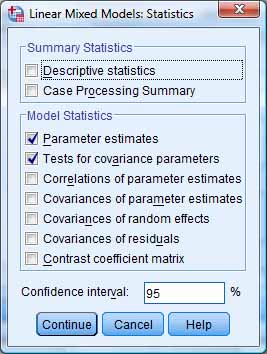
After pressing the Continue button on the initial LMM dialog screen, the screen below appears and the dependent variable, MATHACH, is entered as the dependent variable:



After pressing the Random button on the screen above, the researcher enters the grouping variable, id (which is school id) as the subject groupings/combinations variable. Also, the "Include intercept" box is checked. This means that the intercept for mth achievement is modeled as a random effect of school as the level 2 grouping variable. A "Variance Components" model including intercept is requested (this is the default), as shown below:

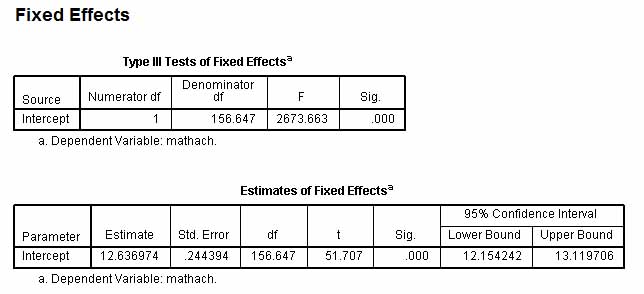


Back under the Statistics button dialog, the researcher selects desired output, such as parameter estimates and tests for covariance parameters (the latter is the minimum output to obtain ICC).



After clicking Continue, OK, the model is run and the output below is obtained.

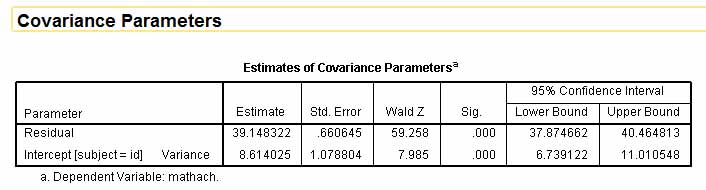
* + - 1. *Fixed effects*. The only fixed effect in an unconditional random effects model is the constant (intercept). Thus in SPSS output there is a table labeled "Tests of Fixed Effects" with one row for the intercept as the only fixed effect. In this example, the estimate column for the Intercept row gives the overall school mean math achievement score. Also listed is a t-test of the significance of the mean, and 95% confidence intervals. For a significant mean, zero will not be within the lower and upper confidence bounds. If the intercept is significant (it almost always is, as here) then we can be 95% confident the intercept is different from 0. Usually this is of no interest.



If under the Statistics button we checked "Parameter estimates," then a second table of "Estimates for Fixed Effects" will appear, where the "Estimate" column will contain the actual estimated value of the intercept along with its confidence interval. If the intercept is significant, the value 0 will not be within the confidence interval.

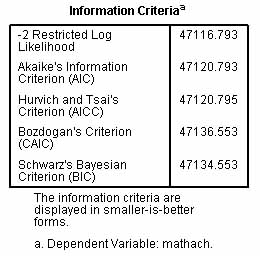
In later models, when additional fixed factors or covariates are added (making the random effects model conditional rather than unconditional), fixed effect parameter estimates include regression slopes. **These fixed effects parameters are interpreted as regression coefficients. For instance, consider a model similar to the example but adding student socio-economic status (SES) as a fixed covariate. If the SES parameter estimate is 2.5, then for a unit increase in SES the dependent variable (ex., test score) will increase 2.5 units - the same interpretation as in ordinary regression. If the fixed effect is a 0, 1 dichotomy such as men=0, women=1, then a coefficient of 2.5 would mean the mean value for women is 2.5 units higher than for men on the dependent variable. The intercept is interpreted as the overall mean of the dependent variable when other factors and covariates are zero**.

* + - 1. *Random effects*. In the "Estimates of Covariance Parameters" table, for the row "id [subject=id] Variance", the Estimate column gives the estimated variance of the intercepts between schools. The Estimate column for the Residual row gives the estimated variance within schools (in this case, among students). Comparing the individual-level variance with the group-level variance informs the researcher about whether the variation in test scores is primarily within or between schools. In this case it is primarily within schools, by a ratio of over 4:1. (Note Singer (1998) and Singer & Willett (2003) state that the t-test of significance for the estimated variance may not be reliable.)
         * *Variance components*, called "covariance parameters" in SPSS output, appear in the "Covariance Parameters" table shown below. The "Intercept[subject=id]" component is the between-group (here, between-school) variance in math achievement explained by differences among schools. **The residual component is the variance in math score not explained by the between-groups effect represents variation within the set of individual students.** This effect is the "within-groups" effect. The "total effect" is, of course, the sum of the between- and within-groups effect.



In this example, although within-school variation is more important, since the school variance component is significant we conclude that math scores do vary by school. This further means that a fixed-effects analysis of scores ignoring the agency effect would violate the assumption of independence of observations since observations are biased up or down depending on school. If the value 0 appeared within the confidence limits for school, the random factor (school in this case) would not be significant and an analysis with just fixed effects factors (ex., SES or race) might be possible.

* + - * + *Wald test*. If the intercept variance component is significant in the null model, as it is above, then means of the dependent variable vary significantly across level 2 units (ex., math scores vary across schools). SPSS uses the *Wald statistic* to assess significance of variance components. The Wald statistic is the ratio of a coefficient to its standard error, resulting in a Z-value which is looked up in a table of the standard normal distribution to generate a corresponding probability (p) value, the normal cut-off for which is .05. The Wald test is discussed further in the section on [logistic regression](http://faculty.chass.ncsu.edu/garson/PA765/logistic.htm#Wald), including a warning about its unreliability for small samples.
        + *Intraclass correlation: interpreting random effects variance components*. Let the variance component estimate for the random factor id (meaning school id) = 8.61. Let the variance component estimate for Residual = 39.15. In this null model, since the school variance component is 18% of the total of both variance components, we would say that the school effect accounts for 18% of the variance in math scores. This is significant because the Wald test for the intercept variance component is significant. We can also say that scores cluster by school, meaning that two students randomly selected from the same school are more likely to have similar scores than a pair of randomly selected students representing different schools. This ratio (.18) of the between-school variance component to the total of variance components is the "intraclass correlation coefficient". Note ICC as computed in this manner applies only to the null model. In other variance components random intercept models with additional random effects, the ratio is the between-school effect controlling for other predictors in the model.
        + *Reliability*. The "Estimates of Covariance Parameters" table for the null model also contributes to assessing the reliability of findings for each particular level 2 unit (ex., school). Given School 3 has 25 students, then its reliability = intercept component/[intercept component + (residual component/group sample size)] = 8.61/[8.61 + (39.15/25)] = .85. Using the common rule of thumb that reliability should be .80 or higher, the researcher would conclude that findings pertaining to this particular school would be reliable.
      1. *Goodness of fit*. Goodness of fit measures for the model appear in the "Information Criteria" table of SPSS LMM output. This table contains five goodness of fit measures: -2 Restricted Log Likelihood (assuming REML rather than ML estimation was chosen), Akaike Information Criterion (AIC), Hurvich and Tsai's Criterion (AICC, meaning AIC Corrected, for finite sample corrected AIC), Bozdogan's Criterion (CAIC, consistent AIC), and Schwarz's Bayesian Criterion (BIC).

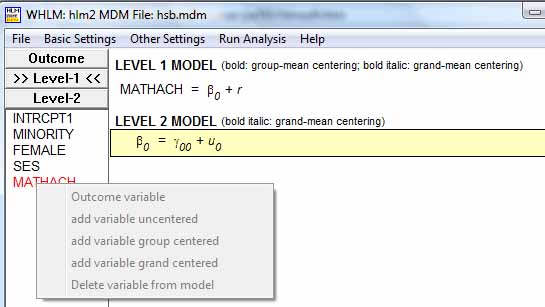


For all five measures, the lower the value the better the model. When comparing models, the model with the lower values is the best-fitting. These measures are discussed further in the section on [structural equation modeling](http://faculty.chass.ncsu.edu/garson/PA765/structur.htm#infotheory). The null or unconditional model serves as a useful baseline model to compare with other models discussed below, using these measures.

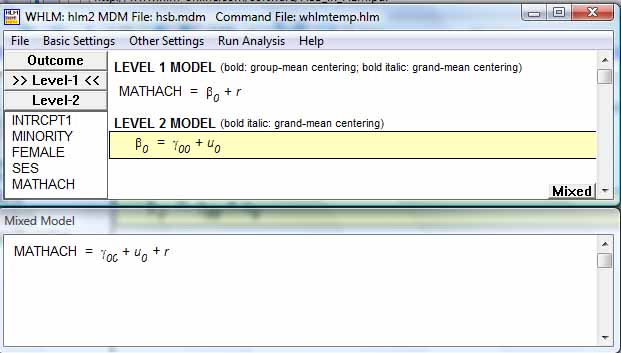
* + - * + *Deviance*, also called model discrepancy, is "-2 Restricted Log Likelihood" (-2RLL) in the SPSS table above, assuming REML estimation is used. If ML estimation is used, it is "-2 Log Likelihood" (-2LL). In the context of multilevel modeling, deviance is a measure of the lack of fit between the model and the data. Therefore, lower deviance is better.
        + *Likelihood ratio tests for model differences*. When considering whether to drop a term from the model, the likelihood ratio test is preferred over the Wald test. The model is run with and without the term in question and the difference taken between the two -2LL coefficients (from the "Information Criteria" table above). This is also called the model chi-square difference. The degrees of freedom are 1 (assuming one term is dropped, otherwise df = number of terms dropped, which is the difference in df between the two models). The df is thus the difference in the number of model parameters, which is listed in the default "Model Dimensions" tables at the top of SPSS output. The probability of a model chi-square difference that large or larger with given degrees of freedom can be looked up in a chi-square table, or can be obtained in SPSS under Transform, Compute, then entering the fomula sig.chisq2(d, df), where d is model chi-square difference and df is the degrees of freedom. If the computed probability is > .05, then the difference in deviance cannot be assumed to be different from 0 and the researcher proceeds with dropping the variable or term in question on the grounds of parsimony (equally good fit with one less variable).

*Likelihood ratio tests for model fixed effects*. If the term in question is a random effect, the likelihood ratio test may be used under either ML or REML estimation. However, if the term is a fixed effect, ML estimation is assumed.

* + - 1. *Saving new variables*. Click the "Save" button in the Linear Mixed Models dialog to have SPSS save to your dataset new variables representing predicted values, standard errors, degrees of freedom, and/or residuals for each case.
  + **The null model and ICC in HLM software.**
    - *Description.* Like other two-level HLM models, the null model is created in the HLM 7's WHLM dialog. If the MDM file was previously saved (on .mdm files and data input in HLM 7, see the [FAQ section](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#hlmentry)), then from the HLM menu select File, Create a new model using an existing MDM file, and open the appropriate .mdm file to arrive at the WHLM modeling dialog.
      1. *The WHLM modeling dialog.* In the example below, the level 1 (individual level) variable MATHACH math achievement scores is selected as the outcome (dependent) variable. As variables are added to the model, HLM shows the level 1 and level 2 model equations on the right.



* + - 1. *The null (intercept-only) model*, also called the *unconditional model*, is used to establish a baseline against which to compare more complex models with level 1 and/or level 2 predictors. Clicking the "Mixed" button at the bottom of the WHLM dialog creates the combined HLM equation shown below, which is the intercept-only (null) model.



* + - * + *Example.* Click "Run analysis" from the menu to generate output for the intercept-only model, shown below in excerpted form. Click File, View output to view. B0 is the level 1 intercept. Y is the dependent, MATHACH math achievement scores in this example. Math achievement in this null model is a function of the intercept plus a residual (error) term. The level 1 intercept, B0, in turn is a function of the grand mean across level 2 units (schools) plus a random error term, signifying the intercept is modeled as a random effect. The id variable contains the id value for any given level 2 unit (school).
        + Program: HLM 6 Hierarchical Linear and Nonlinear Modeling
        + ...
        + -------------------------------------------------------------------------------
        + Module: HLM2S.EXE (6.06.2857.2)
        + ...
        + SPECIFICATIONS FOR THIS HLM2 RUN
        + The data source for this run = hsb.mdm
        + ...
        + The outcome variable is MATHACH
        + ...
        + Summary of the model specified (in equation format)
        + ---------------------------------------------------
        + Level-1 Model
        + Y = B0 + R
        + Level-2 Model
        + B0 = G00 + U0
        + Iterations stopped due to small change in likelihood function
        + \*\*\*\*\*\*\* ITERATION 4 \*\*\*\*\*\*\*
        + Sigma\_squared = 39.14831
        + Tau
        + INTRCPT1,B0 8.61431
        + ...
        + *Sigma-squared* is the level 1 variance in the intercept, reflecting mean MATHACH of students within schools. It corresponds to the "Residual" estimate in the SPSS "Estimates of Covariance Parameters" table.
        + *Tau* is the level 2 variance between schools of the intercept. It corresponds to the "Intercept[subject=id] Variance" estimate in the SPSS "Estimates of Covariance Parameters" table.
        + *Intraclass correlation (ICC).* In a variance components model (which assumes independent random effects with uncorrelated error, hence zero covariance of error terms for random effect slopes and intercepts), the total variance is sigma-squared plus tau: 39.15 + 8.61 = 47.76. Also, as here, if there is only one random effect, there is no term for covariance of error terms. As will be discussed below, however, when there are two or more random effects, and when a variance components model does not apply, the computation of total variance must include covariance of error terms.

The intraclass correlation (ICC) is the percent of variance in MATHACH attributable to the between schools effect, calculated in this example as = 8.61/ 47.76 = .18. The proportion of MATHACH attributable to within-schools effects in this null model is = 1 - .18 = .82. ICC varies from +1.0 when group means differ but within any group there is no variation, to -1/(n-1) when group means are all the same but within-group variation is very large. ICC is sometimes used to assess the utility of applying a hierarchical linear model. At the extreme, when ICC approaches 0 or is negative, LMM is not appropriate. For this example, LMM is appropriate.

* + - * + *Reliability*. HLM treats the intercept as a random effect. A regression was run on each of the 160 level 2 units (schools). The reliability of the intercepts was .901. If the reliability were 1.0, there would be no difference between estimates of slopes and intercepts in HLM compared to OLS regression. In reality, reliability is always less than 1.0. The lower the reliability, the more HLM and OLS estimates will diverge because when calculating the intercept (or in later models, the slopes of the predictors) HLM weights the 160 coefficients such that intercepts or slopes for schools with greater reliability count more than those for schools with lower reliability.
        + ----------------------------------------------------
        + Random level-1 coefficient Reliability estimate
        + ----------------------------------------------------
        + INTRCPT1, B0 0.901
        + ----------------------------------------------------
        + *Likelihood function*. The likelihood function is used to calculate the deviance. Specifically, -2 times the likelihood function is the deviance, sometimes labeled -2LL. As discussed [below](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#hlmdeviance), it is used as a baseline value in likelihood ratio tests used to compare model fit across nested models.
        + The value of the likelihood function at iteration 4 = -2.355840E+004
        + The outcome variable is MATHACH
        + *Fixed effects*. In a null model there are no predictors, so the only fixed effect is the intercept itself, which in this example is the mean of math achievement for students grouped by school. That is, the intercept representing the mean is a random effect of the level 2 grouping variable, school.
        + Final estimation of fixed effects:
        + ----------------------------------------------------------------------------
        + Standard Approx.
        + Fixed Effect Coefficient Error T-ratio d.f. P-value
        + ----------------------------------------------------------------------------
        + For INTRCPT1, B0
        + INTRCPT2, G00 12.636972 0.244412 51.704 159 0.000
        + ----------------------------------------------------------------------------

*Calculating confidence limits.* The Level 1 intercept is 12.64, representing the mean math achievement score among schools. The 95% confidence limits on this value of B0 may be calculated as follows: Since tau is the variance of the intercept, its square root is the standard deviation: s.d. = SQRT(tau) = SQRT(8.61431) = 2.9350. The 95% level corresponds to plus or minus 1.96 standard deviations, so 95% of the 160 school regressions may be expected to have an intercept (mean score) between a high of 12.64 + 1.96\*2.935 = 18.39 and a low of 12.64 - 1.96\*2.935 = 6.88.

*Fixed effects with robust standard errors*. Robust standard errors are advisable when there is misspecification of the distribution of the dependent variable. Therefore significant differences between the ordinary and robust estimates of the standard error may flag a problem with the distribution specified by the researcher (normal is default). This does not appear to be a problem for this example. This specification may be changed in the "Basic Settings" dialog.

Final estimation of fixed effects

(with robust standard errors)

----------------------------------------------------------------------------

Standard Approx.

Fixed Effect Coefficient Error T-ratio d.f. P-value

----------------------------------------------------------------------------

For INTRCPT1, B0

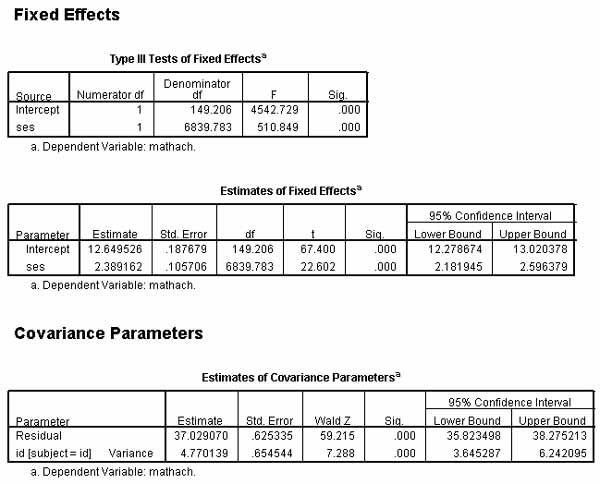
INTRCPT2, G00 12.636972 0.243628 51.870 159 0.000

----------------------------------------------------------------------------

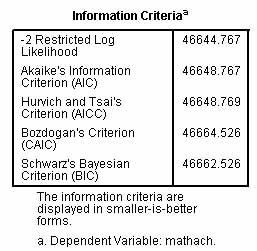
* + - * + *Random effects and model significance*. Random effects are displayed in the variance components table. That the intercept is significant means there are level 2 effects on the dependent, justifying hierarchical linear analysis. By the same token, the null model as a whole is significant, meaning the intercept is significantly different from 0 (also reflected in the fact that 0 is not within the confidence limits calculated above). (The variance components are the tau and sigma-squared values discussed above. The standard deviations are their square roots. Degrees of freedom are (n - 1), where n is the 160 schools in the sample.)
        + Final estimation of variance components:
        + -----------------------------------------------------------------------------
        + Random Effect Standard Variance df Chi-square P-value
        + Deviation Component
        + -----------------------------------------------------------------------------
        + INTRCPT1, U0 2.93501 8.61431 159 1660.23259 0.000
        + level-1, R 6.25686 39.14831
        + -----------------------------------------------------------------------------
        + *Deviance and the likelihood ratio test.* Deviance is the -2 log likelihood (-2LL) coefficient. The difference between the deviance for the null (intercept only) model and a model with level 1 and/or level 2 predictors is used in the likelihood ratio test to assess the effects of adding predictors to the model. The more adding a predictor lowers deviance, the more effect it has on predicting MATHACH. (Likelihood ratio tests are also discussed in the sections on [logistic regression](http://faculty.chass.ncsu.edu/garson/PA765/logistic.htm#lltests) and on [structural equation modeling](http://faculty.chass.ncsu.edu/garson/PA765/structur.htm#diff).)
        + Statistics for current covariance components model
        + --------------------------------------------------
        + Deviance = 47116.793477
        + Number of estimated parameters = 2
        + -----------------------------------------------------------------------------
    - *The null model and ICC in SAS software*. The null and other models in SAS are discussed in Garson, ed. (2012).

**Two-Level Random Intercept Models**

* + **A one-way ANCOVA model with random effects, using SPSS**
    - *Overview*. This is a random intercept model with level 1 covariates. It is similar to the null model discussed above but includes at least one level 1 covariate. It is also a type of "conditional random effects model" since, unlike the null model, the variance in the dependent is conditional on a level 1 predictor. There are no level 2 predictors but the level 1 intercept is still predicted as a random effect of the level 2 grouping variable. Because the slopes of the level 1 predictors are not modeled as random effects, no variables are entered as random effects under the SPSS "Random" button dialog. Rather, level 1 predictors are declared under the "Fixed" button dialog. Because the intercept, reflecting the dependent mean, is modeled as a random effect of the grouping variable, the grouping variable (ex., school id) is entered as the Subjects/Combinations variable in the same dialog.
    - *Example:* For instance, differences in mean math scores may be analyzed as predicted by student socioeconomic status (SES) at level 1, predicting the level 1 intercept but not the SES slope using the level 2 grouping variable, school, as a random effect. There will be one regression of math score on SES for each school, generating as many intercepts as schools. This enables estimation of the between-schools versus within-schools variability in math scores.
    - *SPSS*. For the example where the level 2 nesting factor is ID (school ID) in a study of causes of math scores, then in SPSS select Analyze, Mixed Models, Linear; in the "Linear Mixed Models: Specify Subjects and Repeated" dialog box, assign ID to the Subjects listbox (the Repeated box is left blank). Click Continue. In the "Linear Mixed Models" dialog, assign the dependent variable (MATHACH) to the dependent variable listbox. List ID in the factor(s) box. List SES in the covariates box. Click the Fixed button and make SES a fixed variable. Click the Random button and set (school) ID as the grouping variable in the Combinations area. In the same Random dialog, check "Include intercept" and set the Covariance Type to Variance Components, which is the default for random intercepts models. (Note Unstructured may also be used, but will give a solution characterized by fewer constraints and less parsimony). Click Continue, OK, to run the model.
    - *Fixed and random effects* are interpreted largely as described in the previous section on the null model.
      1. *Fixed effects*. For this example, student SES is shown to be a significant predictor at level 1 of the level 1 dependent, math achievement. Note that the slope of SES in this model may be significantly different from the SES slope in OLS regression.



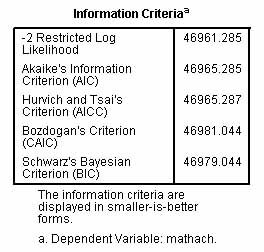
* + - 1. *Random effects*. The "id[subject=id] Variance" estimate reflects the between-schools variation in math scores. The "Residual" estimate reflects the within-schools variation in math scores.
         * *Unexplained variance*. The variance component for the Residual parameter is large and significant, which means there remains significant variance in math scores even after student-level SES is controlled.
         * *Effect size as reduction in between group effects*. Looking at the variance components, which total to 41.80, the between-schools component (4.77) is 11% of the total. We may say that the between-school effect accounts for 11% of the variance in math scores once student-level SES is controlled. We may note that controlling for student-level SES reduced the between-school effect from 18.0% in the null model discussed above to 11.4% in the random intercept model with level 1 covariates (a.k.a. "ANCOVA model with random effects"). This is a 6.6% reduction. Since 6.6/18 rounds to .37, we may say that controlling for student-level SES reduced the between-school effect by about 37%. In another calculation, we may note that the difference in the between-group variance component in the null model (8.59) and in the example random intercept model (4.7701) is 44.5% of that in the null model (8.59). On this basis we may say that controlling for student SES accounts for 44.5% of the between-group variability in student math scores.
         * *Effect size as reduction in within group effects*. Similar calculations can be made for within group effects, using the Residual variance components in the "Estimates of Covariance Parameters" table. The difference in the within-group variance component in the null model (39.24223) and in the example random intercept model (37.02907) is 5.37% of that in the null model (39.24223). On this basis we may say that controlling for student SES accounts for 5.37% of the within-group variability in student math scores. Most of the within-school variability remains, as noted above.
    - *Likelihood ratio test.* The deviance (-2 log likelihood above) is 46644.23 for this model compared to 47116.79 for the null model. Subtracting, we can say the random intercept model with level 1 covariates is a better fit because its deviance is approximately 472.5 lower. From the "Model Dimension" tables (not shown), we can read that in these examples, the null model had 3 parameters and the random intercept model with level 1 covariates had 4 parameters, a difference of 1. Looking in a chi-square table for 1 df, at the .05 significance level the critical value is 3.84; at .01 it is 6.64; at .001 it is 10.83. Since 472.5 is far greater than the critical value for even the .001 level. We may say that the random intercept model with level 1 covariates is different from and better than the null model at a significance level better than .001.



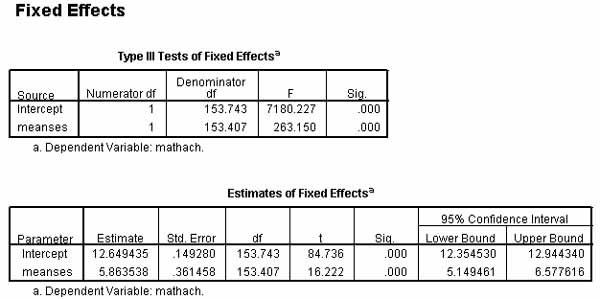
* + **A two-level random intercept regression model (means-as-outcomes regression model), using SPSS software.**
    - *Overview*. Also called a "means-as-outcomes regression model," this is a random intercept model in which the level 1 intercept is predicted on the basis of the level 2 grouping variable and also on the basis of one or more level 2 random effect predictors. There are no level 1 covariates, and accordingly no slopes are estimated. For instance, differences in mean math achievement scores (the intercepts) may be analyzed, predicting the level 1 intercept in terms of the between-groups effect of school and the level 2 random effect variable meanses (representing mean socioeconomic status of schools).
    - *SPSS.* In SPSS, select Analyze, Mixed Models, Linear; in the Specify Subjects and Repeated dialog, enter id (which is school id) as the Subjects variable; in the next "Linear Mixed Models" dialog, enter math achievement as the dependent variable and enter meanses as the covariate; click the Fixed button and set Model to meanses, making sure the Include Intercept checkbox is checked. Then click the Random button and make id the Subject Grouping in the Combinations area, making sure the Covariance Type is Variance Components and Include Intercept is checked. Click the Statistics button to select output options such as Parameter Estimates and Tests for Covariance Parameters. Click OK.

Example notes: The level 2 grouping variable, school id, is entered as the Subjects variable (signifying individuals are independent observations within each school) and under the Random button as the Subject Grouping variable (signifying it defines grouping at level 2). This causes LMM to compute separate regressions for each school, with the intercepts treated as random effects. School id is not entered as a random effect factor as that is already assumed by making it a Subject (grouping) variable. Meanses is a fixed factor and, being level 2, is not a random effect of some higher level. Therefore meanses is not entered under the Random button as a random effect to be modeled.

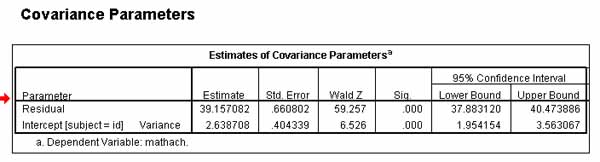
* + - *Information Criteria table*. Goodness of fit measures for the model appear here, primarily of use when comparing models, with lower log likelihood or lower information criteria values reflecting better fit. For these data, fit is better than the null model but not quite as good as the ANCOVA with random effects model (where level 1 ses was a predictor and meanses was not). Likelihood ratio tests of the differences of the random intercept regression model with either the null model or with the ANCOVA with random effects model can be run using the "-2 Restricted Log Likelihood" coefficient below. How such tests are computed is described in the section [above](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#lratiotest).



* + - *Tests of Fixed Effects table*. For the intercept and fixed effects (here, meanses), this table shows the significance level. If Sig. ≤ .05 for the effect, as it is in this example, it is retained in the model and means math achievement score and meanses are significantly related within school id, the grouping variable.



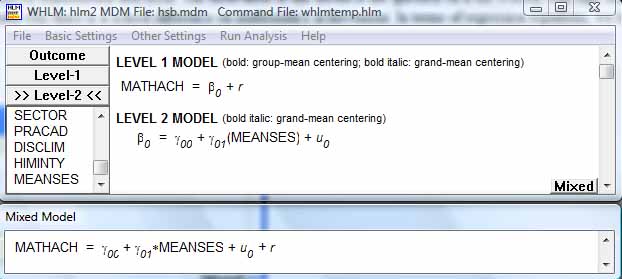
* + - *Estimates of Fixed Effects table*. For the intercept and fixed effects (only meanses in this example), this table shows the parameter estimate, its standard error, its significance level, and its confidence intervals. These estimates will differ from the b coefficients in the corresponding OLS regression because intercept by school id is a random effect. The estimate for the intercept is the estimate of the dependent variable, math achievement score, when meanses is controlled (is zero). This is more easily interpretable by centering meanses prior to analysis, so its mean is zero, in which case the intercept becomes the value of math achievement score when meanses is at its mean. The estimate for meanses is interpreted the same as in regression: math achievement score increases by 5.86 for each unit increase in meanses (here coded -1, 0, +1).
    - *Estimates of Covariance Parameters table*. This table gives the covariance parameter estimates for Residual and for Intercepts within Subject=id, along with a Wald test of significance.
      1. Intercept [subject=id] Variance. If significant, this means that intercepts of math achievement score predicted by meanses vary significantly between schools. That is, this is the between-subjects effect, where school id is the subject variable in this example. We may say there is a significant between-school effect. For this model, the covariance component is Intercept [Subject = id] and is *conditional*, controlling for meanses as a covariate. That is, its share of the total of parameter estimates in the HLM model is the percent of variance in math achievement scores attributable to differences between schools after meanses is controlled.
      2. Residual. If significant, this means that math achievement scores vary significantly within schools. That is, this is the within-subjects effect, where school id is the subject variable.
      3. Comparison of HLM with random effects models (REM). For the same example, HLM enters school id as a subjects variable on the opening SPSS LMM dialog and as the subject groupings/combinations variable under the Random button dialog. REM does neither, but instead enters school id as a factor and under the Random button dialog enters school id as a random effects variable to model, in the model section. One will get the same effect (parameter) estimates under either model, but the standard errors of the intercept will differ somewhat, as will the information criteria goodness-of-fit measures.



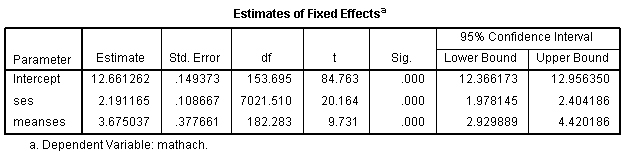
* + - *Model development.* The large residual variance component above indicates the need for a more complete model. The researcher may add additional level 1 (ex., individual) and/or level 2 (ex., school) factors or covariates to the model, and may model their main effects and interactions. These additional variables may be entered as fixed effects and/or random effects (the same variable may be considered both a fixed and a random effect). A typical strategy would be to enter a full factorial model for fixed effects, then drop the effects (higher level interactions are often non-significant, for example) found non-significant in the "Estimates of Fixed Effects" table, then re-run the analysis. Likewise, if the variance of slopes involving a random effect are found to be not significant in the "Estimates of Covariance Parameters" table, that variable may be removed as a random effect (not necessarily as a fixed effect, which it may be also). Information criteria measures (ex., AIC, BIC) may be used to compare models, with lower being better fit.

* + **A two-level random intercept regression model (means-as-outcomes regression model), using HLM software.**
    - *Example.* In this example, a level 1 dependent variable (MATHACH, math achievement score) is modeled by the level 2 grouping effect of schools as well as by a level 2 covariate, MEANSES (the school-level mean SES of students in a given school). As there is no level 1 covariate, there is no level 1 slope to model as a random effect. Rather, the purpose is to see if schools low on mean MEANSES are also low on MATHACH score, or if high on MEANSES are high on MATHACH also.

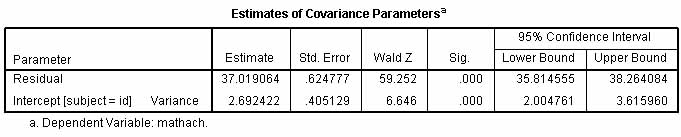
The model as entered in HLM 7's WHLM interface is shown below. For this example, MEANSES is coded as -1, 0, or +1 depending on whether it is low, average, or high in mean SES. MEANSES is entered into HLM as a level 2 variable, uncentered due to this coding. Click "Run analysis" from the HLM menu to generate output for the means-as-outcomes model, shown below in excerpted form. Note: for this example, we have clicked on Other Settings, Output Settings, and have unchecked the default "Reduced output" check box. Consequently the output below has enhanced output. Click File, View output to view.



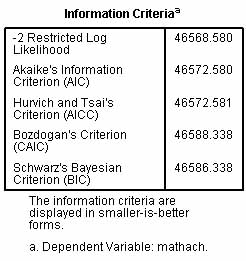
* + - 1. *The model*. At the top of the output, the model is reproduced in print form shown below. In the means-as-outcomes model, MATHACH (Y) is a function of the intercept and a residual term. The intercept is a function of the grand mean of MATHACH across the 160 schools plus a level 2 slope (the G01 b coefficient) times MEANSES, plus an error term.
      2. Module: HLM2S.EXE
      3. The data source for this run = hsb.mdm
      4. ...
      5. The model specified for the fixed effects was:
      6. ----------------------------------------------------
      7. Level-1 Level-2
      8. Coefficients Predictors
      9. ---------------------- ---------------
      10. INTRCPT1, B0 INTRCPT2, G00
      11. MEANSES, G01
      12. The model specified for the covariance components was:
      13. ---------------------------------------------------------
      14. Sigma squared (constant across level-2 units)
      15. Tau dimensions
      16. INTRCPT1
      17. Summary of the model specified (in equation format)
      18. ---------------------------------------------------
      19. Level-1 Model
      20. Y = B0 + R
      21. Level-2 Model
      22. B0 = G00 + G01\*(MEANSES) + U0
      23. *OLS regression estimates*. Least squares estimates are part of the enhanced output normally suppressed when the default "Reduced output" output option is not overridden. OLS regression estimates of intercepts and slopes, unlike HLM estimates, do not treat either as random effects reflecting the random variation among schools (the level 2 units). Comparison of these estimates with LMM estimates further below is one way of assessing the school-level effect on math achievement.
      24. Least Squares Estimates
      25. -----------------------
      26. sigma\_squared = 41.72661
      27. Least-squares estimates of fixed effects
      28. ----------------------------------------------------------------------------
      29. Standard
      30. Fixed Effect Coefficient Error T-ratio d.f. P-value
      31. ----------------------------------------------------------------------------
      32. For INTRCPT1, B0
      33. INTRCPT2, G00 12.712760 0.076215 166.801 7183 0.000
      34. MEANSES, G01 5.716800 0.184286 31.021 7183 0.000
      35. ----------------------------------------------------------------------------
      36. Least-squares estimates of fixed effects
      37. (with robust standard errors)
      38. ----------------------------------------------------------------------------
      39. Standard
      40. Fixed Effect Coefficient Error T-ratio d.f. P-value
      41. ----------------------------------------------------------------------------
      42. For INTRCPT1, B0
      43. INTRCPT2, G00 12.712760 0.149501 85.035 7183 0.000
      44. MEANSES, G01 5.716800 0.326862 17.490 7183 0.000
      45. ----------------------------------------------------------------------------
      46. The least-squares likelihood value = -23600.617867
      47. Deviance = 47201.23573
      48. Number of estimated parameters = 1
      49. *Starting values*. This section, almost entirely truncated below, is part of the enhanced output when requested. Starting values are not interpreted when writing up results but may be inspected to better understand the computation process. It may be useful as a diagnostic tool when trying to understand one potential cause of differences in estimates across software packages.
      50. STARTING VALUES
      51. ---------------
      52. sigma(0)\_squared = 39.14163
      53. ....
      54. *Reliability*. The higher the reliability, the less OLS and HLM estimates will diverge. By rule of thumb, a reliability of .70 or greater is "high." Therefore for these data we would not expect large divergence.
      55. ----------------------------------------------------
      56. Random level-1 coefficient Reliability estimate
      57. ----------------------------------------------------
      58. INTRCPT1, B0 0.740
      59. ----------------------------------------------------
      60. *Fixed effects.* Note that the final estimates for the intercept and MEANSES slope (12.65 and 5.86 respectively below) correspond to the "Estimate" for "Intercept" and "MEANSES" respectively in the "Estimates of Fixed Effects" table in SPSS linear mixed model output for the same model and data. Below, the intercept and the slope of MEANSES are both significant. These estimated coefficients are interpreted as in regression. The intercept of 12.65 is the predicted MATHACH score when MEANSES is 0 (recall MEANSES coded 0 meant medium rather than high or low mean SES for a school). Again, the trivial differences between the ordinary and robust estimates indicate that the default normal distribution assumption for MATHACH (the outcome variable) need not be rejected.
      61. Final estimation of fixed effects:
      62. ----------------------------------------------------------------------------
      63. Standard Approx.
      64. Fixed Effect Coefficient Error T-ratio d.f. P-value
      65. ----------------------------------------------------------------------------
      66. For INTRCPT1, B0
      67. INTRCPT2, G00 12.649436 0.149280 84.736 158 0.000
      68. MEANSES, G01 5.863538 0.361457 16.222 158 0.000
      69. ----------------------------------------------------------------------------
      70. Final estimation of fixed effects
      71. (with robust standard errors)
      72. ----------------------------------------------------------------------------
      73. Standard Approx.
      74. Fixed Effect Coefficient Error T-ratio d.f. P-value
      75. ----------------------------------------------------------------------------
      76. For INTRCPT1, B0
      77. INTRCPT2, G00 12.649436 0.148377 85.252 158 0.000
      78. MEANSES, G01 5.863538 0.320211 18.311 158 0.000
      79. ----------------------------------------------------------------------------
      80. *Random effects*. In the table below, the level 2 intercept, U0 = 2.64, represents the between-schools variance in math score and is the "school effect". R = 39.16 represents the residual within-schools variance. These correspond in a SPSS linear mixed model to the "Estimate" for "Intercept[subject=id] Variance" and "Residual" respectively in the "Estimates of Covariance Parameters" table for the same model and data.
          * *Effect size and confidence limits*. Note that the variance component for the intercept is 2.64, much lower than the 8.61 value for the null intercept-only model, indicating the large impact of adding MEANSES to the model. Put another way, the estimate is now a partial coefficient, controlling for MEANSES. The difference is 5.97 and the ratio 5.97/8.61 = .69. The intercept-only variance component of 8.61 was the between-schools variance in MATHACH. Controlling for MEANSES added to the model reduces between-schools variance by 69%. Put another way, MEANSES explains 69% of the between-schools variance in math achievement scores. However, since the intercept is significant by the chi-square test, significant variation between schools still remains. Confidence intervals on the intercept could also be constructed from the coefficients below using ± 1.96\*sd in the same manner as in previously-discussed models.
      81. Final estimation of variance components:
      82. -----------------------------------------------------------------------------
      83. Random Effect Standard Variance df Chi-square P-value
      84. Deviation Component
      85. -----------------------------------------------------------------------------
      86. INTRCPT1, U0 1.62441 2.63870 158 633.51744 0.000
      87. level-1, R 6.25756 39.15708
      88. -----------------------------------------------------------------------------
      89. *Likelihood ratio test.* The "Deviance" estimate below corresponds to the "-2 Restricted Log Likelihood" value in the "Information Criteria" table in SPSS linear mixed model output for the same model and data. The likelihood ratio test: because the means-as-outcomes model and the intercept-only model both have 2 estimated parameters, degrees of freedom are 0 and while results are computed, the test of difference in model fit is not meaningful. However, as the random intercept regression deviance of 46959.4 is less than the null model deviance of 47116.8, adding MEANSES does result in better model fit, though significance is not tested.
      90. Statistics for current covariance components model
      91. --------------------------------------------------
      92. Deviance = 46959.446959
      93. Number of estimated parameters = 2
      94. Variance-Covariance components test
      95. -----------------------------------
      96. Chi-square statistic = 157.34304
      97. Number of degrees of freedom = 0
      98. P-value = >.500
  + **Random intercepts ANCOVA model**
    - *Overview*. This is a random intercepts model with predictors at both level 1 (ex., student SES) and 2 (ex., school mean SES, school sector - public or parochial). It is random intercepts because only the level 1 intercept (ex., of the dependent, math score) is modeled as a random effect. Level 1 predictor slopes are not modeled as random effects. Predictors from both levels are entered as fixed effects. No variable is entered as a random effect because slopes are not modeled. The grouping variable (ex., schoolID) is entered as the Subject variable and under the Random button dialog in SPSS, as the Subject/Combination variable.
    - *SPSS*. The school id variable id is entered as a Subjects variable in the opening LMM dialog. Mathach, which is math achievement scores, is entered next as a dependent. Both ses (student socioeconomic status at level 1) and meanses (school-level mean ses) are entered as covariates, then under the Fixed button are both entered in the Model area as main effects. Under the Random button, id is entered as the Subjects/Combinations variable and "Include intercept" is checked. Under the Statistics button, parameter estimates and tests of covariance parameters are checked. The resulting SPSS syntax looks like this:
    - MIXED mathach WITH ses meanses
    - /CRITERIA=CIN(95) MXITER(100) MXSTEP(10) SCORING(1) SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE) LCONVERGE(0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
    - /FIXED=ses meanses | SSTYPE(3)
    - /METHOD=REML
    - /PRINT=SOLUTION TESTCOV
    - /RANDOM=INTERCEPT | SUBJECT(id) COVTYPE(VC).
    - *Fixed effects*. The parameter estimates in the "Estimates of Fixed Effects" output table are interpreted as in other random intercept models discussed above. As in other forms of regression, the researcher must be careful to add the phrase "controlling for other variables in the model" when interpreting these estimates. For instance, if both individual level SES and school mean SES variables are significant, as they are below, this means that school mean SES has a significant effect on the dependent, math score, even when student-level SES is controlled.
      1. *Binary factors vs. binary covariates*. SPSS predicts the "0" value of factors but the "1" value of covariates. Results are equivalent but interpretation is affected. Given male=0, female= 1, when coded as a factor, the intercept estimate plus the parameter estimate for sex = 0 is the mean level of the dependent, math score, for males. When coded as a covariate, adding the parameter estimate for sex is the mean level of the dependent for females. The intercept estimate will be different depending on whether sex is entered as a factor or as a covariate. The sex=0 fixed effect estimate when entered as a factor will be opposite in sign to the fixed effect estimate for sex when sex is entered as a covariate. The substantive interpretation will be the same.



* + - *Random effects.* The "Estimates if Covariance Parameters" table will still have only two variance components: in the example, one for "Intercept[subject=id] Variance" (recall id is school id) and one for "Residual". As in previous models, the intercept variance component reflects the between-schools effect on the dependent (math score) and the residual variance component reflects the remaining within-schools variation in math score. Here, the great majority of the variance is a within-school or residual effect, indicating the need for a more complete model. The school effect is significant but minor in comparison.

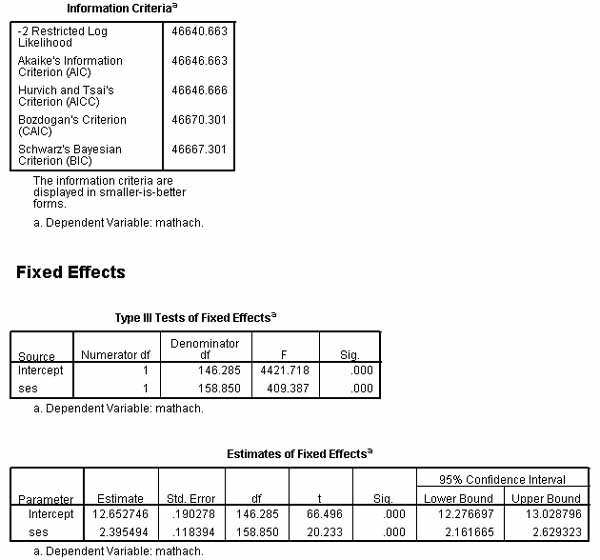


* + - *Deviance.* The deviance, shown as "-2 Retricted Log Likelihood," is lower than models above. A likelihood ratio test, described earlier but not illustrated here, could demonstate whether the difference was significant.

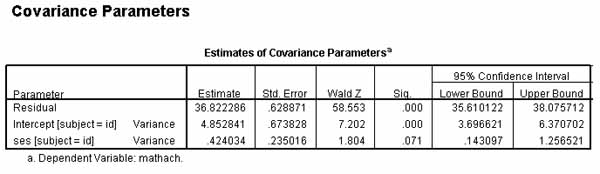


**Two-Level Random Coefficients Models**

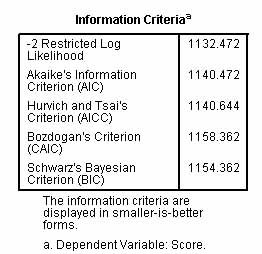
* + - **A random coefficients regression model using SPSS software.**
      1. *Overview*. Random coefficients regression models are a type of random coefficients (RC) model in which the level 1 slope(s) as well as the level 1 intercept is predicted on the basis of the level 2 grouping variable as a random factor. However, there are no level 2 predictors. For instance, differences in mean math scores (reflected in the level 1 intercept) may be analyzed as predicted by the level 1 covariate student socioeconomic status (SES), predicting the slope of SES as well as the level 1 intercept of math score as functions of the between-groups effect of the grouping variable (school) id at level 2. Such random coefficients regression models are unconditional at level 2 but conditional at level 1.
      2. *SPSS.* Select Analyze, Mixed Models, Linear. On the initial dialog, let Subjects=id, where id is the level 2 school id. On the next dialog page, set the Dependent to be mathach and set SES (socioeconomic status of student; this variable is pre-centered in the data) as a covariate. Click the fixed effects button and model SES, meaning SES is a level 1 predictor of mathach. Click the Random button and move id to the Combinations area. Check "Include Intercept" (not the default). Select "Variance Components" as the model assumption (this is the default). Also under the Random button dialog, model SES as a random effect, meaning the slope of the level 1 variable SES will be modeled as a random effect of the level 2 grouping variable id (which is school id), in which a separate regression will be computed for each school. Note SES is thus modeled both as a fixed and a random effect. Note school id is not listed as a random effect as that is already assumed by making it the Subject variable. Continue. Under Statistics, select Parameter Estimates and Tests for Covariance Parameters. Continue. OK.
      3. *Explanation*: Note the following regarding SPSS handling of the random coefficients regression model:
         * The level 1 intercept of math achievement score is modeled as a level 1 effect of student SES, as in ordinary regression, by virtue of listing SES as a fixed effect.
         * The level 1 intercept of math score also is modeled as a random effect of school id by having school id as the Subjects and random effects grouping combinations variable.
         * The level 1 slope of student-level SES is modeled as a random effect of school id by declaring SES to be a random effect to be modeled, under the Random button dialog.
         * The fixed effects output will have two coefficients for this example, reflecting the effects of (school) id as a grouping variable on the two level 1 parameters (intercept and slope of student-level SES).
         * The random effects output will have three coefficients for this example, reflecting the three variance components: (1) the Residual, reflecting within-school random effects after controlling for student SES; (2) the SES component, reflecting between-school random effects on the slope of SES; and (3) the Intercept component, reflecting between-school random effects on the level 1 intercept.
      4. *Goodness of fit*. The "Information Criteria" table illustrated below shows the deviance as "-2 Restricted Log Likelihood" = 46640.66. This is lower than the 47116.79 value in the null model, indicating the model with SES as a predictor is somewhat better, though not as good fit as the random interecepts model which included both student ses and school-level meanses. Likewise, the information theory measures (ex., BIC) are lower, which also is better.



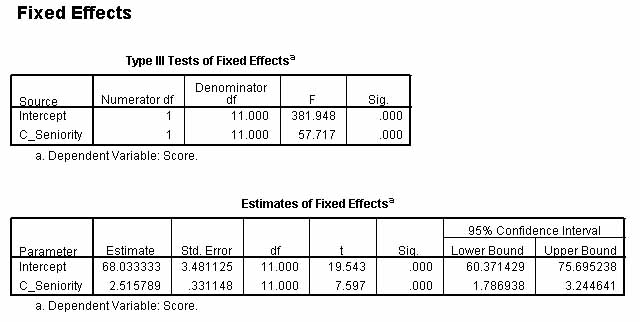
* + - 1. *Fixed effects.* The fixed effects tables also illustrated above show that SES is a significant predictor of math achievement at level 1.
      2. *Random effects*. In the "Estimates of Covariance Parameters" table below, the "Residual" row represents within-school variance in math achievement. The "Intercept[subject=id] Variance" row represents between-school variance in intercepts, where intercepts represent mean math achievement. The "ses[subject=id] Variance" row represents between-school variance in slopes, where slopes represent the strength of the relationships between SES and math achievement.



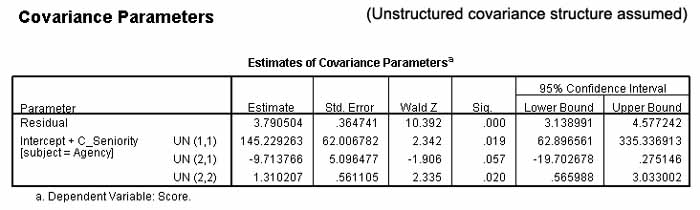
* + - * + *Partition of variance components*. If a variance components model has been assumed, as for this example, then the total variance is the sum of the three covariance parameter estimates: 36.82 + 4.85 + .42 = 42.10. Of this total, most of the variation in math scores is still attributable to the variance of student scores within schools (36.82/42.10 = 87%). Another 11.5% (=4.85/42.10) is attributable to between-school effects - that is, to differences in intercepts (mean achievement scores) between schools, controlling for level 1 SES. This compares with a between-schools effect of about 18% in the uncontrolled null model. About 1% (=.424/42.10) of the variation in math achievement is attributable to between-school differences in slopes (representing the strength of the relation of SES to math scores), controlling for level 1 SES.
        + *R2 estimate.* The variance of math achievement within schools after SES is controlled is 36.82 in this example, shown in the Residual row. In the null model it was 39.14. The difference is about 2.32. Adding individual level SES to the model thus reduces within-schools variation of math achievement by 2.32/39.14 = .06 = 6%. This value (.06) is an estimate of R2 for the random coefficients model with level 1 SES as the only predictor. We may say that 6% of the within-schools variance in math scores in the null model is attributable to between-school effects when SES as a level 1 predictor is controlled.
        + *Significance of the intercept*. That the intercept[subject=id] term above is significant in the Covariance Parameters table means that there is significant variance in intercepts across groups (schools) even after controlling student-level SES. In other words, additional level 1 predictors may be needed.
  + **An unstructured random coefficients regression model using SPSS software.**
    - *Overview*. This section presents an example of random coefficients regression, highlighting the difference between assuming a "variance components" (VC) covariance structure type versus an "unstructured" (UN) type. Let agencies be the sampling unit used as a random effect in a study of performance ratings as impacted by seniority. There are no other level 2 predictors. The random coefficients regression model runs a regression for each agency, with each regression reflecting the relationship of seniority (a level 1 variable) to score (the level 1 dependent) within that agency. Thus the dependent variable is performance rating predicted at level 1 by seniority. The intercept of score is modeled as a random effect of agency as a grouping variable, and likewise the slope of seniority is a random effect of agency.
    - *"Variance Components" (VC) vs. "Unstructured" (UN) covariance structure*. The VC assumption assumes variance components are independent, which means their random effect terms are uncorrelated. Unstructured is preferred when the researcher determines this assumption is unwarranted or if the researcher simply does not know. As the latter is often the case in random coefficients models, Unstructured is a common assumption even though it implements a less parsimonious model.
    - *SPSS*. In SPSS, select Analyze, Mixed Models, Linear; in the Specify Subjects and Repeated dialog, enter Agency as the Subjects variable. In the next "Linear Mixed Models" dialog, enter Score as the dependent variable and enter Seniority (in this example, C\_Seniority is centered Seniority) as the covariate. Click the Fixed button and set Model to C\_Seniority, making sure the Include Intercept checkbox is checked. This means Seniority is a level 1 predictor of Score. Click the Random button. setting Model to C\_Seniority and make Agency the Subject Grouping in the Combinations area. Note Agency is not entered as a random effect factor as that is already assumed by making it a Subject Grouping variable. Set the Covariance Type to Unstructured and check Include Intercept. Click the Statistics button to select output options such as Parameter Estimates and Tests for Covariance Parameters and Covariances of Random Effects; click OK.
      1. *Explanation*: Note the following regarding SPSS handling of the random coefficients regression model assuming an unstructured covariance type:
         * The level 1 intercept of performance score is modeled as a level 1 effect of employee seniority, as in ordinary regression, by virtue of listing C\_Seniority as a fixed effect.
         * The level 1 intercept of performance score also is modeled as a random effect of agency id by having Agency as the Subjects and random effects grouping combinations variable.
         * The level 1 slope of employee-level seniority is modeled as a random effect of agency id by declaring C\_Seniority to be a random effect to be modeled, under the Random button dialog.
         * The fixed effects output will have two coefficients for this example, reflecting the effects of agency id as a grouping variable on the two level 1 parameters (intercept and slope of employee-level seniority).
         * The random effects output will have three coefficients for this example, reflecting the four variance components: (1) the Residual, reflecting within-agency random effects on performance score after controlling for employee seniority; (2 - 4) three more components as follows:
    - *Information Criteria table* gives the -2 restricted log likelihood (a.k.a. model chi-square or deviance), plus the AIC, AICC, CAIC, and BIC goodness of fit measures as before, used when comparing models with lower being better. For instance, for the example data, fit is better (information criteria are lower) for the model run under the assumption of an unstructured covariance structure than a variance components covariance structure, as will be discussed below. Model chi-square can be used to test if the given model is significant overall (that is, significantly different from the model chi-square reported for the null model).



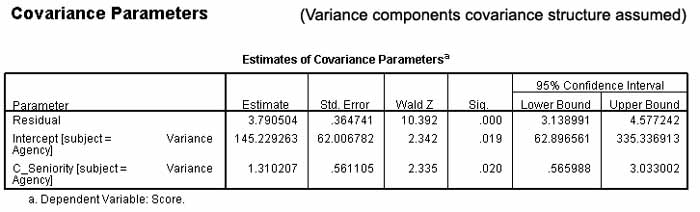
* + - *Tests of Fixed Effects table* gives the significance level for the intercept and for C\_Seniority, the only fixed effect in the model in this example apart from the intercept. If C\_Seniority is significant, as it is here, the researcher concludes seniority does impact score. (Note df (degrees of freedom) is 11 since with 12 agencies, df = number of groups minus 1).



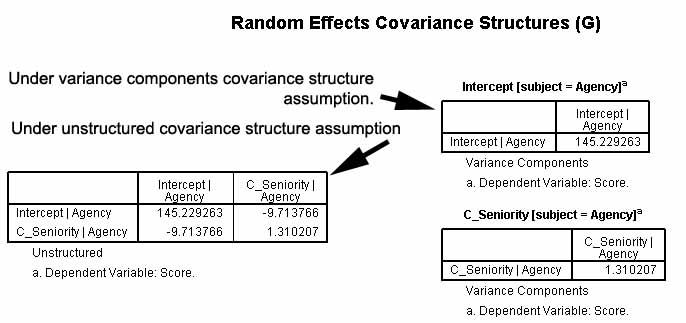
* + - *Estimates of Fixed Effects table*, illustrated above, gives the coefficients for the intercept and for C\_Seniority, the only fixed effect, along with the standard errors, significance, and confidence intervals. The coefficients computed in this table are the same whether an unstructured or a variance components covariance structure is assumed.
      1. *Intercept.* The coefficient for the intercept is the average Agency mean (here a Score of 68) on performance Score when seniority is held constant at 0. If C\_Seniority is centered, then the phrasing becomes "is held constant at its mean," which may be a more meaningful statement.
      2. *Slope.* The coefficient for C\_Seniority is the average slope across all agencies (recall RC models run a regression for each of the agencies) and is interpreted as in regression: for each year increase in C\_Seniority, Score increases by 2.5 units, on the average. Since agency is the level 2 grouping variable and since C\_Seniority and Score are level 1 variables, if the estimate of the slope for C\_Seniority is significant, as it is here, then performance Score and C\_Seniority are related within Agencies. In a model where additional individual or group-level fixed factors and their interactions had been entered in the model under the Fixed button, this table would show the significance of each fixed effect and each interaction term. Thhe researcher would drop non-significant terms from the model one at a time, re-running the RC model without them.
    - *Estimates of Covariance Parameters table*. This table provides the covariance parameter estimates for the random effects, better known as the variance components. In this example these are the within-agency Residual effect and the three between-agency UN() estimates explained below. The Wald statistic and its significance level is also printed for each variance component.



* + - 1. *Residual.* Above, the Residual is interpreted as the within-agencies variance in performance Score controlling for C\_Seniority. The magnitude of the C\_Seniority effect could be estimated by how much the residual covariance estimate was reduced compared that in the null model without C\_Seniority as a covariate. Since the residual is unexplained variance, if C\_Seniority has a large effect on the variability of performance scores within Agencies, the residual variance should drop appreciably in the RC model compared to the null model without C\_Seniority at level 1 but with Agency as a random grouping effect and no other level 2 predictors.
      2. *UN(1,1).* The estimated between-agency variability of intercepts. The higher this value, the more variability among agencies impacts employee performance scores. That is, the higher the intercept variance (the higher UN(1,1)), the more the effect of group-level variables on the dependent - in this example, the more the effect of Agency on Score. This value is identical to the Intercept [subject = agency] effect for the VC model described below.
      3. *UN(2,2).* This is the estimated between-agency variability of slopes. The higher its value, the more variability among agencies impacts how strongly seniority is related to performance scores (that is, how strongly slopes are affected). If significant, the researcher concludes that the variability is not 0 and agencies indeed differ in the slopes, which reflect the strength of the relationship of C\_Seniority and Score. The larger the UN(2,2) estimate, the more the agencies differ in how much C\_Seniority affects Score. This value is identical to the C\_Seniority[subject = agency] effect for the VC model described below.
      4. *UN(2,1).*The estimated covariance of Agency slopes and intercepts. If non-significant, the covariance of the intercepts (plural because the regression is done for each Agency, with intercepts reflecting mean Scores) and slopes cannot be said to differ from 0. That is, agencies with higher mean Scores do not tend to be ones with stronger (or weaker) C\_Seniority-Score relationships on a within-Agency basis. This value has no counterpart for variance components models because VC models assume random effects (in this example, Intercepts and Seniority are random effects) are independent of each other, thus having zero covariance of slopes and intercepts. In this example, UN(2,1) is not significant under the assumption of an unstructured covariance structure, thus warranting use of a VC model. Accordingly, when the Variance Components covariance type is selected under the Random button estimates are presented for three effects in the manner illustrated below:
    - *The corresponding VC model.* In a VC model, variance components are uncorrelated and additive, enabling the percentage calculations below:

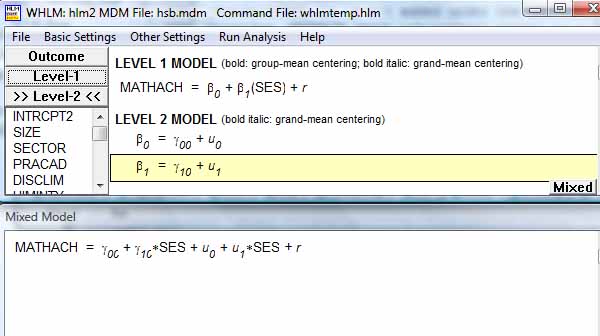


* + - 1. Residual: Same as unstructured model above, for Residual. The estimate of 3.79 is 2.5% of the total of all estimates, which was 150.23. The within-agency effect of C\_Seniority explains 2.5% of the variance in Score. That is, only 2.5% of the variance in Score remains after C\_Seniority and Agency are controlled. This is the within-agency effect.
      2. Intercept [subject = Agency]. Same as unstructured model above, for UN(1,1). The estimate of 145.23 is 96.67% of the total variance, which is 150.23. The between-agency effect of C\_Seniority on Score accounts for over 96% of the variance in Score. That is, differences in Score are largely accounted for by differences in Agency. (For the artificial data in the example, agencies varied widely in average Score). This is the between-agency effect.
      3. Seniority [subject = Agency]. Same as unstructured model above, for UN(2,2). Between-agency differences in how much C\_ Seniority affects Score accounts for only 1% of the total variance. That is, the relation of C\_Seniority to Score is similar across agencies, even if the mean Score differs substantially by Agency. This is the C\_Seniority effect.
      4. There is no coefficient corresponding to UN(2,1) in the unstructured model since in a VC model, components are constrained to have zero covariance.
    - *Random Effects Covariance Structure (G) table*, part of optional output, shows the same information as the Estimates of Covariance Parameters table but in a different format. For this example it is a 2x2 table for which the rows and columens are illustrated below for both the unstructured model and the variance components model. The "Intercept|Agency" row by "Intercept|Agency" column cell displays the variability of Agency intercepts (what above was labeled UN(1,1)). The "C\_Seniority|Agency" row by "C\_Seniority|Agency" column displays the variability of agency slopes (what above was labeled UN(2,2)). The other two cells (the off-diagonal) both display the slope-intercept covariances, labeled UN(2,1) in the Estimates of Covariance Parameters table. Similar information is presented for the VC model.



* + - *Is an RC model really needed?*. If the variability of the intercept (labeled UN(1,1) in unstructured covariance models, illustrated [above](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#mixed51); and labeled "Intercept(Subject=Agency" in variance components models [above](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#mixed52)) is an insignificant percentage of the total of variance components, then the same conclusions will be reached without modeling the level 1 slope as a random effect of the level 2 grouping variable. To reach this conclusion, however, the researcher must also take account of whether variability of the slopes and of the interaction of slopes and intercepts is non-significant (that is, UN(2,2) and UN(2,1) respectively are non-significant) also.
      1. *RC models and ICC*. This is similar to the ICC test of LMM appropriateness discussed [above](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#icctest) except ICC was based on a null model in which only the intercept of the dependent was modeled. The assessment of RC appropriateness involves a model with predictors in which the slopes and interaction of slopes and intercepts are modeled. If slopes and interactions are non-significant, this suggests that it is not necessary to model the dependent variable for within-group effects using an RC model. That is, this means that in the regressions for each agency conducted in the random coefficients regression model, the slopes are not significantly different. A significant ICC means either a random intercepts or random coeffients model is justified; a non-significant ICC does not suffice to reject the appropriateness of an RC model.
      2. *Likelihood ratio test*. Likelihood ratio tests are a better method of assessing RC appropriateness (better than the Wald significance tests from the covariance parameters (vaiance components) tables just discussed). To perform this test for the example above, the researcher takes the -2LL from the "Information Criteria" table for the random coefficients regression model, then re-runs the model removing C\_seniority from the model under the Random button dialog (thus leaving no random effects to model apart from the grouping variable in the Combinations area). To be comparable one would leave the covariance structure type as Unstructured. The degrees of freedom for the chi-square difference between the two -2LL's is 2 in this example (the removed slope for seniority as a random effect, and the removed covariance of this slope with the intercept). As above, the probability of a model chi-square difference this large or larger with 2 degrees of freedom can be looked up in a chi-square table, or can be obtained in SPSS under Transform, Compute, then entering the fomula sig.chisq2(d, df), where d is model chi-square difference and df is the degrees of freedom. If the computed probability > .05, then the RC model is not significantly different from an HLM model and on the basis of parsimony, one need not model the slope of seniority as a random effect in an RC model. However, one would still use agency as the grouping variable and retain whatever individual level and agency level covariates were in the analysis.
  + **A random coefficients regression model using HLM software.**
    - *Example.* This section describes a two-level model featuring a level 1 covariate whose slope is treated as a random effect as is the level 1 intercept. That is, the slope of the level 1 predictor and the level 1 intercept reflect estimates based on one regression for each level 2 unit (each of 160 schools). The modeling of slope as well as intercept as random effects makes this a random coefficients regression model, not just a random intercepts model. This example is similar to the previously discussed null model except that MATHACH (math achievement test score) is predicted by student SES at level 1, as illustrated below. As SES is already standardized, it is not added under the centering option which HLM provides. As in the null model, the intercept of MATHACH is modeled as a function of the school id grouping variable at level 2, but now the slope of student SES is modeled as a random coefficient as well.
    - *HLM set-up*. HLM gives the option of toggling the level 2 error terms on or off - here error terms are right-clicked to make sure they are "on" rather than in the greyed-out "off" position. In the figure below, the U1 error term for the level 2 estimator of the level 1 slope of SES (B1) is turned on. U1 is the level 2 random effect for the level 1 B1 slope, making this a random coefficients regression model. If U1 were toggled off, only the B0 intercept would be estimated by a level 2 random effect (U0), making this a random intercepts model rather than a random coefficients model.

The model: MATHACH (Y = math achievement score) is a function of the intercept (B0) plus the slope times the predictor (B1\*SES) plus a residual error term. The intercept (B0) is a function of the grand mean of MATHACH across level 2 units (schools) plus a random error term. The slope (B1) is a function of the grand mean of SES across schools plus a random error term. Again, clicking the "Mixed" button algebraically substitutes the level 2 equalities into the level 1 equation to create a single mixed equation for the model, shown at the bottom of the figure below.



* + - 1. *The model*. Click "Run analysis" from the menu to generate output for the random coefficients regression model, shown below in excerpted form. Click File, View output to view. At the top of the output is a printed version of the model, shown below.

Module: HLM2S.EXE

The data source for this run = hsb.mdm

The outcome variable is MATHACH

The model specified for the fixed effects was:

----------------------------------------------------

Level-1 Level-2

Coefficients Predictors

---------------------- ---------------

INTRCPT1, B0 INTRCPT2, G00

SES slope, B1 INTRCPT2, G10

Summary of the model specified (in equation format)

---------------------------------------------------

Level-1 Model

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

Level-2 Model

B0 = G00 + U0

B1 = G10 + U1

Iterations stopped due to small change in likelihood function

\*\*\*\*\*\*\* ITERATION 21 \*\*\*\*\*\*\*

Sigma\_squared = 36.82835

Tau

INTRCPT1,B0 4.82978 -0.15399

SES,B1 -0.15399 0.41828

...

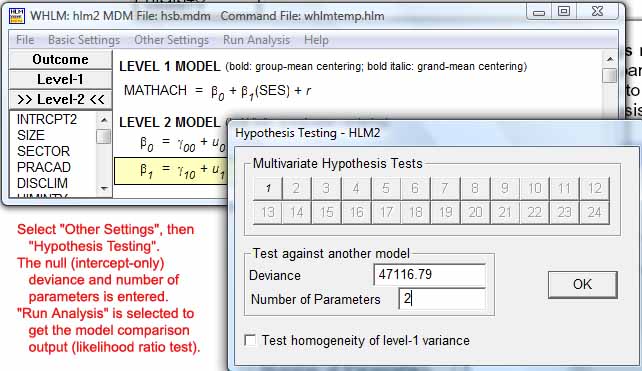
* + - * + *Models without intercepts*. Though not the case here, it may be noted that it is possible to delete the level 1 intercept (B0) from the model. This is done when there is a complete set of dummies as level 1 predictors. For instance, if there is a "MALE" variable coded 0,1 with 1=Male, and there is a "FEMALE" variable coded 0,1 with 1=Female, the model would be be overdetermined (lack positive degrees of freedom needed for solution). To prevent this, the B0 term may be deleted. In the main HLM interface, simply select the Level 1 list of variables, highlight INTRCPT1, then delete it.
        + *Sigma-squared* is the within-schools variance in MATHACH after SES is controlled, here 36.83. Note that sigma-squared is now a partial coefficient, controlling for additional predictors in the model. For the same data and model, sigma-squared corresponds to the "Residual" estimate in the "Estimates of Covariance Parameters" table in SPSS output.
        + *The tau matrix*. The tau values now appear in a matrix and are level 2 interaction effects with the level 1 outcome variable. These coefficients are also partial coefficients, reflecting residual variation after controlling for other predictors in the model.

The intercept value in the upper left of the matrix (here 4.83) is the between-schools variance estimate for the intercept, assuming no level 2 predictors. It corresponds to the "Intercept[subject=id] Variance" estimate in the "Estimates of Covariance Parameters" table of SPSS output for the same model and data. It is the school effect on mean math achievement scores, controlling for student SES.

The tau value in the lower right (here .42) is the between-schools variance estimate for the slopes of SES, assuming no level 2 predictors. Note "slopes" is plural because separate regressions are computed for each level 2 unit, school. This tau value corresponds to the "ses[subject=id] Variance" estimate in the "Estimates of Covariance Parameters" table of SPSS output for the same model and data. It is the school effect on the strength of relationship (slope of student SES as a predictor of MATHACH.

The off-diagonal elements in the tau matrix represent the covariance of the error terms for the level 2 estimates of the level 1 slopes and intercepts (-.15). That the covariance value is negative means that as mean math achievement (intercepts) go down among the 160 schools, the relationship of individual-level SES and math achievement (slopes) becomes more important (stronger, goes up).

* + - 1. *Intraclass correlation (ICC)*. In contrast to the ICC computed for the null model, for a random coefficients regression model the computation is different and the ICC becomes conditional.
         * *Computation*. With two random effects (here the intercept and slope of centered ses), the total of variance components is the within-group variance (sigma-squared = 36.82835; this is the level 1 variance term) plus the level 2 between-group variance terms: 4.82978 (the level 2 intercept variance estimate) + 0.41828 (the level 2 slope variance estimate) plus twice the covariance: 2\*-0.15399. For these data, the total variance sums to 41.76843. For a variance components model, constrained to have zero covariance of the random effect error terms, there would be no 2\*cov term in the ICC formula. For other models the covariance must be taken into account because the variance components are no longer additive in a simple way. The ICC is the between-groups effect on the intercept of the outcome variable (4.82978) divided by total variance (41.76843) = .12.
         * *Conditional interpretation*. We may say that 12% of the variance in math achievement is attributable to the between-groups effect, controlling for student SES. The ICC is conditional because it is now a partial coefficient and its interpretation depends on controlling for other variables in the model.
      2. *R2 estimate.* Sigma-squared, the variance of MATHACH within schools after SES is controlled, is 36.83. In the intercept-only model it was 39.15. The difference is 2.32. Adding individual level (level 1) SES to the model thus reduces within-schools variance of MATHACH by 2.32/39.15 = .06. This value is an estimate of R2 for the random coefficients regression model with SES (a level 1 variable) as the only predictor.
      3. *Reliability.* The intercept is much more reliable than the slope of SES, reflecting the fact that SES is not a powerful predictor.
      4. ----------------------------------------------------
      5. Random level-1 coefficient Reliability estimate
      6. ----------------------------------------------------
      7. INTRCPT1, B0 0.797
      8. SES, B1 0.179
      9. ----------------------------------------------------
      10. *Fixed effects*. The fixed effects table below shows the average intercept and slope across the 160 schools. Here, the predictor is SES and thus the coefficient is a slope. Had the predictor been categorical, the coefficient could be interpreted as differences between means. Either way this a partial coefficient, controlling for variation in students within schools. That standard errors above are very similar whether calculated by the ordinary formula or the "robust" method, indicates that the default assumption that the dependent variable is distributed normally is acceptable. By either calculation, both the intercept and the slope of SES are found to be significantly different from 0 in this model.
      11. Final estimation of fixed effects:
      12. ----------------------------------------------------------------------------
      13. Standard Approx.
      14. Fixed Effect Coefficient Error T-ratio d.f. P-value
      15. ----------------------------------------------------------------------------
      16. For INTRCPT1, B0
      17. INTRCPT2, G00 12.664935 0.189874 66.702 159 0.000
      18. For SES slope, B1
      19. INTRCPT2, G10 2.393878 0.118278 20.240 159 0.000
      20. ----------------------------------------------------------------------------
      21. Final estimation of fixed effects
      22. (with robust standard errors)
      23. ----------------------------------------------------------------------------
      24. Standard Approx.
      25. Fixed Effect Coefficient Error T-ratio d.f. P-value
      26. ----------------------------------------------------------------------------
      27. For INTRCPT1, B0
      28. INTRCPT2, G00 12.664935 0.189251 66.921 159 0.000
      29. For SES slope, B1
      30. INTRCPT2, G10 2.393878 0.117697 20.339 159 0.000
      31. ----------------------------------------------------------------------------
      32. *Random effects*. As both the intercept and slope variance components below are significant, we reject the null hypotheses that slopes and intercepts have no difference between schools. The confidence limits on these components are not printed but are simply plus or minus 1.96 standard deviations around their respective coefficients.
      33. Final estimation of variance components:
      34. -----------------------------------------------------------------------------
      35. Random Effect Standard Variance df Chi-square P-value
      36. Deviation Component
      37. -----------------------------------------------------------------------------
      38. INTRCPT1, U0 2.19768 4.82978 159 905.26472 0.000
      39. SES slope, U1 0.64675 0.41828 159 216.21178 0.002
      40. level-1, R 6.06864 36.82835
      41. -----------------------------------------------------------------------------
      42. *Deviance.* Below we see deviance has dropped somewhat, from 47116.79 in the intercept-only model to 46638.56 in the present random coefficients regression model with SES as a level 1 predictor. This is a difference of 478.23. The intercept-only model had 2 parameters, whereas the present model has 4 - a difference of 2. The critical value of chi-square for 2 degrees of freedom is only 13.82 (read from a chi-square table), far lower than the difference, so we may say the difference between models is significant at ≤ .001, meaning that adding student SES as a predictor improves the model significantly. (Deviance corresponds to the "-2 Restricted Log Likelihood" estimate in the "Information Criteria" table of SPSS output for the same model and data.)
      43. Statistics for current covariance components model
      44. --------------------------------------------------
      45. Deviance = 46638.560929
      46. Number of estimated parameters = 4
      47. -----------------------------------------------------------------------------
      48. *Likelihood ratio test of two models.* HLM provides an automated way of obtaining the likelihood ratio test, though it is available only for HLM2, HLM3, and HGLM modules in HLM software, where Laplace estimation is used. To obtain this test, the researcher must previously have run the intercept-only model and saved its deviance and number of parameters values. Then the researcher selects Other Settings, Hypothesis Testing, and enters these values, as illustrated below. Finally, the researcher selects "Run Analysis".



Having chosen the hypothesis testing option causes the likelihood rato test (corresponding to the p-value of the chi-square statistic in the output below) to be printed in the Deviance section of HLM output (bottom of the output). Again, the likelihood ratio test shows the random coefficients regression model with SES as a predictor to be different from the intercept-only model at a significance level ≤ .001:

Statistics for current covariance components model

--------------------------------------------------

Deviance = 46638.560929

Number of estimated parameters = 4

Variance-Covariance components test

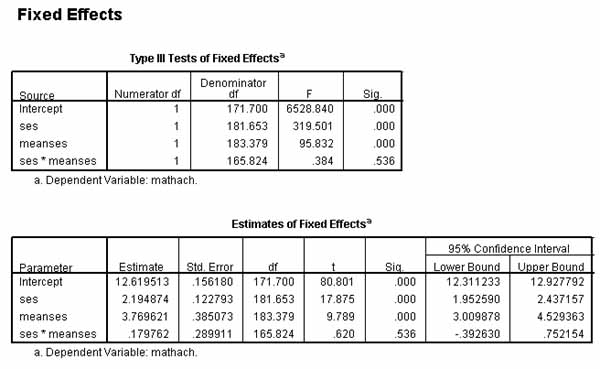
-----------------------------------

Chi-square statistic = 478.22907

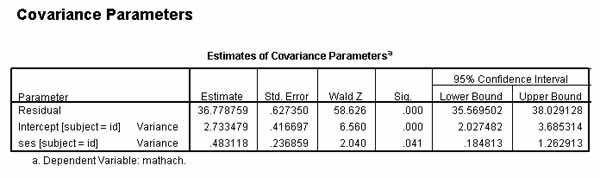
Number of degrees of freedom = 2

P-value = 0.000

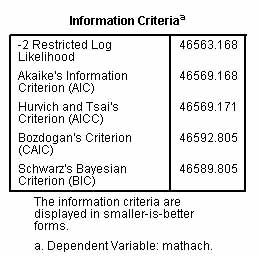
* + **A two-level intercepts-and-slopes-as-outcomes model (full random coefficients model), using SPSS.**
    - *Overview*. A full random coefficients model, also called an intercepts-and-slopes-as-outcomes model, is an extension of a random coefficients regression model to include level 2 predictors in addition to the grouping variable and any level 1 predictors. Thus the level 1 slope(s) as well as the level 1 intercept is predicted on the basis of the level 2 grouping variable as a random effect and also by one or more other level 2 random effect predictors.
    - *Example.* Differences in mean math scores (reflected in the level 1 intercept) may be analyzed as predicted by the level 1 covariate student socioeconomic status (SES), predicting the slope of SES as well as the level 1 intercept as a function of the between-groups effect of the grouping variable school at level 2 and by the level 2 random effect of Meanses (a centered variable representing whether a school is below, near, or above average in mean socioeconomic status of its students). Such full random coefficients regression models are conditional at both levels 1 and 2.
    - *SPSS.* Select Analyze, Mixed Models, Linear. On the initial dialog, let Subjects=id, where id is the level 2 school id. On the next dialog page, set the dependent to be mathach and set ses (socioeconomic status at the student level; this variable is pre-centered in the data) as a covariate. Also set meanses (coded -1/0/+1 for whether a school is below, near, or above average in mean socioeconomic status) as a covariate. Click the fixed effects button and model ses, meaning ses is a level 1 predictor of mathach. Also enter as fixed effects both M\meanses and meanses\*ses. Click the Random button and move id to the Combinations area, meaning it is the grouping variable. Also under the Random button dialog, model SES as a random effect, meaning the slope of the level 1 variable SES will be modeled as a random effect of the level 2 grouping variable (school id), in which a separate regression will be computed for each school. Check "Include Intercept" (not the default). Select "Variance Components" as the model assumption (this is the default). Continue. Under Statistics, select Parameter Estimates and Tests for Covariance Parameters. Continue. OK. Note SES is thus modeled both as a fixed and a random effect.
      1. *Explanation*. The SPSS interface is less than intuitive, but note the following:
         * The level 1 intercept is modeled as a random effect by having school id as the Subject/Combinations grouping variable.
         * The level 1 slope of ses is modeled as a random effect by having school id as the grouping variable and declaring ses to be a random effect to be modeled, under the Random button dialog.
         * The level 1 intercept is also modeled as an effect of level 2 meanses by including meanses as a fixed effect.
         * The level 1 slope of ses is also modeled as an effect of level 2 meanses by including the meanses\*ses interaction term as a fixed effect.
         * Thus both the level 1 intercept and the level 1 slope are modeled as effects of both the level 2 grouping variable (school) and the level 2 covariate (meanses).
         * The fixed effects output will have four coefficients for this example: student-level ses at level 1, school-level meanses at level 2, the ses\*meanses cross-level interaction, and the level 1 intercept for mathach.
         * The random effects output will have three coefficients for this example, reflecting the three variance components: (1) the Residual, reflecting within-school random effects; (2) the ses component, reflecting between-school random effects on the slope of ses; and (3) the Intercept component, reflecting between-school random effects on the level 1 intercept.
         * Managing a full random coefficients model with multiple predictors at each level can become a daunting logical puzzle in SPSS. The HLM 7 software interface, presented in other examples in this section, is far more intuitive and thus much less prone to human error when constructing complex full random coefficients models.
    - *Fixed effects*. In the table below, there are four fixed effects:
      * + Intercept: The mean math score is 12.62 when student ses and school meanses are 0, as in OLS regression.
        + SES: When student-level ses increases by 1 unit, math score is predicted to increase by 2.19 units, controlling for other variables in the model. That is, this is the predicted slope of ses when meanses is 0 (which means at-average as coded in the current example).
        + Meanses: When meanses increases by 1 unit (recall meanses is coded -1, 0, 1), math score is predicted to increase by 3.77 units, controlling for other variables in the model (that is, when ses is 0, which, if centered as it is in this example, signifies at its mean).
        + SES\*Meanses: This interaction effect is the effect of the level 2 variable meanses on the slope of SES at level 1. As meanses increases by 1 unit, the slope of SES is predicted to increase by .18. However, this is not significant in this example.



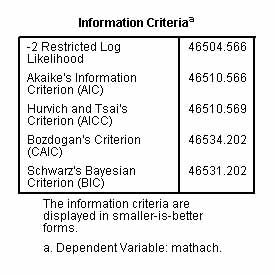
* + - *Random effects / variance components*. In the table below, there are three fixed effects:
      1. Residual: This is the largest variance component and represents within-school variance in math scores controlling for other variables in the model. As other variables explain more, this partial coefficient will decline. That it is high in this example shows the need for additional explanatory variables.
      2. Intercept[subject=id] Variance: This much smaller variance component represents the between-school effect on the level 1 intercept, which represents mean math score.
      3. ses[subject=id] Variance: This small and marginally significant variance component represents the between-school effect on the slope of SES. For these data, modeling SES as a random effect of level 2 makes little difference and a more parsimonious model might be judged better.



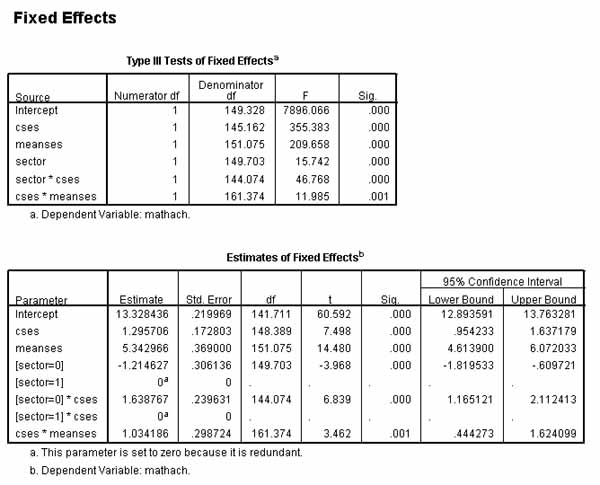
* + - *Model fit*. The model fit measures in the table below may be used for purposes of model comparison, with lower coefficients representing better fit.
      1. *Likelihood ratio test*. In the table below, the deviance (-2LL) is 46563.17, with 7 parameters (number of parameters is displayed in the "Model Dimension" table, not shown). In the random coefficients regression model discussed previously, the deviance was 46640.66, with 5 parameters. This is a deviance (model chi-square) difference of 77.43 and a degrees of freedom difference of 2 parameters. In a chi-square table with 2 degrees of freedom, the critical value is 5.99. As 77.43 is far larger, we can say that the full random coefficients model is significantly better fit than the previous random coefficients regression model. This method of comparison is called the "likelihood ratio test" or the "model chi-square difference test."
      2. *Model development.* This model might be compared to yet other models, such as one dropping the interaction term (ses\*meanses) as a fixed effect. Dropping it would mean that the level 1 slope would still be a function of the level 2 grouping variable, school, but would no longer also estimated by the level 2 covariate, meanses. The deviance for this modified model, which has 6 parameters (one less, having dropped the interaction term), is 46562.91. Comparing the two models, the chi-square difference is .25 and df = 1. The critical value of chi-square of 1 degree of freedom is 3.84, much larger than .25. We conclude that there is no significant difference in model fit and therefore choose the modified model as the better one on parsimony grounds.



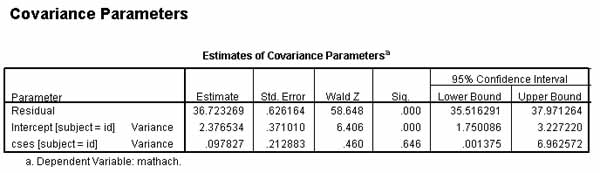
* + **A two-level intercepts-and-slopes-as-outcomes model (full random coefficients model) with a dichotomous predictor at level 2, using SPSS.**
    - *Example.* Where student is level 1 and school id is level 2, This example adds a level 2 predictor, the dichotomous variable "sector," (coded 0=public, 1=parochial), to the model predicting math achievement (mathach) at level1 from student socioeconomic status, renamed "cses" to indicate that it is centered. In this model, cses is a fixed effect as it would be in OLS regression. The level 2 variables, meanses (school average ses) and sector, are fixed effects also: they are not random effects of some higher level. The cross-level interaction effects, meanses\*cses and sector\*cses, must also be added as fixed effects as they are used to estimate the effects of meanses and sector on the slope of cses at level 1. In general, in SPSS and SAS syntax, such cross-level interactions are used whenever level 1 slopes are to be modeled as random effects.
    - *SPSS.* In SPSS, select Analyze, Mixed Models, Linear; in the Specify Subjects and Repeated dialog, enter id (which is school id) as the Subjects variable; in the next "Linear Mixed Models" dialog, enter mathach as the dependent variable; enter cses, meanses, and sector as covariates; click the Fixed button and set Model to cses, meanses, sector, meanses\*cses, and sector\*cses, making sure the Include Intercept checkbox is also checked. Then click the Random button and make id the Subject Grouping in the Combinations area, with the Covariance Type set to the default Variance Components type and making sure Include Intercept is checked. Also under the Random button, model cses as a random effect. Click the Statistics button to select output options such as Parameter Estimates and Tests for Covariance Parameters. Click OK.
    - *Information Criteria table*. This table displays goodness of fit measures, where lower is better fit. For these data, fit is better than the null model or previous models. This could be tested using the previously-described likelihood ratio test.



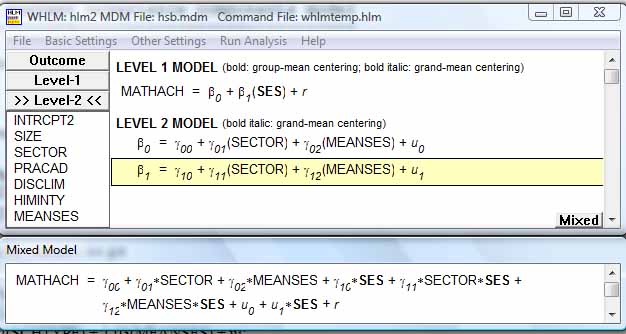
* + - *Tests and Estimates of Fixed Effects tables*. These tables show all fixed effects are significant. In the "Estimates of Fixed Effects table" below, there are two sets of coefficients, though SPSS does not clearly list them in sets.
      1. *Coefficients pertaining to the level 1 intercept of math achievement predicted by student's centered ses*. The higher the intercept, the higher the mean math score. The "Intercept" means that average math score is 13.33 when predictors are 0. The "meanses" means that this intercept increases by 5.34 units when meanses increases 1 unit (recall meanses is coded -1, 0, 1 and that one regression is run for each school in a random coefficients model), controlling for other variables in the model. The "[sector=0]" value means that public schools (sector=0) compared to parochial schools (sector = 1) lowers the intercept (representing mean math scores) by 1.21, controlling for other variables in the model. That is, parochial schools have higher mean math scores.
      2. *Coefficients pertaining to the level 1 slope of cses as a predictor of math achievement*. The higher the slope, the stronger the relation of cses to math score. The "cses" value is the estimated mean slope of cses across all schools when sector and meanses are 0, and under these conditions, when cses increases 1 unit, math score increases by 1.3 units. The "cses\*meanses" value means that when meanses increases by 1 unit, the slope of cses increases by 1.03 units, controlling for other variables in the model. The "[sector=0]\*cses" value means that public schools (sector = 0) compared to parochial schools (sector = 1) increase the slope of cses by 1.64 units, controlling for other variables in the model. That is, the relation of cses to math score is stronger in public schools.



* + - *Estimates of Covariance Parameters table*. This table gives the covariance parameter estimates for Residual and for Intercepts within Subject=id, along with a Wald test of significance.
      1. Residual. If significant, this means that math achievement scores vary significantly *within* schools, controlling for other variables in the model. That is, this is the within-subjects effect, where school id (not student id!) is the subject variable. For these data, almost 94% of the variance in math achievement scores is explained by variance among students within schools, assuming a variance components model. That residual within-groups variance is large indicates the need for additional predictors.
      2. Intercept [subject=id] Variance. Recall id is school id, not student id. If this intercept term is significant, this means that intercepts of predicted math achievement score vary significantly between schools. That is, this is the between-subjects effect, where school id is the subject variable. We may say there is a significant between-school effect on mean math scores. In a variance components model, its proportion of the total of parameter estimates is the percent of variance in math achievement scores attributable to differences between schools after other predictors in the model are controlled. The between-group variance component divided by the total of variance components is .06. We may say 6% of the total variance in math achievement is due to the between-schools effect on mean math scores (on the level 1 intercepts).
      3. Cses[subject=id] Variance. This coefficient reflects the between-schools effect on the slope of the level 1 predictor, cses. For these data, this coefficient is not significant, indicating that the slopes of cses do not vary significantly when, under a random coefficients model, separate regressions are run for each school. Put another way, we may say there is no variance between schools in the slope of cses. Put a third way, the strength of relationship between cses and math achievement does not vary significantly between schools.

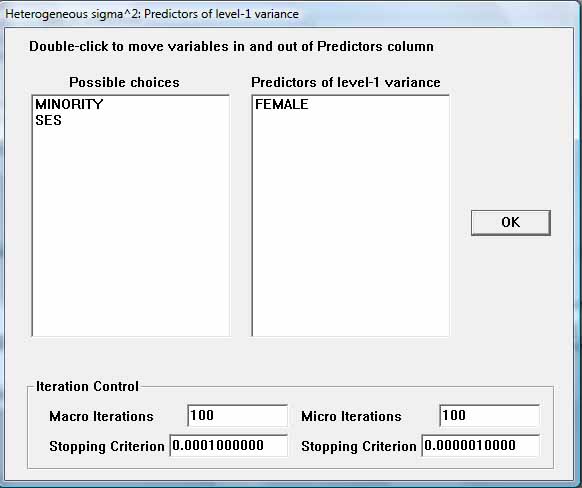


* + **A two-level random coefficients intercepts-and-slopes-as-outcomes model, using HLM software.**
    - *Example*. In this model there is a level 1 predictor, student-level SES. Both its slope and the level 1 intercept of MATHACH are modeled as random effects of the level 2 grouping variable, school, and of level 2 covariates. The two level-2 predictors are SECTOR and MEANSES. SECTOR is coded 0 = public school, 1 = parochial school. MEANSES is a trichotomized measure of whether a school is low, medium, or high on the mean SES of its students. In the figure below, boldface indicates a variable (here, SES) which is centered (in this case, group centered, meaning MEANSES is subtracted from each SES). The centering option is offered by HLM as each variable is added to the model.

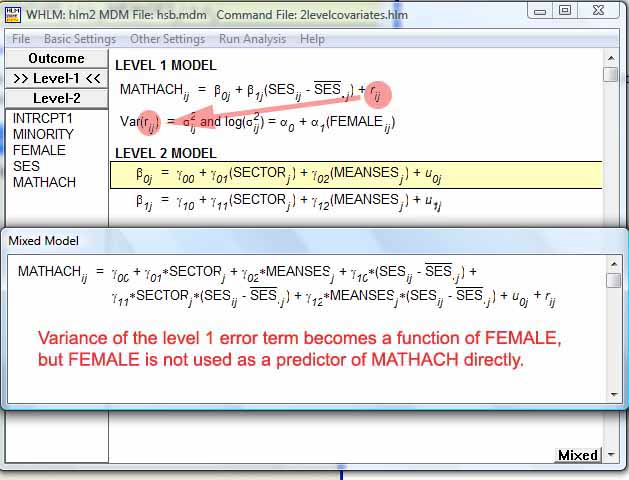


* + - 1. *The model.* Click "Run analysis" from the menu to run the intercepts-and-slopes-as-outcomes model. For this run, the default reduced output was accepted. At the top of the output, the model is reproduced in printed format as shown below. In this example, MATHACH (Y) is a function of the intercept (B0) plus the B1 coefficient times (SES - MEANSES) plus a residual error term. The B0 intercept is a function of a level-2 intercept, a level-2 slope (G01) times the level 2 variable SECTOR, plus another level-2 slope (G02) times the level 2 variable MEANSES plus an error term (U00. The level 1 B1 slope is a function of similar level 2 intercepts, slopes, and variables. In summary, the level 1 slopes and intercept are modeled by schools as the level 2 grouping variable and by level 2 covariates.
      2. Module: HLM2S.EXE
      3. The data source for this run = hsb.mdm
      4. The outcome variable is MATHACH
      5. The model specified for the fixed effects was:
      6. ----------------------------------------------------
      7. Level-1 Level-2
      8. Coefficients Predictors
      9. ---------------------- ---------------
      10. INTRCPT1, B0 INTRCPT2, G00
      11. SECTOR, G01
      12. MEANSES, G02
      13. \* SES slope, B1 INTRCPT2, G10
      14. SECTOR, G11
      15. MEANSES, G12
      16. '\*' - This level-1 predictor has been centered around its group mean.
      17. The model specified for the covariance components was:
      18. ---------------------------------------------------------
      19. Sigma squared (constant across level-2 units)
      20. Tau dimensions
      21. INTRCPT1
      22. SES slope
      23. Summary of the model specified (in equation format)
      24. ---------------------------------------------------
      25. Level-1 Model
      26. Y = B0 + B1\*(SES) + R
      27. Level-2 Model
      28. B0 = G00 + G01\*(SECTOR) + G02\*(MEANSES) + U0
      29. B1 = G10 + G11\*(SECTOR) + G12\*(MEANSES) + U1
      30. *Sigma-squared and tau.* The tau values appear in a variance-covariance matrix, where the covariance of the SES intercept and slope error terms for the 160 schools in this example is .19 and the corresponding correlation is .32. On the main diagonal, the variance estimate for the intercepts is 2.37 and for the slopes is .15. Sigma-squared and tau values here are identical to those in the variance components table discussed below with regard to random effects.
      31. Sigma\_squared = 36.70313
      32. Tau
      33. INTRCPT1,B0 2.37996 0.19058
      34. SES,B1 0.19058 0.14892
      35. Tau (as correlations)
      36. INTRCPT1,B0 1.000 0.320
      37. SES,B1 0.320 1.000
      38. *Reliability.* The reliability is high for the intercept but low for the slope. When one regression is run for each school, there is more variation in slopes than in intercepts. HLM compared to OLS will be more different for slope estimates than for intercept estimates.
      39. ----------------------------------------------------
      40. Random level-1 coefficient Reliability estimate
      41. ----------------------------------------------------
      42. INTRCPT1, B0 0.733
      43. SES, B1 0.073
      44. ----------------------------------------------------
      45. *Fixed effects*. The intercept-and-slopes-as-outcomes model used the level 2 variables SECTOR and MEANSES to predict the level-1 intercept (B0) and the level 1 slope of group-centered SES (B1). In the table below we see that SECTOR and MEANSES are both significant predictors of both B0 and B1. Recall SECTOR is the public vs. parochial school binary variable, and MEANSES is a trichotomized measure of whether a school is low, medium, or high on average SES of its students. The significant p values for SECTOR mean that public and parochial schools differ significantly on both intercepts (average math achievement controlling for other variables in the model) and slopes (strength of the relation of SES to math achievement controlling for other variables in the model). Likewise, the significant p values for MEANSES indicate that schools with different level 2 SES averages also differ in these same ways. In the "Final Estimation of Fixed Effects" table above, there are two sets of coefficients:
          * *Coefficients pertaining to the level 1 intercept of math achievement predicted by student's centered ses*. The higher the intercept, the higher the mean math score. The "INTRCPT2, G00" coefficient (corresponding to "Intercept" in SPSS output) means that average math score is 12.10 when predictors are 0. The "SECTOR, G01" coefficient (corresponding to "[sector=0]" in SPSS) means that when sector increases by one unit, going from public schools (sector=0) to parochial schools (sector = 1), the intercept(representing mean math scores) increases by 1.23, controlling for other variables in the model. That is, parochial schools have higher mean math scores. The "MEANSES, G02" coefficient (corresponding to "meanses" in SPSS) means that this intercept increases by 5.33 units when meanses increases 1 unit (recall meanses is coded -1, 0, 1 and recall that one regression is run for each school in a random coefficients model), controlling for other variables in the model. That is, the MEANSES coefficient of 5.33 indicates that the greater the mean SES of a school, the greater that school's mean math achievement score.
          * *Coefficients pertaining to the level 1 slope of centered ses as a predictor of math achievement*. The higher the slope, the stronger the relation of cses to math score. The INTRCPT2, G10" coefficient (corresponding to the "cses" value in SPSS output) is the estimated mean slope of cses across all schools when sector and meanses are 0, and under these conditions, when centered ses increases 1 unit, math score increases by 2.94 units. The "SECTOR. G11" coefficient (corresponding to the "[sector=0]\*cses" value in SPSS) means that as sector increases 1 unit (going from public schools (sector = 0) to parochial schools (sector = 1)), the slope of centered ses by decreases by 1.64 units, controlling for other variables in the model. That is, the relation of cses to math score is stronger in public schools. The MEANSES, G12" coefficient (corresponding to the "cses\*meanses" value in SPSS output) means that when meanses increases by 1 unit, the slope of cses increases by 1.03 units, controlling for other variables in the model. That is, the positive slope coefficient MEANSES indicates that the higher the mean SES of a school, the stronger (higher slope) the relation of SES to MATHACH, controlling for other predictors in the model.
      46. Final estimation of fixed effects:
      47. ----------------------------------------------------------------------------
      48. Standard Approx.
      49. Fixed Effect Coefficient Error T-ratio d.f. P-value
      50. ----------------------------------------------------------------------------
      51. For INTRCPT1, B0
      52. INTRCPT2, G00 12.096006 0.198734 60.865 157 0.000
      53. SECTOR, G01 1.226384 0.306272 4.004 157 0.000
      54. MEANSES, G02 5.333056 0.369161 14.446 157 0.000
      55. For SES slope, B1
      56. INTRCPT2, G10 2.937981 0.157135 18.697 157 0.000
      57. SECTOR, G11 -1.640954 0.242905 -6.756 157 0.000
      58. MEANSES, G12 1.034427 0.302566 3.419 157 0.001
      59. ----------------------------------------------------------------------------
      60. Final estimation of fixed effects
      61. (with robust standard errors)
      62. ----------------------------------------------------------------------------
      63. Standard Approx.
      64. Fixed Effect Coefficient Error T-ratio d.f. P-value
      65. ----------------------------------------------------------------------------
      66. For INTRCPT1, B0
      67. INTRCPT2, G00 12.096006 0.173699 69.638 157 0.000
      68. SECTOR, G01 1.226384 0.308484 3.976 157 0.000
      69. MEANSES, G02 5.333056 0.334600 15.939 157 0.000
      70. For SES slope, B1
      71. INTRCPT2, G10 2.937981 0.147620 19.902 157 0.000
      72. SECTOR, G11 -1.640954 0.237401 -6.912 157 0.000
      73. MEANSES, G12 1.034427 0.332785 3.108 157 0.003
      74. ----------------------------------------------------------------------------
      75. *Random effects*. The variance components table has one row for each random effect. Here the random effects are the error term for the level-2 intercept (U0), the error term for the level 2 slope (U1), and the error term for the level 1 equation (R).
          * U0. The level 2 error term U0 estimates the level 1 intercept. It is significant, indicating that average math achievement (which is what the intercept reflects) varies significantly between schools.
          * U1. The level 2 error term U1 estimates the slope of the level 1 SES variable. The variance for the slope of SES is .15 (same coefficient as in the tau matrix above). That its p > .05 means it cannot be said to be different from 0. Thus we conclude that there is no variance between schools in the slope of centered SES, where centered SES = SES - MEANSES. That is, the relationship of SES to MATHACH cannot be said to differ between schools.
          * R. While the level 1 error term does not have a significance test, it is associated with by far the largest variance component (36.70). This means that most of the variance in math achievement scores is not explained by a model predicting these scores from centered SES adjusted for level 2 MEANSES and SECTOR.
      76. Final estimation of variance components:
      77. -----------------------------------------------------------------------------
      78. Random Effect Standard Variance df Chi-square P-value
      79. Deviation Component
      80. -----------------------------------------------------------------------------
      81. INTRCPT1, U0 1.54271 2.37996 157 605.29503 0.000
      82. SES slope, U1 0.38590 0.14892 157 162.30867 0.369
      83. level-1, R 6.05831 36.70313
      84. -----------------------------------------------------------------------------
      85. *Intraclass correlation (ICC)*. With the residual within-schools effect (R) and the two random between-schools effects (here the intercept and slope of centered ses), the total of variance components is the within-group variance (sigma-squared = 36.70313) plus the level 2 between-group variance terms: 2.37996 (the level 2 intercept variance estimate) + 0.14892 (the level 2 slope variance estimate) plus twice the covariance: 2\*0.19058. For these data, the total variance sums to 39.61327. For a variance components model, constrained to have zero covariance of the random effect error terms, there would be no 2\*cov term. For other models the covariance must be taken into account because the variance components are no longer additive in a simple way. The ICC, which unlike that in the null model, is now a partial coefficient, is the between-groups effect on the intercept of the outcome variable (2.37996) divided by total variance (39.61327) = .06. We may say that 6% of the variance in math achievement is attributable to between-schools effects on mean math achievement, controlling for other variables in the model.
      86. *Calculating the level 2 effect.* One may calculare the total level 2 effect as the sum of the level 2 intercept effect (2.37996) plus the level 2 slope effect (0.14892) plus twice the covariance of intercept and slope error terms at level 2 (2\*0.19058), this quantity divided by the total of variance components as calculated above for ICC (39.61327). Making this calculation yields the value .0735. We may say that taking level 2 into account explained 7.35% of the total variance in math achievement. This is percentage often is considered a better effect size measure for LMM than is the ICC. (Note that in variance components models the 2\*cov term drops out because VC models constrain random effects to have zero covariance. HLM uses an unstructured rather than VC covariance type assumption).
      87. *Likelihood ratio test.* In the Other Settings, Hypothesis Testing dialog, the deviance for the intercept-only model was entered as the comparison for the likelihood ratio test. Deviance for the intercept-and-slopes-as-outcomes model is significantly lower than for the intercept-only model by the likelihood ratio test, indicating it is a significantly better model, though not by a lot.
      88. Statistics for current covariance components model
      89. --------------------------------------------------
      90. Deviance = 46501.875643
      91. Number of estimated parameters = 4
      92. Variance-Covariance components test
      93. -----------------------------------
      94. Chi-square statistic = 614.91783
      95. Number of degrees of freedom = 2
      96. P-value = 0.000

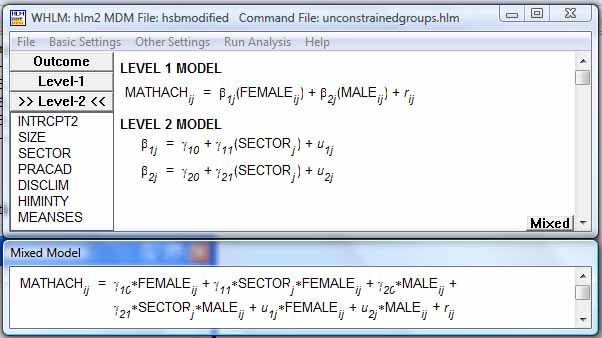
* + - *A heterogenous two-level random coefficients intercepts-and-slopes-as-outcomes model, using HLM software*. This model is the same as the previous one, but introduces HLM software's ability to construct heterogenous models. HLM allows the variance of the level 1 random error term to be modeled using a grouping variable which is not otherwise a predictor variable. In the current example, this is done by using the level 1 gender variable FEMALE, which is hypothesized to influence level 1 residual variance but is not used as a level 1 covariate in the actual model.
      1. *Heterogeneity of variance test.* HLM supports a heterogeneity of variance test to detemine if such models are needed. Select Other Settings, Hypothesis Testing,and check the "Test homogeneity of level-1 variance" checkbox, as illustrated [above](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#lrtest). When significant, as below for the intercepts-and-slopes-as-outcomes model just discussed, the researcher rejects the null hypothesis that level 1 variance of residuals is homogenous. Lack of homogeneity may be due to a variety of causes including outliers, non-normal (heavily kurtotic) data, omitting one or more important level 1 variables, or treating an included level 1 variable as a fixed effect when it is not.
      2. Test of homogeneity of level-1 variance
      3. ----------------------------------------
      4. Chi-square statistic = 244.08638
      5. Number of degrees of freedom = 159
      6. P-value = 0.000
      7. *Identifying variables associated with heterogeneity*. The HLM heterogeneity of variance test does not tell the researcher which other predictors may be affecting residual variance. This can be explored using SPSS to view the variances of OLS residuals by a variable such as gender. Select Analyze, Regression, Linear; set Dependent = MATHACH, Independent = SES, Selection variable = FEMALE with Rule FEMALE = 0 in a first run, the FEMALE = 1 in a second run. In the ANOVA table output, residual sum of squares for males (FEMALE = 0) is 148,250 and for females (FEMALE = 1) it is 144,022. (Alternatively, the Save button would allow residuals to be saved on each run and then their variances could be viewed under the Descriptive statistics option.) Although this is not a great difference, for pedagogical reasons FEMALE is used to illustrate heterogenous linear mixed modeling.
      8. *Requesting heterogenous variance modeling*. Heterogenous variance models are any of the other types of models with the heterogenous sigma option selected. For instance, one could adapt the intercepts-and-slopes-as-outcomes model above to model the level 1 error term using FEMALE. To do this in HLM software, select Other Settings, Estimation Settings from the menu and then click the Heterogenous Sigma^2 button. (Recall Sigma2 reflects within group variance.) This brings up the dialog below.



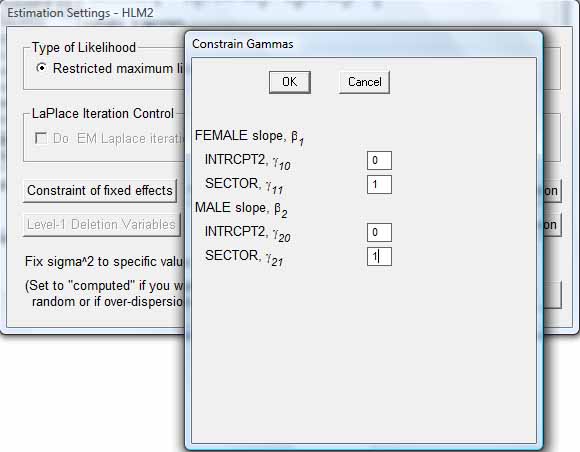
* + - 1. *The heterogenous model.* Back in the main HLM interface, a variance equation for the level 1 residual is added to the level 1 equation section. Otherwise the intercepts-and-slopes-as-outcomes model is unchanged.



* + - 1. *Likelihood ratio test*. To compare the heterogenous model with the previous homogenous variance model, a likelihood ratio test is employed. The deviance for the intercepts-and-slopes-as-outcomes model without heterogenous variance modeled was 46501.88. Asking under Other Settings, Hypothesis Testing, for a model comparison as described in previous sections, the output for the heterogenous variance modified model contains slightly different parameter estimates which are significant by this model comparison test:
      2. Statistics for current covariance components model
      3. --------------------------------------------------
      4. Deviance = 46482.093344
      5. Number of estimated parameters = 11
      6. Model comparison test
      7. -----------------------------------
      8. Chi-square statistic = 19.78230
      9. Number of degrees of freedom = 7
      10. P-value = 0.006
  + **Testing slope invariance across groups in RC models (constraining fixed effects), using HLM software.**
    - *Purpose.* With HLM it is possible to compare an unconstrained model with a constrained model. In the example below, the gender dummy variables Male and Female are level 1 determinants of math achievement. The slopes of the two gender variables are made a function of the level 2 variable SECTOR (0=public, 1=parochial). There is no level 1 intercept because the two gender dummy variables are a complete categorical set, not leaving out a category (see discussion [above](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#nointercepts)). In the unconstrained model, SECTOR may affect the Male slope differently from the Female. In the constrained model, SECTOR is constrained to affect both slopes equally. One may look at the deviance measure to see which model fits the data better (lower deviance is better).



* + - 1. *Requesting a constrained model in HLM software.* The model above may be run unconstrained, then SECTOR can be constrained to be equal for Male and Female. To do this, Other Settings, Estimation Settings is chosen from the HLM menu, then the "Constraint of fixed effects" button is clicked, bring up the dialog below. A "0" entry leaves a parameter unconstrained. A pair of "1" entries forces an equality constraint. Here, the level 2 determinants of the slopes of the level 1 variables Male and Female are constrained to be equal.



* + - 1. *Fixed effects.* In the unconstrained model, the fixed effects table illustrates the way SECTOR influences on level 1 slopes for FEMALE and MALE differ:
      2. Final estimation of fixed effects
      3. (with robust standard errors)
      4. ----------------------------------------------------------------------------
      5. Standard Approx.
      6. Fixed Effect Coefficient Error T-ratio d.f. P-value
      7. ----------------------------------------------------------------------------
      8. For FEMALE slope, B1
      9. INTRCPT2, G10 10.684432 0.298122 35.839 158 0.000
      10. SECTOR, G11 2.932540 0.446512 6.568 158 0.000
      11. For MALE slope, B2
      12. INTRCPT2, G20 12.174859 0.322616 37.738 158 0.000
      13. SECTOR, G21 2.597771 0.487027 5.334 158 0.000
      14. ----------------------------------------------------------------------------

In the constrained model, the slopes are the same. The level 2 G21 slope is not displayed as it is equal to the G11 slope:

Final estimation of fixed effects

(with robust standard errors)

----------------------------------------------------------------------------

Standard Approx.

Fixed Effect Coefficient Error T-ratio d.f. P-value

----------------------------------------------------------------------------

For FEMALE slope, B1

INTRCPT2, G10 10.723664 0.295717 36.263 158 0.000

SECTOR, G11 \* 2.804823 0.417646 6.716 158 0.000

For MALE slope, B2

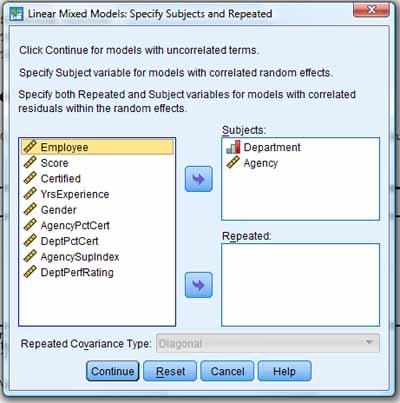
INTRCPT2, G20 12.103608 0.313462 38.613 159 0.000

----------------------------------------------------------------------------

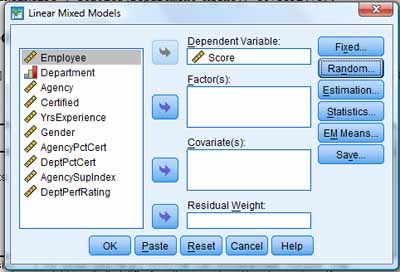
* + - 1. *Random effects.* The unconstrained and constrained variance components are very close.
      2. Final estimation of variance components (unconstrained model):
      3. -----------------------------------------------------------------------------
      4. Random Effect Standard Variance df Chi-square P-value
      5. Deviation Component
      6. -----------------------------------------------------------------------------
      7. FEMALE slope, U1 2.41260 5.82064 121 481.99916 0.000
      8. MALE slope, U2 2.64370 6.98917 121 483.25462 0.000
      9. level-1, R 6.22438 38.74285
      10. -----------------------------------------------------------------------------
      11. Final estimation of variance components (constrained model):
      12. -----------------------------------------------------------------------------
      13. Random Effect Standard Variance df Chi-square P-value
      14. Deviation Component
      15. -----------------------------------------------------------------------------
      16. FEMALE slope, U1 2.40847 5.80071 121 484.11557 0.000
      17. MALE slope, U2 2.63048 6.91943 121 483.35444 0.000
      18. level-1, R 6.22449 38.74426
      19. -----------------------------------------------------------------------------
      20. *Unconstrained deviance*. Below the deviance computed for the unconstrained model. It is entered as the comparison for the constrained model in the likelihood ratio test below.
      21. Statistics for current covariance components model
      22. --------------------------------------------------
      23. Deviance = 47012.437355
      24. Number of estimated parameters = 4
      25. *Likelihood ratio test*. The constrained and unconstrained models have the same number of estimated parameters, so the likelihood ratio test has 0 degrees of freedom, making the test unavailable. However, we can see the deviance values for the unconstrained model above and the constrained model below are practically identical. As assuming equal level 2 slopes is more parsimonious than assuming heerogenous ones, the researcher may conclude that there is slope invariance across sectors.
      26. Statistics for current covariance components model
      27. --------------------------------------------------
      28. Deviance = 47014.952691
      29. Number of estimated parameters = 4

**Three-Level Random Coefficients Models**

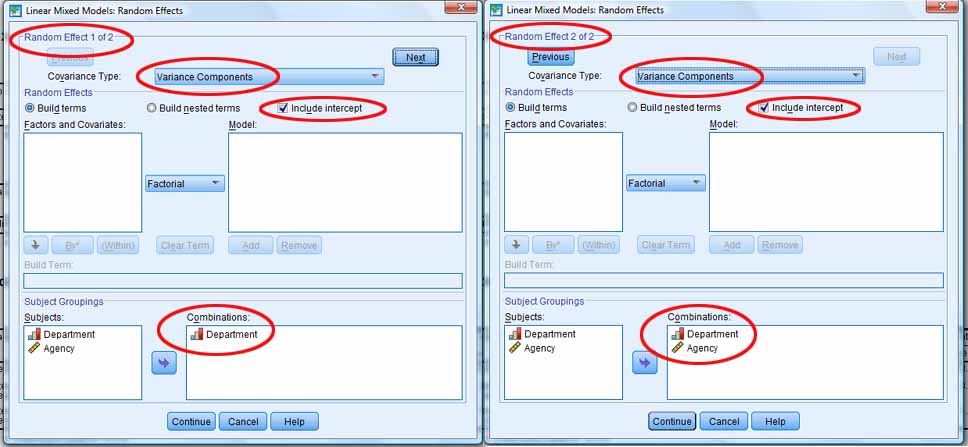
* + **A three-level null linear mixed model, using SPSS**
    - *Description.* This is a type of random intercept model. It is also a type of two-way Anova with random effects. There are no predictors at any of the three levels but the dependent variable at level 1 is made a random effect of the level 2 and level 3 grouping variables.
    - *Example*. Let level 1 be the employee level, level 2 the agency level, and level 3 the department level. Employees are nested within agencies, and agencies are nested within departments. Let "Score" be the continuous dependent variable, representing performance scores of employees. We wish to see if there is an agency and/or department effect on performance scores. The SPSS syntax for this example is:
    - MIXED Score
    - /CRITERIA=CIN(95) MXITER(100) MXSTEP(5) SCORING(1) SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE) LCONVERGE(0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
    - /FIXED=| SSTYPE(3)
    - /METHOD=ML
    - /PRINT=G SOLUTION TESTCOV
    - /RANDOM=INTERCEPT | SUBJECT(Department) COVTYPE(VC)
    - /RANDOM=INTERCEPT | SUBJECT(Department\*Agency) COVTYPE(VC).
    - *Subject variables*. As illustrated below, both Department and Agency (which are both id variables) are entered as Subject variables.



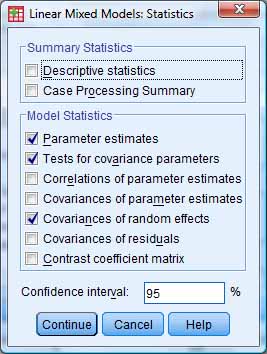
* + - *Dependent variable*. Score, a continuous variable representing level 1 employee performance score, is entered as the dependent variable on the next SPSS dialog screen.



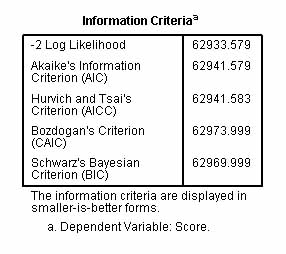
* + - *Independent variables*. Apart from the grouping variables, there are no factors or covariates. No variables are entered as fixed effects under the Fixed button.
    - *Random effects*. There are two random effects: (1) the effect of the level 3 grouping variable, Department; and (2) the effect of the level 2 grouping variable, Agency, nested within Departments. After clicking the Random button, these two effects must be entered. This is accomplished by entering the first (below, left), then clicking the Next button and entering the second (below, right). Note the "Include intercept" checkboxes are checked since we are modeling the intercept of Score. The covariance type is set to "Variance Components" as this is the default for random intercept models, which the null model is. One could set the type to "Scaled Identity", which is the type often used when modeling the interaction of a random factor (ex., Department) with a fixed grouping factor (ex., Agency entered also as a factor and a fixed effect, not done in this example). For this example, the VC and ID covariance types would lead to identical results.



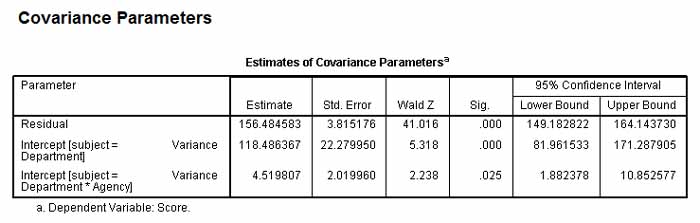
* + - *Estimation method*. As discussed [above](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#estimation), the default estimation method is REML. REML is better for comparing variance components but cannot be used to compare regression coefficients in the fixed effects portion of LMM output. In this example it is set to ML (maximum likelihood estimation) because we wish consistency with examples comparing fixed effects. ML is chosed under the Estimation button, not illustrated.
    - *Statistical options* are selected under the Statistics button in SPSS, the dialog for which is illustrated below: However, if no options are selected, the default output will still generate the two needed tables: the "Information Criteria" table and the "Estimates of Covariance Parameters" table (see below). However, to get significance tests in this table we must check "Tests for Covariance Parameters".



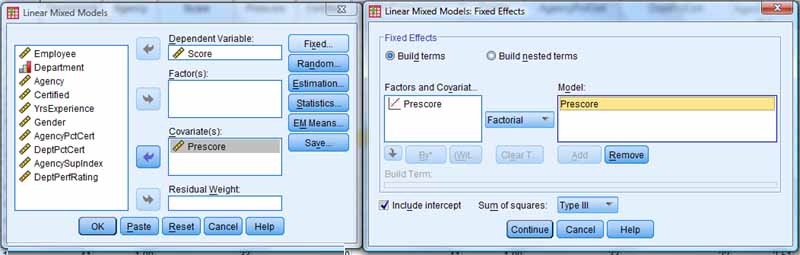
* + - *Information Criteria table*. In later models, when predictor variables are added in addition to the grouping variables (Agency, Department), we expect all these measures to go down. The likelihood ratio is the "-2 Log Likelihood" estimate and is the baseline needed when, for later models, a likelihood ratio test is performed to see if the model with predictors is significantly better than the model without predictors. That is, -2LL will be used to test the overall significance of the researcher's model. The other information theory measures in this table are explained in the [separate module for structural equation modeling](http://faculty.chass.ncsu.edu/garson/PA765/structur.htm#infotheory), in which they also appear.



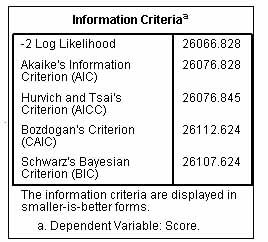
* + - *Estimates of Covariance Parameters table*. Below, the variance in employee performance scores associated with differences in departments (118.49) is larger than the variance associated with differences in agencies (4.52). The within-subjects (employee level) variance in scores (156.48) is shown to be a higher variance component than either the Department or the Agency effect. This means most of the variance in scores needs to be explained by predictors at some level other than the grouping variables. The sum of variance components is 279.49. Dividing by this denominator, we find the Department effect accounts for 42.4% of the variance in performance scores while the Agency effect accounts for 1.6%. Some 56.0% of the variance in scores remains unexplained (the Residual component divided by the total). All effects are significant, as shown by the Wald test. Because of significant Agency and Department effects, hierarchal linear modeling is justified.



* + **Three-level random intercept model with a level 1 control covariate, using SPSS**.
    - *Example*. This is a model in which a level 1 covariate is added as a control. The example is the same as in the null model above, except that "prescore" is added as a level 1 control variable. "Prescore" represents the previous performance score for each employee. Predicting score controlled for prescore means, in essence, that we are predicting change in performance score.
    - *SPSS.* The SPSS syntax for this model is the same as for the null model, except adding prescore as a fixed effect:
    - MIXED Score WITH Prescore
    - /CRITERIA=CIN(95) MXITER(100) MXSTEP(5) SCORING(1) SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE)
    - LCONVERGE(0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
    - /FIXED=Prescore | SSTYPE(3)
    - /METHOD=ML
    - /PRINT=G SOLUTION TESTCOV
    - /RANDOM=INTERCEPT | SUBJECT(Department) COVTYPE(VC)
    - /RANDOM=INTERCEPT | SUBJECT(Department\*Agency) COVTYPE(VC).
    - *Entering a control covariate*. In this example most menu specifications are the same as for the null model, except prescore is entered as a covariate and then under the Fixed button, is declared a fixed effect.



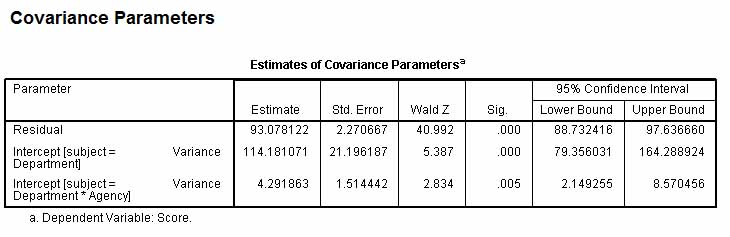
* + - *"Information Criteria" table*. Lower values on this table correspond to better model fit. Note, however, that by controlling prescore, we are now predicting change in performance score - in essence, a different dependent variable. Performing a likelihood ratio test comparing the deviance in this model with that in the null model is questionable since the the dependent variable in the null model was performance score whereas in the present model it is, in essence, change in performance score. Researchers might wish to treat the present model as the baseline, not the null model.



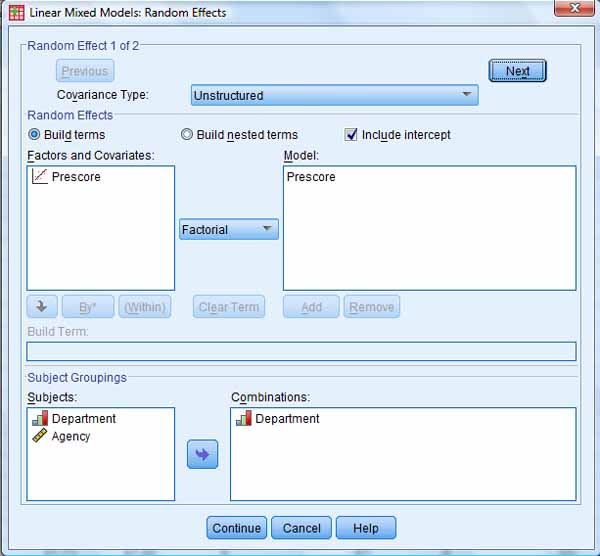
* + - *"Estimates of Fixed Effects" table*. Since prescore was entered as a fixed effect, output includes a fixed effects table showing, indeed, that prescore is highly significant.



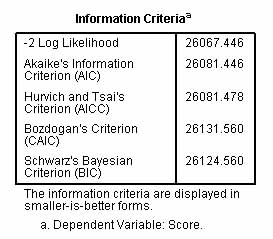
* + - *"Estimates of Covariance Parameters" table* . With prescore as a control variable in the model, the unexplained or within-subjects residual variance in employee performance scores shrinks to be 44.0% of the total while the Department effect rises to become 54.0% and the Agency effect 2%. (Percentages, as before, are simply the corresponding variance components divided by the sum of all components in a VC model.)



* + *A three-level random coefficient model with a level 1 control covariate, using SPSS*.
    - *Description.* This model is identical to the one above except that the level 1 control variable, prescore, is made a random effect. Having determined that the Agency effect is small in the previous example, a possible Department effect not only on performance score (the intercept of the dependent) but also on the strength of relationship between prescore and score (that is, on the slope or b coefficient of prescore, the control variable).
    - *SPSS.* The SPSS syntax is below:
    - MIXED Score WITH Prescore
    - /CRITERIA=CIN(95) MXITER(100) MXSTEP(5) SCORING(1) SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE) LCONVERGE(0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
    - /FIXED=Prescore | SSTYPE(3)
    - /METHOD=ML
    - /PRINT=G SOLUTION TESTCOV
    - /RANDOM=INTERCEPT Prescore | SUBJECT(Department) COVTYPE(UN)
    - /RANDOM=INTERCEPT | SUBJECT(Department\*Agency) COVTYPE(VC).
    - *Setting a random effect.* Under the Random button , Prescore is declared to be a random effect to be modeled by Department. The Department effect is the first of the two random effect screens, which is changed to look as below. There are two changes in this dialog window: (1) declaring Prescore to be a random effect to be modeled by Department as a grouping variable, and (2) declaring the covariance type to be unstructured (UN), which is the commonly used for random coefficients models (here, where the b coefficient of Prescore is modeled as a random effect).

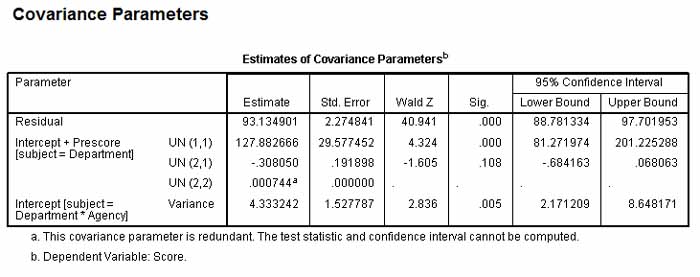


* + - *"Information Criteria" table*. The information criteria measures show very little change from the random intercept model. This strongly suggests what is uncovered in the "Estimates of Covariance Parameters" table - that there is little effect. That is, Department has a strong effect on the intercept (level) of performance score, but no significant effect on how strongly prescore is related to score (that is, on the b coefficient).

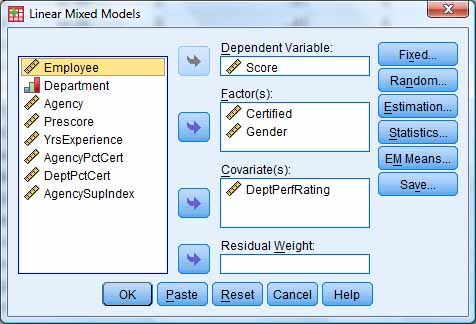


* + - *"Estimates of Fixed Effects" table*. Though not illustrated here, this table continues to show that prescore is highly significant as a predictor of score.
    - *"Estimates of Covariance Parameters" table*. Since an unstructured covariance type was specified, the variance components include the UN(1,1), UN(2,1), and UN(2,2) components. As discussed above in the section on the [random regression models with unstructured covariance assumption](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#rccompare), these refer respectively to components having to do with the intercept, the intercept-slope covariance, and the slope.

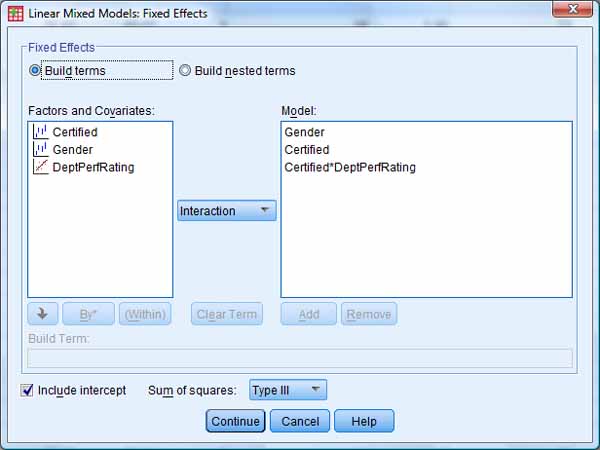
In the output below, the "Residual" variance component is large and significant, meaning that much of the within-subjects variance in employee performance remains unexplained. The UN(1,1) component shows that there is a large and significant between-departments effect explaining the variance in employee performance scores, even after prescore is controlled (that is, there is a large and significant Department effect on change in Score). The UN(2,2) component is so small that it is redundant: it does not explain any additional variance in performance score once prescore is controlled and the Department effect is taken into account. The covariance on intercepts and coefficients, UN(2,1), is also non-significant. In sum, Department heavily influences average performance score even after prescore is controlled, but it does not influence the strength of relationship between score and prescore.



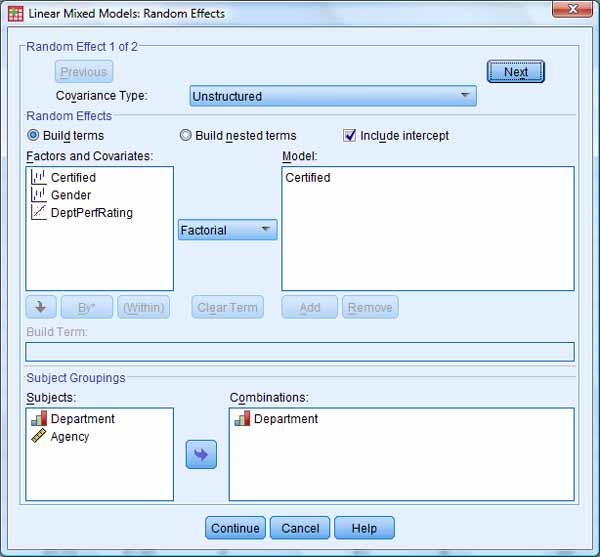
* + **Three-Level Random Coefficient Model with a Level 1 and Level 2 Covariates, Using SPSS**.
    - *Example*. In this example we drop prescore and go back to simply predicting employee performance score (rather than score controlling for prescore, which is makes change in score the dependent). We then add two level 1 covariates: Gender (Male=0, Female=1) and Certified (0=not certified, 1=certified). We then consider a level 3 covariate, DeptPerfRating, which is department-level performance rating. We do no consider DeptPerfRating to be a direct effect on level 1 score and hence do not add it as a fixed effect. However, we add the interaction term Certified\*DeptPerfRating to see if Departmental Performance Rating affects the b coefficient of Certified as a level 1 predictor of Score. We also model Certified as a random effect of the grouping variable Department. Finally, we invoke the EM Means button to ask for estimated marginal means output for Gender, discussed below.
    - *SPSS.* The SPSS syntax for this model is:
    - MIXED Score BY Certified Gender WITH DeptPerfRating
    - /CRITERIA=CIN(95) MXITER(100) MXSTEP(5) SCORING(1) SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE) LCONVERGE(0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
    - /FIXED=Gender Certified Certified\*DeptPerfRating | SSTYPE(3)
    - /METHOD=ML
    - /PRINT=G SOLUTION TESTCOV
    - /RANDOM=INTERCEPT Certified | SUBJECT(Department) COVTYPE(UN)
    - /RANDOM=INTERCEPT | SUBJECT(Department\*Agency) COVTYPE(VC)
    - /EMMEANS=TABLES(Gender) COMPARE ADJ(LSD).
    - *Entering variables*. The initial "Subjects and Repeated" screen remains the same in this example. The second screen, on which types of variables are declared is shown below. Note that because Gender and Certified are entered as factors rather than as covariates, by default SPSS will predict the "0" levels (Gender=0=Male, Certified=0=Not Certified) and used the "1" level as the reference level. If entered as covariates, the "1" level would be predicted.



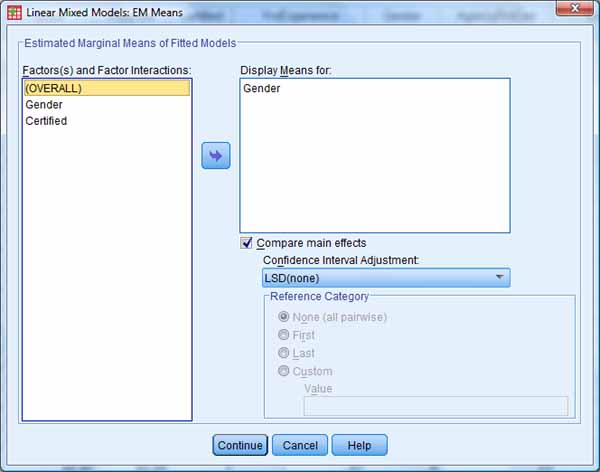
* + - *Fixed effects*. The fixed effects screen, invoked by the Fixed button, is shown below. Note the level 1 covariate interaction with the level 3 covariate: Certified\*DeptPerfRating. This will test the effect of DeptPerfRating on the b coefficient of Certified as a level 1 predictor.



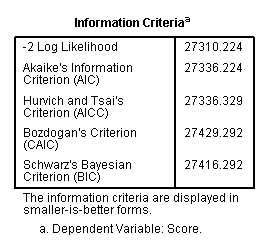
* + - *Random effects*. On this screen, shown below, Certified is modeled as a random effect of the Department grouping variable.



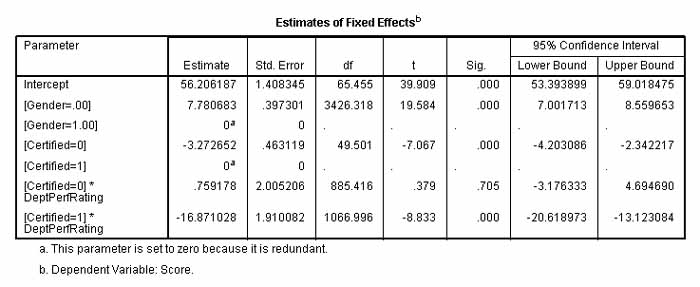
* + - *EM Means*. On this screen, shown below, estimated marginal means are requested for Gender.



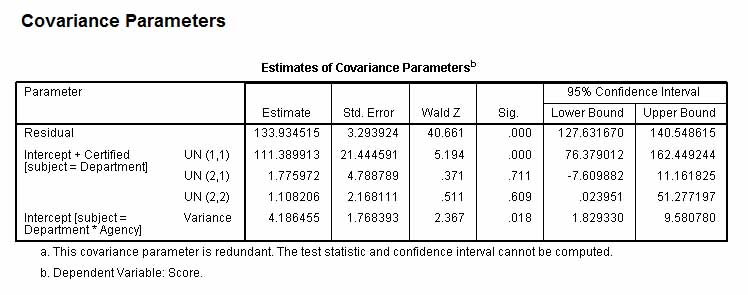
* + - *Information Criteria table*. This table shows a substantial drop in deviance (-2LL) compared to the null model. One could use the two deviance coefficients to test whether the current model significantly differs from the null model (it does). We would not use -2LL to compare non-nested models, such as the current one with the one previously discussed. Information theory measures like BIC are sometimes used to compare non-nested models, with lower BIC being better. As the immediately preceding model, which included prescore as a predictor, had a slightly lower BIC, it was the more explanatory of the two models. Both the current model and the preceding one are markedly better than the null model. By the BIC criterion, the current model with several effects (but not prescore) is about as powerful as the previous model which included prescore as a predictor of score.



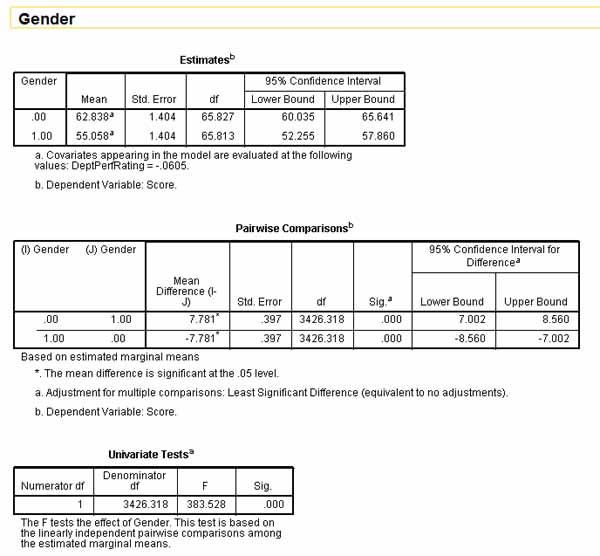
* + - *Estimates of Fixed Effects table*. The table below shows the main effects of Gender and Certified on employee performance score are significant. The interaction term of Certified at level 1 with DeptPerfRating at level 3 is significant for those who are certified (Certified=1) but not for those who are not.



* + - *Random effects*. On the "Estimates of Covariance Parameters" table, random effects are shown. The significant variance components are the Residual, representing the unexplained within-subjects variance in employee performance score; UN(1,1), reprsenting the between-Departments effect on the intercept of performance score; and Intercept[subject=Department\*Agency], representing the between-Agency effect on the intercept. UN(2,2), representing the between-Department effect on the slope of Certified, was not significant, nor was UN(2,1), representing the covariance of slopes and intercepts.



* + - *Estimated Marginal Means*. This part of the output shows the effect of Gender on mean performance score. The "EMM Estimates table" gives the mean, standard error, degrees of freedom, and 95% confidence level for each level of gender (male=0, female=1). Below we see males and females differ appreciably (62.8 for males, 55.1 for females) on mean Score.
      1. *Pairwise comparisons.* In the "Pairwise Comparisons table" (which would be of greater interest if there were more categories than the two reflecting Gender) gives the mean difference between categories, the standard error of the difference, its significance level, and the 95% confidence interval for the difference. For this example, the difference between men and women mean scores is significant and therefore can be assumed to differ from 0. Of course, if the comparison were done on a variable with more than two categories (ex., ethnicity), then there will be more than one pairwise comparison to assess.
      2. *Test of the gender effect.* The "Univariate Tests" table gives an F test of significance for the effect of Gender on the dependent variable, based on the linearly independent pairwise comparisons among the estimated marginal means. For Gender, which has only two categories and thus only one pairwise comparison, this significance level will be the same as that reported in the Pairwise Comparisons table. However, for a variable with more than two categories, the Univariate Tests significance level tests the significance of the predictor variable overall (ex., ethnicity), whereas the Pairwise Comparisons table will report a significance level for each pair of categories (ex., Irish vs. Italian ethnicity) but no overall significance test of ethnicity.



**Longitudinal, Growth, and Repeated Measures Models**

* + **Terminology**
    - *Nomenclature*. Longitudinal, growth, and repeated measures models can be handled by linear mixed modeling software. While these are highly overlapping terms, each has connotations. "Repeated measures modeling" tends to refer to two-level studies with measures over time at level 1 and subjects (usually individuals) at level 2. "Growth modeling" tends to refer to studies in which there is an explicit time variable and a purpose of the study is to understand time patterns (growth). "Longitudinal modeling" is an umbrella term used to refer to any study in which measures are collected for more than one time period.
  + **An intercept-only longitudinal model, using SPSS**.
    - *Example*. In in this and subsequent SPSS examples, the following variables are used:
      1. empl\_id: an id variable for the individual employee
      2. empl\_ses: the socioeconomic status score for the employee, centered such that 0 = mean ses
      3. agcy\_id: an id variable for the employing agency
      4. agcy\_ses: the socioeconomic status mean score for the agency
      5. gender: 0 = male, 1 = female
      6. seniority: 0 = lacks seniority, 1 = has seniority
      7. trainyrs = number of years of training
      8. test\_seq: test sequence number: 0, 1, or 2, representing a time variable with equal intervals between administrations of a performance test
      9. seq\_sq: the time variable squared
      10. score: the performance test score of the employee, which is the dependent variable at level 1
    - *Description*. In this simple longitudinal model, the dependent variable is score at level 1, grouped by employee id as the subjects variable at level 2. **The only random effect in this model is the intercept.** **Specifically, time is neither a random effect nor a repeated measure** (ensuing models will explore these). Rather, time (test\_seq) is treated as a fixed effect entered as a covariate (since it is a metric, administered at equal intervals). Compared to OLS regression, this model will not change the actual estimates of the intercept or the b coefficients, but it will change the standard errors of the slopes and hence the significance tests.
    - *SPSS*. By checking the "Include intercept" box on the "Random Effects" dialog screen and not entering any other factors or covariates, the only random effect is the intercept grouped by subject (Combinations=empl\_id in the "Subject Groupings" area. As shown below, select Analyze, Mixed Models, Linear and then as illustrated below.
      1. Under "Specify Subjects and Repeated" enter the id variable as "Subjects"
      2. In the "Linear Mixed Models" dialog, enter performance score as the "Dependent Variable" and make the time variable(s) the "Factor(s)" or "Covariate(s)", depending on whether they represent equal intervals (here they are covariates)
      3. Click the Fixed button and Model the time variable(s), making sure to check "Include Intercept"
      4. Click the Random button and set the "Covariance Type" to "Unstructured" or other assumption. (However, as there are no random effects modeled, SPSS will return the notification "Warnings: The covariance structure for random effect with only one level will be changed to Identity.")
      5. Make the id variable the "Subject Groupings" in the Combinations area. Be sure to check "Include Intercept".

The SPSS syntax looks like this:

MIXED score WITH test\_seq seq\_sqr

/CRITERIA=CIN(95) MXITER(100) MXSTEP(10) SCORING(1) SINGULAR(0.000000000001)

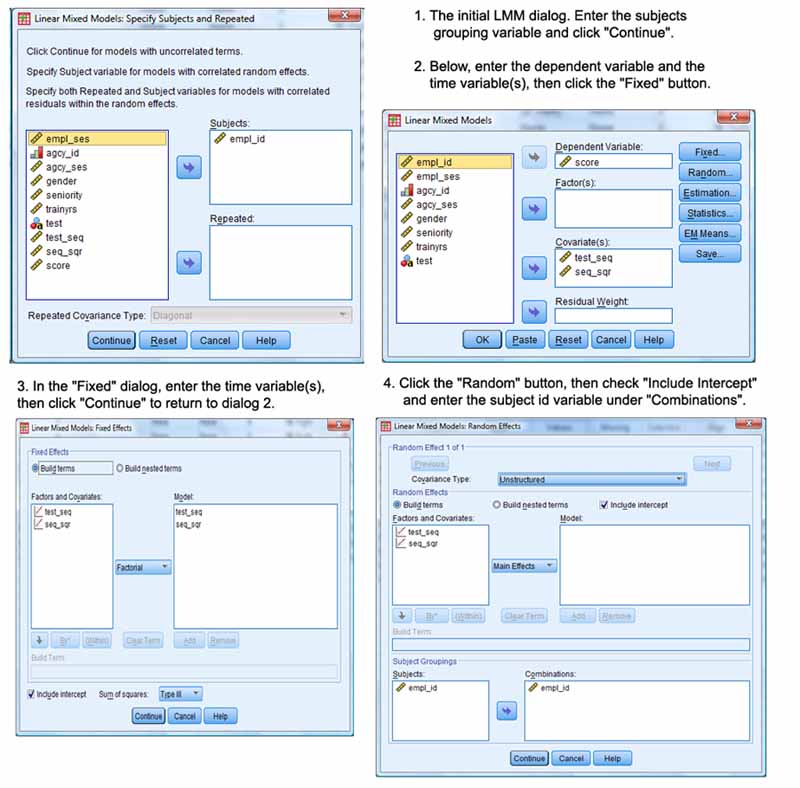
HCONVERGE(0, ABSOLUTE) LCONVERGE(0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)

/FIXED=test\_seq seq\_sqr | SSTYPE(3)

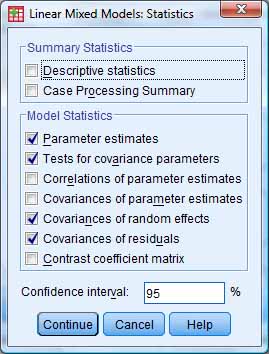
/METHOD=REML

/PRINT=SOLUTION TESTCOV

/RANDOM=INTERCEPT | SUBJECT(empl\_id) COVTYPE(UN).



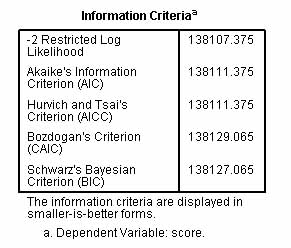
After these steps, click the Statistics button and select the desired output; click Continue; then click OK to run the model.



Note that employee id is entered twice, as both the Subjects variable and the Subjects-Combinations variable. The potential combinations variables are those entered on the initial "Subjects and Repeated" screen, here just empl\_id. In other examples, the researcher might have entered more than one grouping variable on the "Subjects and Repeated" screen, then select the one(s) wanted for a particular computer run on the "Random" screen in the "Combinations" area.

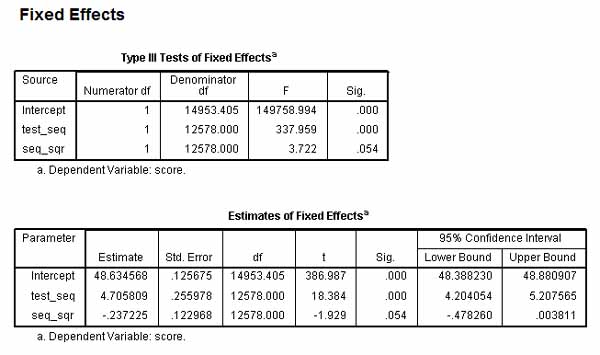
Time is modeled as a fixed effect. In this example test score is level 1 and empl\_id as Subjects-Combinations variable is level 2.

* + - *The Information Criteria table.* The information criteria effect size measures for the intercept-only model may be used as a baseline to compare other later more complex models. Here the baseline AIC is 138,111.4.

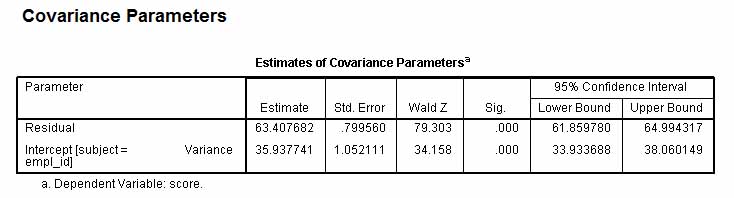


* + - *Fixed Effects.* This table contains the same estimates as the regression table for the intercept and the b coefficients for the time variables. The standard errors for these b coefficients are different, however, because the clustering effect of employees is now adjusted. In this table, for the example data, we test to see if the dependent variable (score) is linearly or quadratically related to time as represented respectively by the variables test\_seq or seq\_sq. If the F test for the time variable is significant, as it is here for test\_seq, then performance score varies by time of measurement within the same individuals. That is, there is a "time effect" by which subjects perform better at later testing times (assuming a positive relationship).

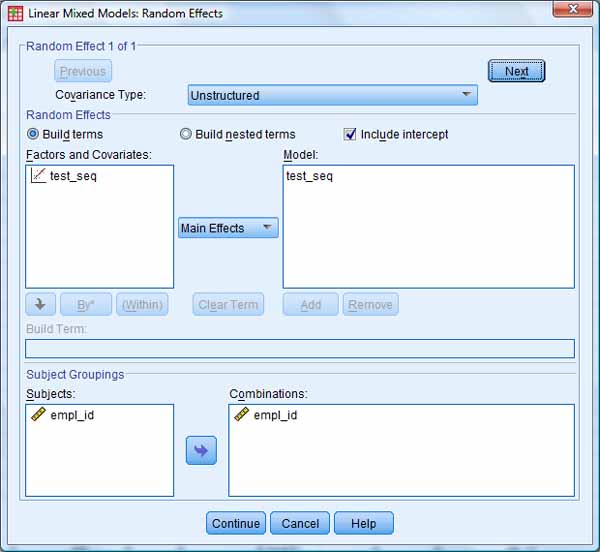
For this example, there is a linear time effect but not a nonlinear effect (at least not a quadratic effect, which is the most common type of nonlinear effect). That the linear effect is significant may lead us to investigate the growth model when additional variables are introduced. That the nonlinear effect is not significant may lead us to drop time-squared (seq\_sq) from further analysis.



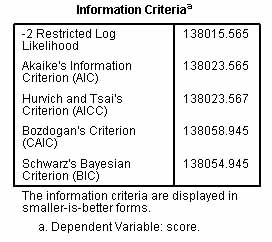
* + - *The "Estimates of Covariance Parameters" table.* In the example below we come to the unsurprising finding that the intercept as a random effect is significant. This means that between-subjects effects do impact score. Put another way, mean scores of employees differ.
      1. *Unstructured covariance structure model.* In the UN model in the example, we can state that the variability in score is partly explained by the between subjects (intercept) effect, but even more of the variability is explained by the within subjects (residual) effect, which in a growth model is the time effect since test times are grouped by employee. Put another way, the residual variance estimate is large and significant means there is a time effect, most likely because employees tend to score higher in successive tests, and this effect is greater than the effect of variation in score between individual employees.
      2. ***Variance components covariance structure model.* In a VC model (not the case here since "Unstructured" was assumed), the variance of the intercept term divided by total variance (intercept + residual variance) is the percent of variance in score explained by between-subject effects in a variance components model. For example, if the model were recalculated on a variance components assumption and the same estimates resulted, then 35.94/(35.94+63.41) = 36% of the variance in performance score would be attributable to variability between subjects. The variance components assumption is required because it supports additivity of components.**



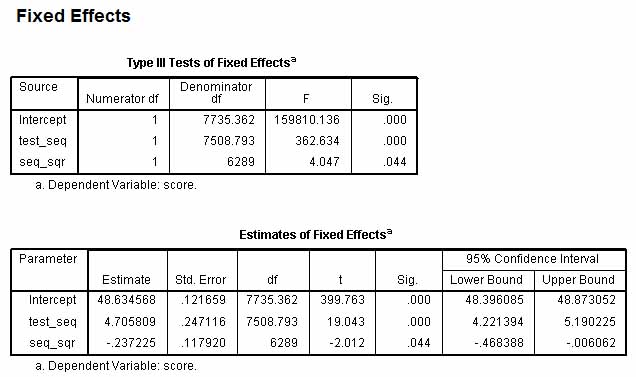
* + **An intercept-plus-time (random coefficients) growth model with time as a random effect, using SPSS**.
    - *Modeling time as a random effect*. In this model, time is made a random effect but is still not declared as a repeated variable. The estimates of the intercept and b coefficients for the time variables (time\_seq and seq\_sq) will be the same as in the OLS regression and intercept-only models. Whereas in the intercept-only model only the standard errors of the b coefficients changed, now in the intercept + time model, the standard errors of the intercept as well as of the b coefficients change (and so also the significance tests). That is, by making time a random effect we ask that the standard errors of the intercept as well as of the slopes of the fixed time variables be adjusted for the clustering effect of empl\_id as the "Subjects" variable. The standard errors for this model will differ from the regression model and from the intercept-only model.
    - *SPSS*. The SPSS syntax looks like this:
    - MIXED score WITH test\_seq seq\_sqr
    - /CRITERIA=CIN(95) MXITER(100) MXSTEP(10) SCORING(1) SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE) LCONVERGE(0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
    - /FIXED=test\_seq seq\_sqr | SSTYPE(3)
    - /METHOD=REML
    - /PRINT=SOLUTION TESTCOV
    - /RANDOM=INTERCEPT test\_seq | SUBJECT(empl\_id) COVTYPE(UN).
    - *Modeling time as a random effect*. Time (the test\_seq variable in the example) is modeled as a random effect as illustrated below. There will now be two random effects: the intercept and time. (In the intercept-only model above, only the intercept was a random effect). Time is made a random effect by moving it into the "Model:" area after clicking the "Random" button, as illustrated below. In spite of having been not significant in the intercept-only model, the quadratic time term, seq\_sq, is retained as a fixed factor and other model settings remain as described above for the intercept-only model. This will cause the random effects to be "Residual" and also "Intercept + test\_seq[subject=empl\_id]" in this example, where test\_seq is the linear time variable.



* + - *Information Criteria table.* For the baseline intercept-only model the baseline AIC was 138,111.4. With time modeled as a random effect, AIC is now 138,023.6. As lower AIC is better, this is a modest improvement.



* + - *Fixed effects* In this intercept + time model, the "Estimates" column in the "Estimates of Fixed Effects" table contains the estimates of the intercept and of the slopes for the time variables. The regression coefficient (slope) for the time variable (test\_seq) indicated the number of performance points employees change above the mean (the intercept = 48.63) on the average for each unit increase in time as measured by the variable test\_seq, controlling for other variables in the model. That is, an employee's score can be expected on the average to increase 4.71 points between test time = 1 and test time = 2. Since quadratic time (seq\_sq) goes from 1 to 4 in this interval, we then must subtract -.24\*3, so the increase in score is roughly 4 points. We conclude that test scores do increase modestly over time even when the standard errors of the slope and intercepts are adjusted for the clustering effects of empl\_id as a "Subjects" variable. We also conclude there is a marginally significant negative quadratic time effect, meaning the propensity for scores to increase over time decelerates with time.

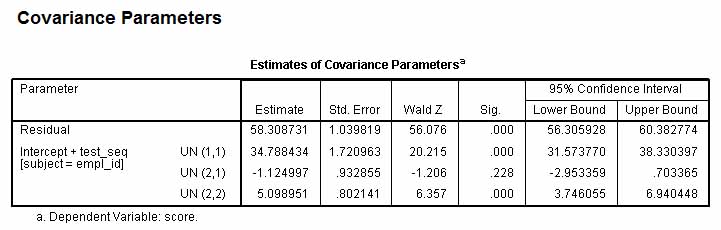


Note that quadratic time (time\_seq) has now flipped back below the critical .05 significance level and may be considered significant when the intercept is adjusted for the clustering effect of empl\_id by virtue of declaring test\_seq to be a random effect under the "Random" button dialog.

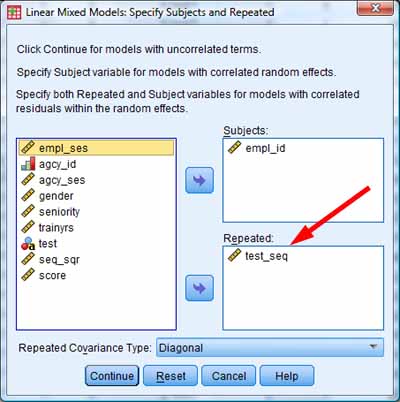
* + - *Random effects* are shown in the "Estimates of Covariance Parameters" table. The estimates in this table lack an easily-interpretable absolute meaning but are used instead to assess variance components in the model. If a "Variance Components" covariance type assumption has been made, this is made easier as variance components are then additive and percentages may be imputed to each source of variation. The task is more challenging in this example, where an "Unstructured" type was selected. In the "Estimates of Covariance Parameters" table there are parameters for "Residual" and for "Intercept + test\_seq[subject=empl\_id]") with rows for three "UN()" parameters discussed below.
      * + *Residual.* The "Residual" covariance estimate reflects within-subjects variation. For this example, what is "within" is the three test occasions per employee. Therefore the residual reflects the estimated variability in performance score across time for the average employee. That is, it is the time effect, which here is shown to be significant.

The estimate for the residual term can be compared with the corresponding residual estimate for the intercept-only model to see if there has been a change (typically the researcher defines improvement as a lower estimate). For this example, the residual estimate for the intercept-only model was 63.41. For the intercept+time model, it is 58.31. This means that the time effect on the variance of score is less when time is modeled as a random effect of empl\_id as a level 2 grouping variable, in comparison to the intercept-only model. Later a similar comparison may be used to assess effects of adding one or more covariates to the model.

* + - * + UN(1,1): this is the between-subjects variance estimate for the error term for the intercepts, which represent mean scores. Its significance test tests if mean scores vary significantly between employees. For these data, mean employee score across the three test occasions does vary significantly between employees at the .000 level. This means that there is significant unexplained variability (reflected in variability of the error term for the intercepts) between employees in mean score across time. A more complex model with covariates may be needed.
        + UN(2,2): this is the between-subjects variance estimate for the error terms for the slopes for the time variables. It is significant in this example, meaning the regression slopes for the time variables also vary significantly between employees.
        + UN(2,1): this is the between-subjects estimate of covariance between intercepts and slopes, the significance test for which tests intercept-slope interaction. That is, this is the estimate for the covariance between the error terms for the time variables and the error terms for the intercept, which represents mean test score. Since not significant for this example, the error terms for the intercepts and for the slopes of the time variables do not covary significantly. This means that error in estimating the slopes is unrelated to error in estimating the intercept or vice versa. This in turn implies no significant intercept-slope interaction.



* + **A repeated measures intercept-plus-time model using SPSS**
    - *Modeling time as a repeated measure and random effect*. The previous model modeled time as a random effect. In this model it is modeled as a repeated measure as well as a random effect:
      1. Time (test\_seq) is made a random effect variable under the "Random" button dialog in SPSS, as in the previous model. Time is also listed as a fixed effect. Specifying time as a random effect to be modeled means one regression will be performed for each subject. Time as a fixed effect is a mean effect (regression slope) around which the random effect of time varies, as measured by the standard error of the time slopes in the fixed effects table. By modeling time as a random effect, the standard error of the fixed effect time slopes are adjusted for clustering of scores by the level 2 grouping variable, empl\_id. Note that a random effect model adjusts standard errors and significance tests for clustering of score by individuals. It models between-subjects variance.
      2. We must specify a covariance type assumption for time as a random effect. Following common practice, this example assumes an unstructured (UN) covariance type, which allows both variances and covariances to be computed and reported (these are the "UN()" terms in the output section). The default diagonal structure does not do this.
      3. Time is also made a repeated measure on the initial "Subjects and Repeated" LMM screen in SPSS. This adjusts standard errors and significance tests for correlated residuals within the random effect, time. Knowing the score of subject 1 in time 1 helps predict their score in time 2. Declaring time as a repeated variable adjusts standard errors to control for this autocorrelation. Note that repeated measures adjusts standard errors and significance tests for patterned (hence non-independent) relationships across time, the repeated variable. Repeated measures models within-subject variance.



* + - *SPSS*. This model predicts score from time and quadratic time, where both of these fixed effects are adjusted for time as a random effect and time as a repeated measure; and adjusting for employee id as the subject (combination, grouping) variable for test scores. The SPSS menu settings for the repeated measures intercept + time model are ...
      1. On the initial "Subjects and Repeated" LMM screen, enter empl\_id as the "Subjects" variable; enter the time variable, test\_seq, as the "Repeated" variable; and accept "Diagonal" as the default repeated covariance type.
      2. On the main "Linear Mixed Models" dialog, enter score as the "Dependent" and enter test-seq and seq\_sq (and other continuous predictors, if any) as "Covariates".
      3. Click the Fixed button and enter test\_seq and seq\_sq in the "Model:" area; leave "Include intercept" checked; exit by clicking "Continue".
      4. Click the Random button and select "Unstructured" as the covariance type; enter test\_seq in the "Model:" area; make empl\_id the Subjects-Combinations variable; and check "Include intercept"; Continue.

SPSS syntax looks like this:

MIXED score WITH test\_seq seq\_sqr

/CRITERIA=CIN(95) MXITER(100) MXSTEP(10) SCORING(1) SINGULAR(0.000000000001)

HCONVERGE(0, ABSOLUTE) LCONVERGE(0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)

/FIXED=test\_seq seq\_sqr | SSTYPE(3)

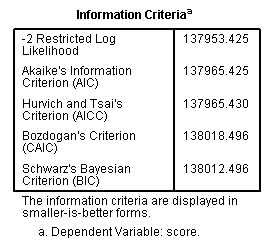
/METHOD=REML

/PRINT=SOLUTION TESTCOV

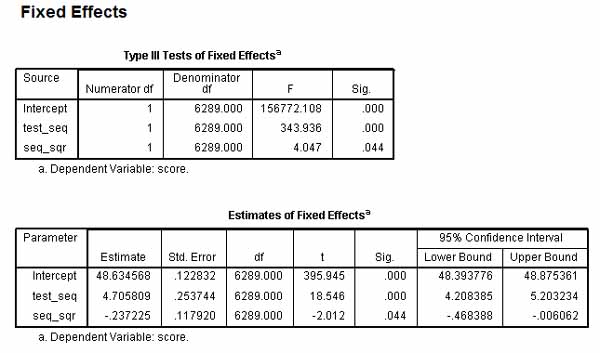
/RANDOM=INTERCEPT test\_seq | SUBJECT(empl\_id) COVTYPE(UN)

/REPEATED=test\_seq | SUBJECT(empl\_id) COVTYPE(DIAG).

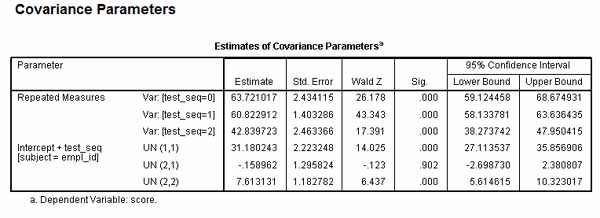
* + - *Information criteria*. Adding time as a repeated measure, not just a random effect, further improves model fit, reflected in a lower AIC.



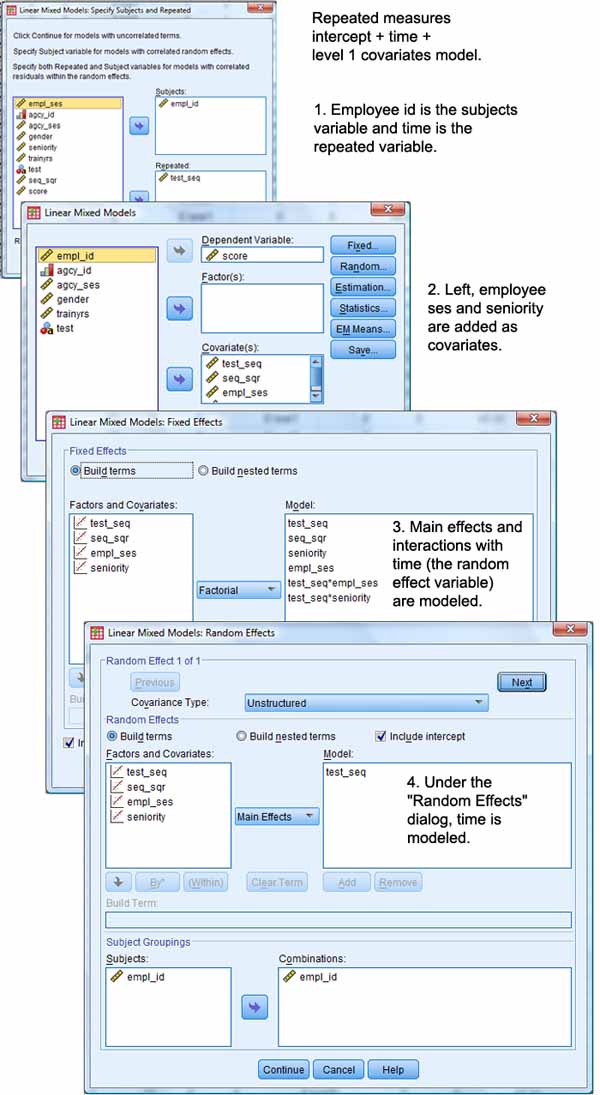
* + - *Fixed effects*. The actual estimates of the intercept for the dependent (score) and for the predictors (test\_seq, seq\_sq) do not change compared to the regression model or the random effects intercept-only and intercept + time models above. However, the standard errors for the intercept and for the time variable do change because they are adjusted for the random effect of time clustered by employee (making time a modeled random effect with empl\_id as subjects-combinations variable does this); and they are adjusted for autocorrelation across time (making time the repeated measure does this). Note the standard errors for the intercept and the slopes are all lower than for the unadjusted regression model. All are significant, even time-squared, whose larger standard error in the regression model and consequent non-significance misleadingly suggested no effect. LMM thus leads to different and more accurate inferences compared to regression models.



* + - *Random Effects*. The "Covariance Parameters" table now has two components: an upper "Repeated Measures" section, and the lower random effects section similar to that seen in previous models. With the addition of the repeated measures parameters, the single "Residual" parameter disappears.
      1. Repeated Measures. The exact format of this section varies depending on what covariance structure was assumed. For this example it is "Diagonal". The estimates here substitute for the "Residual" parameter in examples above, except we get one estimate per time period (per level of the repeated variable, time). These reflect the residual variance in score after taking into account between-subjects effects described below. As such they reflect within-subjects variance, which here is significant for each time period. We can say that at each time period there is significant residual variance to explain beyond the covariates thus far in the model.
      2. UN(1,1): as before, this is the between-subjects variance estimate for the error term for the intercepts of score. Here, mean employee score across the three test occasions does vary significantly between employees at the .000 level. As the intercept reflects unexplained between-subject variance, we can say there remains significant unexplained variability between employees in mean score across time. A more complex model with additional fixed covariates may be needed.
      3. UN(2,2). as before, this is the between-subjects variance estimate for the error terms for the slopes for the time variables. It is significant in this example, meaning the regression slopes for the time variables also vary significantly between employees.
      4. UN(2,1). this is the between-subjects estimate of covariance between intercepts and slopes. As it is non-significant, we can conclude that the error terms for the intercepts and for the slopes of the time variables do not covary significantly, hence there is no significant intercept-slope interaction.



* + **A repeated measures intercept-plus-time-plus-level 1 covariates model using SPSS**.
    - It is, of course, possible to add other level 1 covariates as predictors, seeking to improve model fit and to reduce unexplained within-subjects variance. In this model we .....
      1. Retain empl\_id as the "Subjects" variable, retain time-seq as the "Repeated" variable, and keep the repeated covariance structure as diagonal.
      2. We keep time and quadratic time (test\_seq and seq\_sq) as fixed covariates, but add employee socioeconomic status. (empl\_ses) and employee seniority as additional fixed factors. We retain an intercept for the fixed effects.
      3. Since we will have time as a random effect, we also add test\_seq\*empl\_ses and test\_seq\*seniority interactions to the fixed effects.
      4. We keep test\_seq as a random effect with an unstructured covariance type and we keep empl\_id as the subjects-combinations variable. We retain an intercept for the random effects.



The SPSS syntax looks like this:

MIXED score WITH test\_seq seq\_sqr empl\_ses seniority

/CRITERIA=CIN(95) MXITER(100) MXSTEP(10) SCORING(1) SINGULAR(0.000000000001)

HCONVERGE(0, ABSOLUTE) LCONVERGE(0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)

/FIXED=seq\_sqr seniority test\_seq empl\_ses test\_seq\*empl\_ses test\_seq\*seniority | SSTYPE(3)

/METHOD=REML

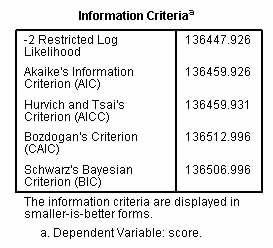
/PRINT=SOLUTION TESTCOV

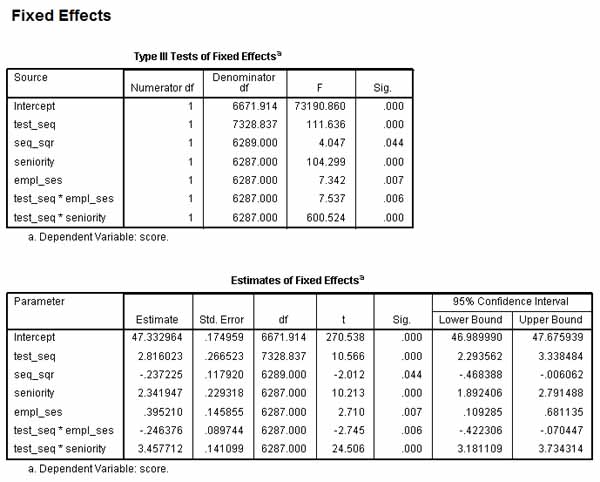
/RANDOM=INTERCEPT test\_seq | SUBJECT(empl\_id) COVTYPE(UN)

/REPEATED=test\_seq | SUBJECT(empl\_id) COVTYPE(DIAG).

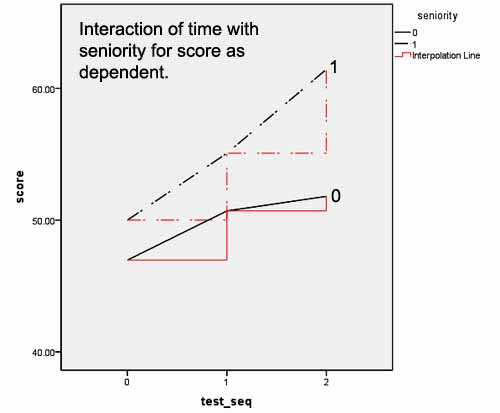
* + - *Information criteria.* As lower AIC is better fit, the repeated intercept + time +level 1 covariates model is shown to be an improvement. This is also true by BIC, which has a stronger adjustment for model parsimony.

|  |  |  |
| --- | --- | --- |
| MODEL | AIC | BIC |
| Intercept only | 138,111.38 | 138,127.07 |
| Intercept + time | 138,023.57 | 138,054.95 |
| Repeated intercept + time | 137,965.43 | 138,012.50 |
| Repeated intercept + time + level 1 covariates | 136,459.93 | 136,507.00 |

* + - 
    - *Fixed effects.* In the fixed effects table below, we see that time, seniority, and employee ses are each positive and significant. The interactions of seniority and employee ses with time are also significant. (In a different run, not illustrated, the interaction of seniority with employee ses was not significant). Note "significant" here means "significant controlling for other variables in the model". Estimated scores rise with time, with seniority, with employee ses. Scores increase with the interaction of time with seniority but decrease with the interaction of time and employee ses. Other observations can be made from the "Fixed Effects" table:
      1. The average growth rate in scores between testing intervals is the regression coefficient for test\_seq, the time variable: 2.82, controlling for other variables in the model.
      2. The intercept is the score when other variables in the model are controlled (held at 0). Thus, employees with no seniority (coded 0) and mean ses (coded 0 since empl\_ses is centered) on the first test occasion (test\_seq = 0) can be expected to have a score of 47.33.
      3. Since the estimates in the "Fixed Effects" table are regression slopes, an employee with seniority (coded 1) and mean ses (coded 0 since ses is centered) on the first test occasion (test\_seq = 0) can be expected to have a score of 47.33 + 2.34 = 49.67.
      4. On the second test occasion (test\_seq = 1), an employee with seniority (coded 1) and mean ses (coded 0) can be expected to have a score of 47.33 + 2.34 + 3.46 = 53.13 due to the interaction effect of time and seniority.
      5. That the test\_seq\*seniority interaction is positive and significant means that employees with seniority increase in score between test intervals more than employees without seniority. This can be seen visually in the graph below. Non-parallel or crossing lines indicate interaction.

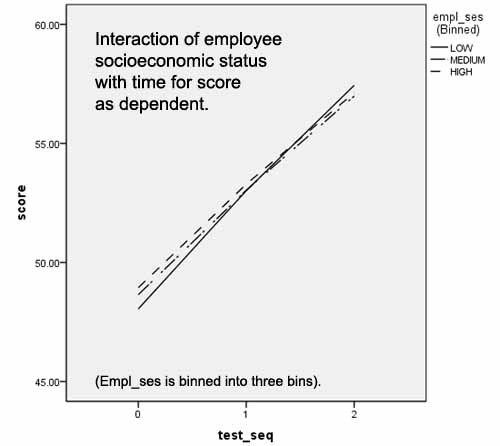


* + - 1. *Graphing interactions.* This graph is created by the SPSS syntax: GRAPH /LINE(MULTIPLE)=MEAN(score) BY test\_seq BY seniority.



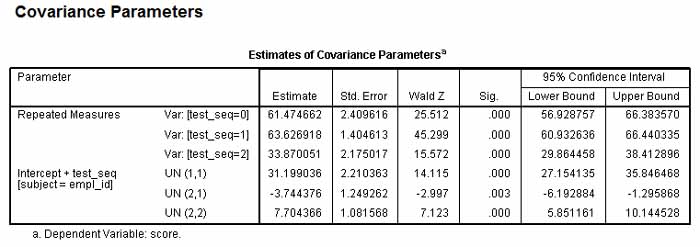
* + - 1. That the test\_seq\*empl\_ses interaction is negative and significant means that employees with higher socioeconomic status increase in score between test intervals at a lower rate compared to employees with mean ses (empl\_ses = 0, where ses is coded centered on 0). Note that the negative coefficient need not mean absolute decrease in score as the time effect is positive and large compared to the negative test\_seq\*empl\_ses interaction effect.

Viewed graphically after binning empl\_ses into high, medium, and low categories, this is seen in the graph below. However, although there is interaction (indicted by crossing lines) and although significant, we can see a near-parallel pattern. The large sample size enables a small interaction to be significant.

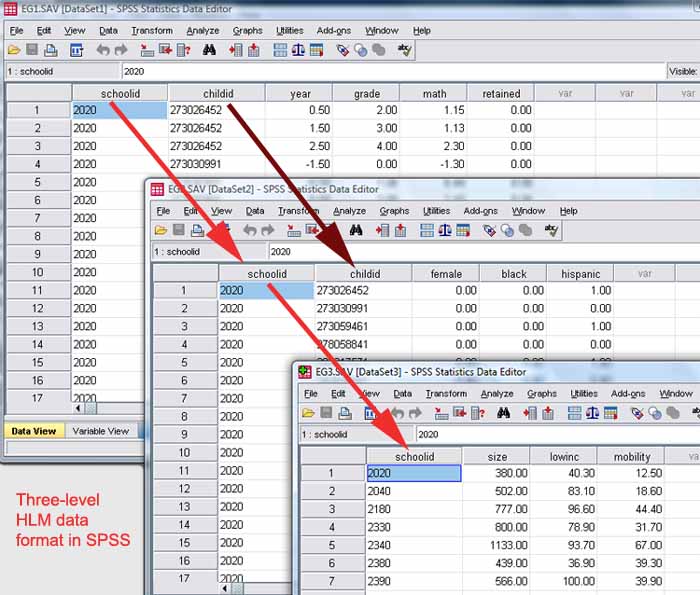


* + - 1. Graphing is not necessary (though visually helpful) because similar information is contained in the confidence limits. For any effect, main or interaction, the effect is significant if 0 is not withn the confidence limits. For the example data, although zero is not quite within the confidence limits for the time\*ses interaction (test\_seq\*empl\_ses) and for the quadratic time term (seq\_sq), both of which are significant due to large sample size (n = 18,870), this indicates effects which are only marginally different from no effect. The researcher may wish to drop these terms in spite of significance in order to create a more parsimonious model.
    - *Random effects* The estimates in this table partition population variability into within-subjects components (the "Repeated" estimates, which replace the "Residual" estimate in random effects models discussed above) and various between-subjects components (the "UN" terms in the unstructured covariance model below). Note that the estimates do not have a clear absolute meaning. Rather the covariance parameter estimates are used to understand the relative importance of different variance components in the model.

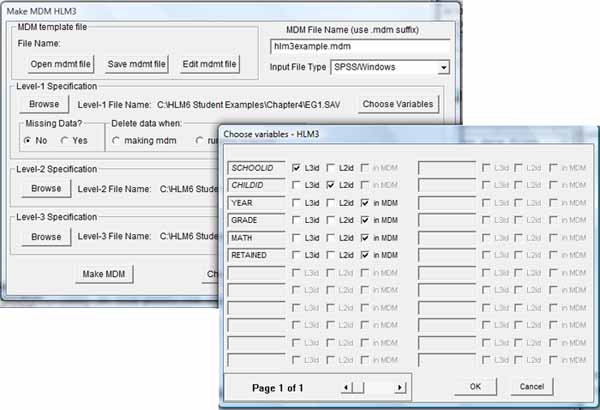
In this example, a"diagonal" covariance type is assumed for the repeated effects and an "unstructured" covariance type is assumed for the random effects. The diagonal type assumes independence of the correlated residuals between time periods, as discussed above. (Note that a variance components (VC) covariance structure type is not available for repeated measures models.) The unstructured type allows non-zero covariances among random effects. In the example, there are two random effects: the intercept and test\_seq (the time variable). This is reflected, of course, in the RANDOM clause of the SPSS syntax:    /RANDOM=INTERCEPT test\_seq | SUBJECT(empl\_id) COVTYPE(UN)



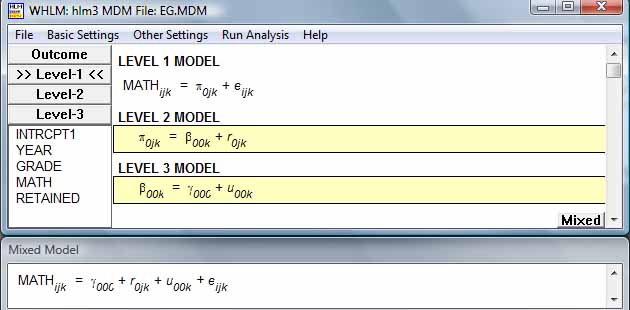
* + - * + Repeated Measures. There is one estimate per time period (per level of test\_seq). These reflect the residual variance in score across time after taking into account between-subjects effects described below. As such they reflect within-subjects variance, which here is significant for each time period. We can say that at each time period there is significant residual variance to explain beyond the covariates thus far in the model. This residual is lowest for the third test occasion, and this drop in the estimate for the third time period means that the covariates added to the model had the most effect on explaining residual variance for the third test occasion.
        + UN(1,1): this is the between-subjects variance estimate for the error term for the intercepts, which represent mean scores. Its significance test tests if mean scores vary significantly between employees. Mean employee scores across the three test occasions in this example do vary significantly between employees at the .000 level. This means that there is significant unexplained variability (reflected in variability of the error term for the intercepts) between employees in mean score across time. A more complex model with additional covariates may be needed.
        + UN(2,2): this is the between-subjects variance estimate for the slopes for the second random variable in the model, namely test\_seq, which is the time variable (the first is the intercept). It is here shown to be significant, meaning the linear regression slope for the time variable varies significantly between employees around the mean growth parameter (2.82, the b coefficient of test\_seq in the fixed effects table).
        + UN(2,1): This is the the estimate for the covariance between the second random effect (the time variable, test\_seq) and the first (the intercept), representing mean test score. Since significant for this example, the error terms for the intercepts and for the slopes of the time variable do covary significantly.
        + *Notation for unstructured models*. For unstructured covariance models, the UN(i,j) terms refer to the covariance between the ith and jth random effects. When i = j, the UN() term refers to variance rather than covariance.Thus UN(3,3) would be the between-subjects variance estimate for the slopes for a third random effect. UN(3,1) would be the estimate for the covariance between a third random effect and the first (the intercept). If significant, the error terms for the intercepts and for the slopes of the third random effect variable would covary significantly. Etc.
        + *Non-significant covariances*. Note that if the covariance estimates had been non-significant, the researcher might recalculate the model using a variance components (VC) covariance type assumption, since VC assumes no covariance between random effects.
    - *Using baseline covariates to measure change*. Additional covariates serve as control variables. In the current example, a baseline performance score, "prescore", could be added as a covariate fixed effect (with the same value for any given individual across time periods). By adding prescore to the covariate list and to the model under the Fixed Effects button, the researcher would be modeling change in performance score. The dependent becomes score controlling for prescore. If the prescore\*time interaction were also added (added to the Model list under the Fixed button but not added to the covariate list), then in the "Tests of Fixed Effects" table a significant positive prescore\*time interaction would indicate that as prescore increases so does the rate of increase in performance score (the dependent).
* **A three-level longitudinal null (intercept-only) model using HLM software.** 
  + *Example.* In this example, there are three levels: tests nested by student, with students nested within schools. The null model models the dependent variable, MATH math achievement scores, as a function of the grouping variables at levels 2 (students) and three (schools).
  + *Data set-up.* This example uses sample files provided in the HLM7 Student Examples/Chapter4/ folder provided with HLM software: EG1.SAV, EG2.SAV, and EG3.SAV. Each .SAV file contains the data for its own level, hence three .SAV files. There is and must be a common link field, in this case "schoolid", linking all three levels. That is, "schoolid" is the level 3 link variable. There is and must also be a second common link field linking levels 1 and 2. Here "childid" is the level 2 link variable.
    - *Sorting.* In HLM, the level-1 and level-2 files must be sorted in the same order of level-2 ID nested within level-3 ID. That is, in this case, the schoolid order in the level 3 file (EG3.SAV) must be the schoolid order in the level 1 and level 2 files (EG1.SAV and EG2.SAV). The childid order in the level 2 file must be the childid order in the level 1 file. Failure to sort properly will lead to incorrect results.
      1. Level 1 file (EG1.SAV): There are four level 1 variables on 7242 observations collected on 1721 children in 60 schools, with data from the end of grade one and then annually until grade six. Level 1 is thus time series data for each child on the four variables. Not every child is measured on every occasion (maximum is 6 measurement occasions), in which case there is no data row for that child/occasion. The four variables are:
         * YEAR: For the 6 years studied, YEAR is centered by subtracting 3.5 so it ranges from -2.5 to +2.5.
         * GRADE: Actual grade minus 1, so Grade 1 is 0, etc.
         * MATH: A math test in an IRT scale score metric:
         * Descriptive Statistics
         * N Minimum Maximum Mean Std. Deviation
         * math 7230 -5.22 5.77 -.5369 1.53470
         * Valid N (listwise) 7230
         * RETAINED: An binary indicator variable flagging if a child is retained in grade for a given year (1 = retained, 0 = not retained) rather than passed.
      2. Level 2 file (EG2.SAV): Three level 2 variables (black, hispanic, female), not counting the two link variables (schoolid, childid). All three variables are coded 0,1, where "1" indicates thr trait is present.
      3. Level 3 file (EG3.SAV: Three level 3 variables (size, lowinc, mobility), all at the interval level, describing school characteristics:
         * SIZE: number of enrolled students
         * LOWINC: Percent of students from low income families
         * MOBILE: Percent of students moving during the academic year
      4. Summary. Level 1 is the year-of-measurement level. Level 2 is the student level. Level 3 is the school level. We thus have years within students within schools nesting.



* + - *Making the MDM file.* After creating the needed three SPSS raw data files as described above, the researcher follows conversion steps similar to those for HLM2 illustrated [below](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#makemdm), except checking the HLM3 radio button. As in HLM2 for 2-level models, the researcher completes the MakeMDM dialog illustrated below, but this time indicating the level 3 as well as level 2 link variables. (The figure below shows variable selection for the level 1 file, EG1.SAV, but the researcher repeats the process for EG2.SAV and EG3.SAV.) When the researcher clicks the "Make MDM" button, a file called HLM3MDM.STS is created containing descriptive statistics on the data. The "Check Stats" button should be clicked to see if data look right and if the sample size seems low (this may be due to not sorting properly). After clicking "Done" on the "MAKE MDM - HLM3" dialog, the HLM working file is created under the name given by the researcher. From this point on, the .SAV files are no longer used.



* + *The null (unconditional) three-level model.* As in HLM2, a first step often is to create and run the null (intercept-only) model as a basis for comparison with later more complex models.
    - Creating the intercept-only model in HLM.
      1. Select File, Create a new model using an existing MDM file, then select the MDM file created from the SPSS .SAV files as discussed previously. For this example, provided with HLM software, select c:\HLM7 Student Examples\Chapter 4\EG.MDM.
      2. Then as illustrated below, in the Level 1 pane, select MATH as the outcome variable. Click "Mixed" to display the combined equation. In the intercept-only model, the intercept is modeled as a random coefficient. Specifically, at level 1 (across measurement periods), the dependent variable MATH is a function of the intercept and an error term. At level 2, the level 1 intercept is a function of the level 2 intercept and an error term. At level 3, the level 2 intercept is a function of the level 3 intercept and an error term. The level 1 intercepts are assumed to vary randomly across the level 2 units, which in this example are the 1,721 children (sets of measurement periods by child) within schools. The level 2 intercepts are assumed to vary randomly across the level 3 units, which here are the 60 schools.



* + - 1. From the HLM menu, select File, Save As, and assign a filename to save the command file just constructred. For example, save to "eg0.hlm".
      2. Select "Run Analysis" from the main HLM menu bar.
    - Output for the intercept-only model. Partial output is shown below, with comments in italics.
      1. *The data header* contains information on the count of cases at each level. This should be examined as part of verifyng the data were read in properly. Here there were 60 schools with 1,721 children and 7,230 measurements.
      2. Module: HLM3S.EXE
      3. The data source for this run = C:\HLM7 Student Examples\Chapter4\EG.MDM
      4. The command file for this run = C:\HLM7 Student Examples\Chapter4\EG0.hlm
      5. The maximum number of level-1 units = 7230
      6. The maximum number of level-2 units = 1721
      7. The maximum number of level-3 units = 60
      8. The outcome variable is MATH
      9. *The model* is then reproduced in print form.
      10. Summary of the model specified (in equation format)
      11. ---------------------------------------------------
      12. Level-1 Model
      13. Y = P0 + E
      14. Level-2 Model
      15. P0 = B00 + R0
      16. Level-3 Model
      17. B00 = G000 + U00
      18. *Sigma-squared and tau.* Sigma-squared represents the level 1 variance of the intercept. It is the variance of the level 1 error term, E. Tau-pi represents the level 2 variance of the intercept. It is the variance of the level 2 error term, R0, in the model specification in output above. As such it contributes to the estimated slope at level 1. Tau-beta represents the level 3 variance of the intercept. It is the variance of the level 3 error term, U00 in the model specification in output above. As such it contributes to the estimate of the slope at level 2. In more complex models, tau values are presented in a matrix, but in the null model there is only the intercept at each level and one tau value per level. Both sigma-squared and tau appear later in the variance components table as estimates of random effects.
      19. Sigma\_squared = 1.52393
      20. Standard Error of Sigma\_squared = 0.02900
      21. Tau(pi)
      22. INTRCPT1,P0 0.57038
      23. Tau(pi) (as correlations)
      24. INTRCPT1,P0 1.000
      25. Standard Errors of Tau(pi)
      26. INTRCPT1,P0 0.03338
      27. Tau(beta)
      28. INTRCPT1
      29. INTRCPT2,B00
      30. 0.31767
      31. Tau(beta) (as correlations)
      32. INTRCPT1/INTRCPT2,B00 1.000
      33. Standard Errors of Tau(beta)
      34. INTRCPT1
      35. INTRCPT2,B00
      36. 0.06636
      37. *Reliability.* The level 1 reliability is the reliability of the level 1 intercepts across the 1,721 children (recall children are the level 2 units). The larger the number of level 1 units per level 2 unit, and the larger the level 1 variance component (sigma-squared = 1.52) relative to the level 2 variance component (tau-pi = .57), the closer to 1.0 will be the reliability. In this example, the level 1 variance component is large relative to the level 2 variance component (1.52/.57 = 2.67) but there are not very many level 1 units per level 2 unit (7230/1721 = 4.2). The former ratio drives reliability up toward 1.0 and the latter ratio drives reliability down, resulting in overall moderate reliability of .604.

The level 2 reliability is the reliability of the level 2 intercepts across the 60 schools (recall schools are the level 3 units). Reliability will approach 1.0 as the ratio of the level 2 variance component is large relative to the level 3 variance component (tau-pi/tau-beta = .57/.32 = 1.78) and when the number of level 2 units is large relative to level 3 units adjusting for level 2 reliability ((1721 children/60 schools)\*.604 = 17.3). Here while the former ratio is lower than for level 1 reliability, the latter ratio is much larger, resulting in the stronger reliability estimate of .871.

----------------------------------------------------

Random level-1 coefficient Reliability estimate

----------------------------------------------------

INTRCPT1, P0 0.604

----------------------------------------------------

----------------------------------------------------

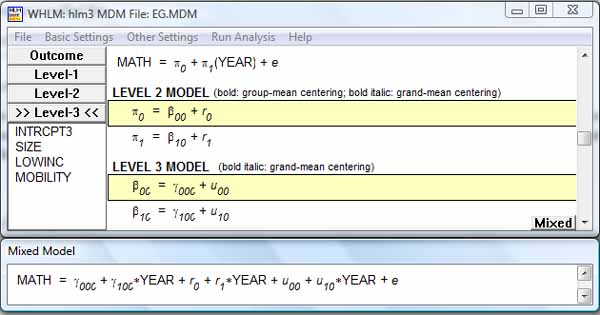
Random level-2 coefficient Reliability estimate

----------------------------------------------------

INTRCPT1/INTRCPT2, B00 0.871

----------------------------------------------------

* + - 1. *Fixed effects.* The estimated MATH intercept of -.51 is significant at < .001. The intercept represents the mean MATH score when no predictors are in the model.
      2. Final estimation of fixed effects:
      3. ----------------------------------------------------------------------------
      4. Standard Approx.
      5. Fixed Effect Coefficient Error T-ratio d.f. P-value
      6. ----------------------------------------------------------------------------
      7. For INTRCPT1, P0
      8. For INTRCPT2, B00
      9. INTRCPT3, G000 -0.510018 0.077970 -6.541 59 0.000
      10. ----------------------------------------------------------------------------
      11. Final estimation of fixed effects
      12. (with robust standard errors)
      13. ----------------------------------------------------------------------------
      14. Standard Approx.
      15. Fixed Effect Coefficient Error T-ratio d.f. P-value
      16. ----------------------------------------------------------------------------
      17. For INTRCPT1, P0
      18. For INTRCPT2, B00
      19. INTRCPT3, G000 -0.510018 0.077964 -6.542 59 0.000
      20. ----------------------------------------------------------------------------
      21. *Random effects*. Below, random effect variance components are shown. Total variance sums to 2.41 (sigma-squared plus tau-pi plus tau-beta = 1.524 + .570 + .318 = 2.412.)
          * The level 1 variance component, E, is 1.52. This is the largest variance component, being (63% of the total), showing that the greatest variance in MATH scores is at level 1, which means by year within children within schools. The variance in MATH score attributable to variance in YEAR (the level 1 unit) is thus 1.524/2.412 = 63.2%. Most of the variance is due to children within schools improving MATH score from measurement year to measurement year.
          * The level 2 variance component is .57. This component is 24% of the total and reflects variance in MATH scores between children within schools. The variance in score attributable to variation at level 2 (the child level) is .570/2.412 = 23.6%. This is the child effect.
          * The level 3 variance component is .32. This is the smallest variance component (13% = [.318/2.412 = 13.2%]) of the total), showing that between-schools variance (the level 3 component) explains the least variance in MATH. This is the school effect. Since schools explains relatively little of the total variance in MATH, a more complex model with predictor variables beyond the intercept seems warranted.
          * Final estimation of level-1 and level-2 variance components:
          * ------------------------------------------------------------------------------
          * Random Effect Standard Variance df Chi-square P-value
          * Deviation Component
          * ------------------------------------------------------------------------------
          * INTRCPT1, R0 0.75524 0.57038 1661 4253.88860 0.000
          * level-1, E 1.23447 1.52393
          * ------------------------------------------------------------------------------
          * Final estimation of level-3 variance components:
          * ------------------------------------------------------------------------------
          * Random Effect Standard Variance df Chi-square P-value
          * Deviation Component
          * ------------------------------------------------------------------------------
          * INTRCPT1/INTRCPT2, U00 0.56362 0.31767 59 573.08980 0.000
          * ------------------------------------------------------------------------------
          * *Likelihood ratio test.* The deviance is a measure of model fit, with lower being better fit. The intercept-only (null model) deviance is compared with the deviance in other more complex models, with lower deviance being better fit. The likelihood ratio test will test the significance of the difference in deviance values between models. For the null model, however, the deviance is the baseline measure and no likelihood ratio test is appropriate.
          * Statistics for current covariance components model
          * --------------------------------------------------
          * Deviance = 25305.980829
          * Number of estimated parameters = 4
  + **A three-level unconditional linear growth model using HLM software.**
    - *Example.* This is a three-level random coefficients model very similar the three-level longitudinal null model discussed above, except YEAR (entered uncentered) is added as a level 1 covariate and its slope is modeled by grouping at higher levels. That is, YEAR as well as the level 1 intercept of MATH are modeled as random effects. As such, this is a "linear growth model" because it is based on linearly increasing YEAR, and is unconditional because not conditioned on additional predictors.
    - *The model.* In this model, illustrated below, the level 1 outcome variable MATH is a function of a level 1 intercept, a slope times YEAR, and a random error term. The level 1 intercept is a function of a level 2 intercept and a random error term. The level 1 slope for YEAR is also a function of a level 2 intercept and a random error term. Finally, the level 2 intercepts are in turn functions of level 3 intercepts and error terms. (The equations can be displayed in more detail by selecting File, Preferences, and checking the "Use level subscripts" checkbox). Note that all error terms are toggled "on".



The model is shown in printed form at the top of the HLM output below.

Module: HLM3S.EXE

The data source for this run = C:\HLM7 Student Examples\Chapter4\EG.MDM

...

The outcome variable is MATH

Summary of the model specified (in equation format)

---------------------------------------------------

Level-1 Model

Y = P0 + P1\*(YEAR) + E

Level-2 Model

P0 = B00 + R0

P1 = B10 + R1

Level-3 Model

B00 = G000 + U00

B10 = G100 + U10

* + - *Sigma-squared and tau.* Below, sigma-squared is level 1 variance within level 2 units (children). In the intercept-only model sigma-squared was 1.52, but now it has dropped to .30, a relatively small magnitude. That is, sigma-squared is the variance in the level 1 error term once YEAR is placed in the level 1 equation. The variance of MATH by observation YEAR is small within the measures for any given child within any given school, once the linear effect of measurement YEAR is controlled.

Tau-pi is now a 2-by-2 matrix because there are two random effects modeled at level 2: the level 1 intercept and YEAR. The variance associated with the level 1 intercept (.64) is much greater than the variance associated with YEAR (.01). The covariance of intercept error with error for slope of YEAR was .05 in the level 2 model.

Tau-beta is also a 2-by-2 matrix because it too models two random effects: level 3 models the two level 2 intercepts - one for the level 1 intercept and one for the level 1 slope of YEAR. Again, the variation associated with the intercepts is much greater than that for YEAR. Here at level 3, the correlation of intercepts and slopes is moderate (.399) and weaker than at level 2 (where it was .055). That is, the tendency of high average MATH score to be associated with a high effect of measurement YEAR was more pronounced at the child level (level 2) than at the school level (level 3). Tau and pi values appear in the variance components output for random effects, discussed below.

Sigma\_squared = 0.30148

Standard Error of Sigma\_squared = 0.00660

Tau(pi)

INTRCPT1,P0 0.64049 0.04676

YEAR,P1 0.04676 0.01122

Tau(pi) (as correlations)

INTRCPT1,P0 1.000 0.551

YEAR,P1 0.551 1.000

Tau(beta)

INTRCPT1 YEAR

INTRCPT2,B00 INTRCPT2,B10

0.16531 0.01705

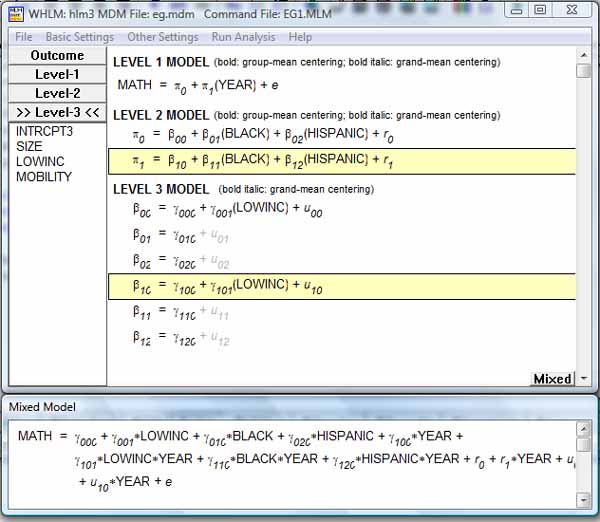
0.01705 0.01102

Tau(beta) (as correlations)

INTRCPT1/INTRCPT2,B00 1.000 0.399

YEAR/INTRCPT2,B10 0.399 1.000

* + - *Reliability.* Reliability coefficients for each level are output, interpreted as in previous models.
    - ----------------------------------------------------
    - Random level-1 coefficient Reliability estimate
    - ----------------------------------------------------
    - INTRCPT1, P0 0.839
    - YEAR, P1 0.190
    - ----------------------------------------------------
    - ----------------------------------------------------
    - Random level-2 coefficient Reliability estimate
    - ----------------------------------------------------
    - INTRCPT1/INTRCPT2, B00 0.821
    - YEAR/INTRCPT2, B10 0.786
    - ----------------------------------------------------
    - *Fixed effects.* The effect of YEAR on MATH score is significant at the .000 level. For each unit increase in measurement YEAR, MATH increases .763.
    - Final estimation of fixed effects:
    - ----------------------------------------------------------------------------
    - Standard Approx.
    - Fixed Effect Coefficient Error T-ratio d.f. P-value
    - ----------------------------------------------------------------------------
    - For INTRCPT1, P0
    - For INTRCPT2, B00
    - INTRCPT3, G000 -0.779309 0.057829 -13.476 59 0.000
    - For YEAR slope, P1
    - For INTRCPT2, B10
    - INTRCPT3, G100 0.763029 0.015263 49.993 59 0.000
    - ----------------------------------------------------------------------------
    - Final estimation of fixed effects
    - (with robust standard errors)
    - ----------------------------------------------------------------------------
    - Standard Approx.
    - Fixed Effect Coefficient Error T-ratio d.f. P-value
    - ----------------------------------------------------------------------------
    - For INTRCPT1, P0
    - For INTRCPT2, B00
    - INTRCPT3, G000 -0.779309 0.057830 -13.476 59 0.000
    - For YEAR slope, P1
    - For INTRCPT2, B10
    - INTRCPT3, G100 0.763029 0.015260 50.000 59 0.000
    - ----------------------------------------------------------------------------
    - *Random effects.* The following observations may be made based on the variance components table below:
      1. The variance component for the level 1 intercept, with a value of .64, is the largest component and is significant. The higher levels (child and school) have a significant effect on mean MATH score, even after controlling for YEAR.
      2. The random effect on the slope of YEAR is significant but quite small (.01). The child and school effects on the slope of YEAR are minor.
      3. The variance component for the residual level 1 error term is appreciable (.30). There there is significant residual variance in MATH scores within children within schools, indicating the need for additional variables in the model.
      4. At level 3, the variance component for the intercept(.16) is much larger than for YEAR (.01), again suggesting the need for additional variables in the model.
      5. *Calculating the effect of adding YEAR to the null model:* Sigma-squared (the residual error E term in the variance components table below) for the intercept-only model was 1.52, representing the residual variance of the level 1 error term when there were no predictors but accounting for nesting at child and school levels. Now for the unconditional growth model, it is .30. The reduction is 1.22. The reduction divided by the intercept-only sigma-squared = 1.22/1.52 = .802. That is, 80.2% of the level 1 residual variance in the intercept-only model is accounted for by adding YEAR to the level 1 equation.
      6. *Variance proportions.* The remaining variation in MATH score after linear effects of YEAR at level 1 are controlled are .302 for level 1 (within measurement years), .640 for level 2 (between children), and .165 for level 3 (between schools). This is total variance of 1.107. Dividing the level variance component by total gives variance proportions of .272 for level 1, .578 for level 2, and .149 for level 3. If this were a variance components (VC) or diagonal (DIAG) model, without covariance of random effects, we could say that the variance in MATH score after the linear effect of measurement YEAR is controlled is 27.2% associated with nonlinear and residual effects of YEAR, the level 1 unit; 57.8% associated with variation among children within schools; and 14.9% associated with between-school variation, in this model. As HLM software assumes an unstructured rather than VC or DIAG covariance type (unless the "diagonalize tau" option is selected), such a calculation is at best a "ballpark estimate" for this example due to covariance of random effects. One could, of course, compute the model with and without diagonalizing tau and examine differences in model fit, which, if trivial, would justify the foregoing calculation.
    - Final estimation of level-1 and level-2 variance components:
    - ------------------------------------------------------------------------------
    - Random Effect Standard Variance df Chi-square P-value
    - Deviation Component
    - ------------------------------------------------------------------------------
    - INTRCPT1, R0 0.80030 0.64049 1661 13679.62589 0.000
    - YEAR slope, R1 0.10595 0.01122 1661 2132.50756 0.000
    - level-1, E 0.54907 0.30148
    - ------------------------------------------------------------------------------
    - Final estimation of level-3 variance components:
    - ------------------------------------------------------------------------------
    - Random Effect Standard Variance df Chi-square P-value
    - Deviation Component
    - ------------------------------------------------------------------------------
    - INTRCPT1/INTRCPT2, U00 0.40658 0.16531 59 488.30922 0.000
    - YEAR/INTRCPT2, U10 0.10498 0.01102 59 377.43020 0.000
    - ------------------------------------------------------------------------------
    - *Likelihood ratio test.* Deviance is reduced from 25306 in the intercept-only model to 16326 in the unconditional linear growth model here. By the likelihood ratio test (model chi-square test), this difference is significant at the .000 level, confirming that adding YEAR to the model leads to better fit to the data.
    - Statistics for current covariance components model
    - --------------------------------------------------
    - Deviance = 16326.231292
    - Number of estimated parameters = 9
    - Model comparison test
    - -----------------------------------
    - Chi-square statistic = 8979.74871
    - Number of degrees of freedom = 5
    - P-value = 0.000
  + **A three-level conditional linear growth model.**
    - *Example.* This is a more complex random coefficients model which retains YEAR as a level 1 predictor of MATH, but adds additional predictors: BLACK and HISPANIC from the children level (level 2) and LOWINC from the school level (level 3).
    - *The model.* In this model, illustrated below, the level 1 outcome variable MATH is a function of a level 1 intercept, a slope times YEAR, and an error term. The level 1 intercept is a function of a level 2 intercept and an error term. The level 1 slope for YEAR is also a function of a level 2 intercept and an error term. Finally, the level 2 intercepts are in turn function of level 3 intercepts and error terms. (The equations can be displayed in more detail by selecting File, Preferences, and checking the "Use level subscripts" checkbox). Note that all error terms are toggled "on", as described above in the section on HLM2.



Module: HLM3S.EXE

The outcome variable is MATH

Summary of the model specified (in equation format)

---------------------------------------------------

Level-1 Model

Y = P0 + P1\*(YEAR) + E

Level-2 Model

P0 = B00 + B01\*(BLACK) + B02\*(HISPANIC) + R0

P1 = B10 + B11\*(BLACK) + B12\*(HISPANIC) + R1

Level-3 Model

B00 = G000 + G001(LOWINC) + U00

B01 = G010

B02 = G020

B10 = G100 + G101(LOWINC) + U10

B11 = G110

B12 = G120

* + - *Sigma-squared and tau.* Sigma-squared and tau values are variance components estimates and appear again the variance components table discussed below. Sigma-squared reflects the level 1 variance in the intercept, which in turn is the mean value of MATH when predictors are 0.

Tau-pi is is a variance/covariance matrix for the two random effects modeled at level 2: the level 1 intercept and YEAR. The variance associated with the level 1 intercept (.62) is still much greater than the variance associated with YEAR (.01). The covariance of intercept and slope of YEAR error was .05 in the level 2 model.

Tau-beta is also variance-covariance matrix, but as modeled at level 3 for the intercept and YEAR intercepts at level 2. Again, the variation associated with the intercepts (.078) is much greater than that for the level 2 YEAR intercept (.008). Here at level 3, the correlation of intercept and slope error is weak (.033) compared to level 2 (where it was .561), indicating that the tendency of high average MATH score to be associated with a high effect of measurement YEAR was much more pronounced at the child level (level 2) than at the school level (level 3)

Sigma\_squared = 0.30162

Standard Error of Sigma\_squared = 0.00660

Tau(pi)

INTRCPT1,P0 0.62231 0.04657

YEAR,P1 0.04657 0.01106

Tau(pi) (as correlations)

INTRCPT1,P0 1.000 0.561

YEAR,P1 0.561 1.000

Standard Errors of Tau(pi)

INTRCPT1,P0 0.02451 0.00491

YEAR,P1 0.00491 0.00196

Tau(beta)

INTRCPT1 YEAR

INTRCPT2,B00 INTRCPT2,B10

0.07808 0.00082

0.00082 0.00798

Tau(beta) (as correlations)

INTRCPT1/INTRCPT2,B00 1.000 0.033

YEAR/INTRCPT2,B10 0.033 1.000

Standard Errors of Tau(beta)

INTRCPT1 YEAR

INTRCPT2,B00 INTRCPT2,B10

0.01991 0.00441

0.00441 0.00194

* + - *Reliability.* The reliability for the intercept is much greater than for the slope of YEAR. The higher the reliability, the less discrepancy between OLS and LMM estimates for the given coefficient.
    - ----------------------------------------------------
    - Random level-1 coefficient Reliability estimate
    - ----------------------------------------------------
    - INTRCPT1, P0 0.835
    - YEAR, P1 0.188
    - ----------------------------------------------------
    - ----------------------------------------------------
    - Random level-2 coefficient Reliability estimate
    - ----------------------------------------------------
    - INTRCPT1/INTRCPT2, B00 0.702
    - YEAR/INTRCPT2, B10 0.735
    - ----------------------------------------------------
    - *Fixed effects.* The slope of YEAR is .87, meaning that as measurement YEAR increases 1 unit, MATH increases by .87 - a significant average growth rate. At level 3, which is the school level, BLACK and HISPANIC both significantly affect the intercept term, which reflects average MATH score when other variables in the model are controlled. At the school level, BLACK and HISPANIC do not significantly affect the slope term for YEAR, which reflects the rate of increase in MATH score across measurement YEARs. In contrast, LOWINC significantly affects both intercept and slope (both average score and rate of increase).
    - Final estimation of fixed effects
    - (with robust standard errors)
    - ----------------------------------------------------------------------------
    - Standard Approx.
    - Fixed Effect Coefficient Error T-ratio d.f. P-value
    - ----------------------------------------------------------------------------
    - For INTRCPT1, P0
    - For INTRCPT2, B00
    - INTRCPT3, G000 0.140628 0.113814 1.236 58 0.222
    - LOWINC, G001 -0.007578 0.001396 -5.428 58 0.000
    - For BLACK, B01
    - INTRCPT3, G010 -0.502091 0.076842 -6.534 1718 0.000
    - For HISPANIC, B02
    - INTRCPT3, G020 -0.319381 0.081918 -3.899 1718 0.000
    - For YEAR slope, P1
    - For INTRCPT2, B10
    - INTRCPT3, G100 0.874501 0.037287 23.453 58 0.000
    - LOWINC, G101 -0.001369 0.000499 -2.744 58 0.009
    - For BLACK, B11
    - INTRCPT3, G110 -0.030918 0.022274 -1.388 1718 0.165
    - For HISPANIC, B12
    - INTRCPT3, G120 0.043085 0.024368 1.768 1718 0.077
    - ----------------------------------------------------------------------------
    - *Random effects.* The variance components at the school level (level 3) are significant for both the intercept and YEAR terms when YEAR is 0 (recall YEAR was centered, so 0 corresponds to measurement YEAR 3.5). The table below reflects the effects of the school level on intercepts at level 2, which in turn affect the level 1 intercept and slope of YEAR. Differences among schools significantly affect both the intercept (mean MATH score) and the slope of YEAR (rate of score increase over measurement periods). The level 3 variance components are controlling for LOWINC and are lower as a result compared to the unconditional linear growth model previously discussed, where the intercept component was .165 and the YEAR component was .011 .
      1. *Calculating the effect of adding BLACK, HISPANIC, and LOWINC to the unconditional linear growth model*. Sigma-squared (E, within-groups residual error, in the table below) for the unconditional growth model was 0.30, representing the variance of the level 1 error term when only the linear effect of YEAR was controlled. Now for the conditional growth model, it is still .30. The reduction is close to zero. That is, none of the level 1 variance remaining in the unconditional growth model is accounted for by adding BLACK, HISPANIC, and LOWINC to the model. Note that in a full analysis, variables would have been added one at a time in each subsequent model.
      2. Final estimation of level-1 and level-2 variance components:
      3. ------------------------------------------------------------------------------
      4. Random Effect Standard Variance df Chi-square P-value
      5. Deviation Component
      6. ------------------------------------------------------------------------------
      7. INTRCPT1, R0 0.78886 0.62231 1659 13364.57298 0.000
      8. YEAR slope, R1 0.10518 0.01106 1659 2126.73092 0.000
      9. level-1, E 0.54920 0.30162
      10. ------------------------------------------------------------------------------
      11. Final estimation of level-3 variance components:
      12. ------------------------------------------------------------------------------
      13. Random Effect Standard Variance df Chi-square P-value
      14. Deviation Component
      15. ------------------------------------------------------------------------------
      16. INTRCPT1/INTRCPT2, U00 0.27943 0.07808 58 254.96395 0.000
      17. YEAR/INTRCPT2, U10 0.08935 0.00798 58 277.26967 0.000
      18. ------------------------------------------------------------------------------
      19. *Likelihood ratio test.* Below, the likelihood ratio test shows the conditional linear growth model to be significantly better than the null model. For the intercept-only three-level model, deviance was 25306. For the unconditional growth model, with only YEAR as a level 1 predictor, deviance was 16326. In the current conditional model, deviance is lower still at 16239. Lower deviance is better fit. Adding HISPANIC, BLACK, and LOWINC to the model has only improved fit marginally. Nonetheless, a separate likelihood ratio test, not shown, shows that the improvement is significant.
      20. Statistics for current covariance components model
      21. --------------------------------------------------
      22. Deviance = 16239.207232
      23. Number of estimated parameters = 15
      24. Model comparison test
      25. -----------------------------------
      26. Chi-square statistic = 9066.77360
      27. Number of degrees of freedom = 11
      28. P-value = 0.000

**Multivariate Models**

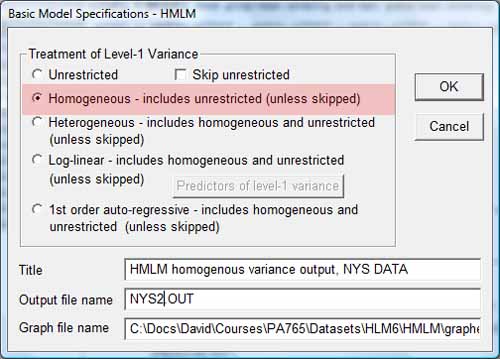
* + **Overview.**
    - *HLM software supports multivariate linear mixed modeling*. "Multivariate" in this context refers to models having more than one dependent variable. HLM software supports two modules for multivariate modeling: (1) HMLM (one-level multivariate normal models) and HMLM2 (multiple level 1 outcomes where level 1 units are nested within higher-level units). The models supported by these multivariate modules are essentially similar to univariate models discussed above. It is important to note, however, that the null model (the "unrestricted model") cannot be meaningfully defined in HMLM2 and thus likelihood ratio tests cannot be applied.
    - *Applicability to repeated measures designs*. Multivariate HLM may be applied to repeated measures data. For example, students nested within schools may be measured at T observation times. Level 1 would be the observations, with multiple observation data rows per student. Level 2 would be the students and their associated variables such as gender. In such a model, some of the T observations may be missing if missing at random (so missingness does not alter the covariance matrix), meaning not all students would have T data rows.
  + **Example: National Youth Survey data**.
    - *Source.* The example below is adapted from the NYS.MDM example file and related files supplied with HLM software. In these data, eleven-year-olds were interviewed starting in 1976 and continuing until 1980, for a total of five consecutive years (T = 5). Thus there are five data rows per student ID. Multivariate HLM is appropriate because the dependent variable, which was attitude toward deviant behavior (ATTID), has five values. In total, there were 1,079 observations of 239 students. Level 1 is the 1,079 data rows on observations and level 2 is the 239 data rows on the students themselves.
    - *Creating the latent dependent variable.* Let Y = ATTIT. Let Y1 = ATTIT in year 1, Y2 = ATTIT in year 2, etc. In the data file, ATTIT is the attitude score for any given student for any given year. But in the model, Y = ATTIT is a latent variable measured by the five observed dependent variables Y1 through Y5. In the data file, Y1 through Y5 do not appear. Rather, when the researcher lets HLM know ATTIT is the dependent variable and that (student) id is the ID variable, and when HLM discovers 5 data rows per student, it does the behind-the-scenes work of creating Y1 through Y5 as indicators of Y, which is ATTIT in the model. That is, HMLM combines the five observed dependent variables into one latent dependent variable for purposes of modeling.
    - *Variables.* The following variables are used:

Level 1 variables

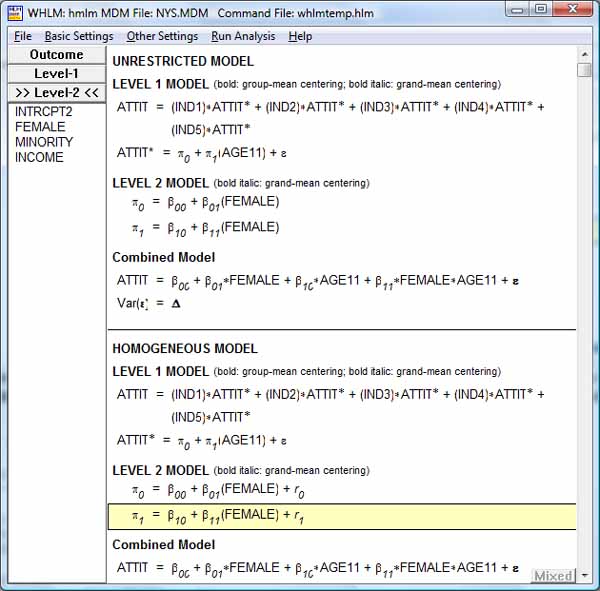
* + - 1. ID, the student id.
      2. ATTIT is the dependent variable, a scale of attitudes toward nine deviant behaviors (ex., property damage, marijuana use, alcohol use). Values ranged from 0.0 to 1.24, with a mean of .33. Higher values meant more tolerance of deviant behavior. Males had higher values than women. Students in later observation years had higher values than in initial years.
      3. AGE11 is student age at time T minus 11 years. Thus if age was 11, AGE11 was 0. As each student has 5 data rows, AGE11 will be 0, 1, 2, 3, and 4 sequentially in these rows, assuming the student is observed in all five years. If not observed, the row for that year will be missing.
      4. AGE11S is the value of AGE11 squared.
      5. AGE13 is the same as AGE111, except subtracting 13. Thus the five observation rows for student id=1 will be -2, -1, 0, 1, and 2.
      6. AGE13S is the value of AGE13 squared.
      7. EXPO is an index of exposure to deviant peers (students were asked how many of their friends engaged in the nine deviant behaviors from the ATTIT scale). This variable is standardized around a mean of 0, ranging from -.37 to +1.04
      8. IND1 to IND5 are indicator variables indicating whether a given measurement was available at a given time or not. Student 1's first data row would have values of 1,0,0,0,0 on IND1 through IND5 respectively; the second data row would be 0,1,0,0,0; etc. However, if a student was absent for a given year, the data row for that student for that year would be absent from the dataset and the student would have only four data rows, assuming present for the other four years.

Level 2 variables

* + - 1. ID, student id with one row per student
      2. FEMALE, coded 1 = female, 0 = male (117 females, 122 males)
      3. MINORITY, codes 1 = minority, 0 = other (50 minority, 189 other)
      4. INCOME, coded 1 to 10
  + **The unrestricted (null) multivariate linear mixed model.**
    - *The model.* This model, which is shown from HMLM output below, does not restrict variances of deviant behaviors attitudes to be equal across students. The next section will present a homogenous variances model for comparison. Unrestricted models have more parameters and are less parsimonious than homogenous variances models. The model. Y is ATTIT, the dependent variable. Y is modeled at level 1 as a function of the IND indicator variables and as a function of AGE11. The level 1 intercept for ATTIT (which is labeled P0) is modeled as a function of the level 2 variable, FEMALE. Likewise, the level 1 slope of AGE11 (labeled P1) is modeled as a function of the level 2 variable, FEMALE. The "Var(E) = D" expression simply means that the variance of Y\* equals D and that the complete covariance matrix will be estimated. This model assumes multivariate normally distributed error with a mean of 0.
    - Level-1 Model
    - Y = IND1\*Y1\* + IND2\*Y2\* + IND3\*Y3\* + IND4\*Y4\* + IND5\*Y5\*
    - Y\* = P0 + P1\*(AGE11) + e
    - Level-2 Model
    - P0 = B00 + B01\*(FEMALE)
    - P1 = B10 + B11\*(FEMALE)
    - Var(E) = D
    - *Variance, correlation and standard error matrices*. After applying an iterative algorithm (8 iterations for this example), HMLM outpurs the D matrix, representing estimated variances; then the D matrix represented as estimated, model-implied correlations; then the matrix of estimated standard errors. For most researchers, this is all intermediate output of interest only to spot anomalies. For instance, it is slightly anomalous that ATTIT in year 1 is more correlated with ATTIT in year 5 (r=.430) than it is with closer years.
      1. *Reading the D matrix.* The five rows and five columns represent the five observation years in this example. The diagonal in the D matrix is the variance of ATTIT for that year (ex., the variance of ATTIT in observation year 2 is .044). The off-diagonal coefficients are the covariances between any pair of years, with the upper triangle mirroring the lower. The "D (as correlations)" matrix represents the standardized covariances, which are correlations. The correlation of any year with itself is, of course, shown to be 1.0. For these data it can be seen that the correlations between later years is higher than between earlier observation years.
      2. \*\*\*\*\*\*\* ITERATION 8 \*\*\*\*\*\*\*
      3. D
      4. IND1 0.03446 0.01600 0.01803 0.02030 0.02365
      5. IND2 0.01600 0.04364 0.02667 0.02315 0.02569
      6. IND3 0.01803 0.02667 0.07147 0.05134 0.04640
      7. IND4 0.02030 0.02315 0.05134 0.08358 0.06428
      8. IND5 0.02365 0.02569 0.04640 0.06428 0.08780
      9. D (as correlations)
      10. IND1 1.000 0.412 0.363 0.378 0.430
      11. IND2 0.412 1.000 0.478 0.383 0.415
      12. IND3 0.363 0.478 1.000 0.664 0.586
      13. IND4 0.378 0.383 0.664 1.000 0.750
      14. IND5 0.430 0.415 0.586 0.750 1.000
      15. Standard Errors of D
      16. IND1 0.00341 0.00298 0.00368 0.00403 0.00420
      17. IND2 0.00298 0.00425 0.00420 0.00444 0.00461
      18. IND3 0.00368 0.00420 0.00666 0.00616 0.00611
      19. IND4 0.00403 0.00444 0.00616 0.00791 0.00716
      20. IND5 0.00420 0.00461 0.00611 0.00716 0.00834
    - *Fixed effects*. In the fixed effects table below, the following observations may be noted:
      1. That the "INTRCPT2, B00" coefficient is significant at the .000 level means that estimated mean ATTIT is different from 0.0 after controlling for other variables in the model. The actual mean ATTIT (from Analyze, Descriptive statistics procedure, not shown) was .33 on a scale that went from 0.0 to 1.24. Controlling for other variables in the model, which means at AGE11=0=age 11 and FEMALE=0=male, mean ATTIT was .23.
      2. That "FEMALE, B01" is significant at the .028 level means that the level 2 variable, FEMALE, has a significant effect on the intercept of ATTIT at level 1. When FEMALE increases 1 unit (corresponding to replacing a male with a female), mean ATTIT decreases by .05 units, controlling for AGE11 at AGE11=0=age 11.
      3. That "INTRCPT2,B10" is significant at the .000 level means that the slope of AGE11 differs from 0.0 after controlling for other variables in the model (for FEMALE). When FEMALE=0=male, the mean slope of AGE11 is estimated to be .065. (not part of HMLM output, this compares to a slope of .064 in OLS regression on males and females combined).
      4. That "FEMALE, B11" is not significant (p = .242) means that the level 2 variable, FEMALE, does not have a significant effect on the slope of AGE11. When FEMALE increases 1 unit (corresponding to replacing a male with a female), the slope of FEMALE as a predictor of the level 1 slope of AGE11 changes by only -.011, which is not significant.
      5. *In summary*, in this unrestricted model, age is related to attitude toward deviant behavior, controlling for sex. Sex does not affect how strongly age is related to deviant behavior attitude, but sex does affect mean attitude scores.
    - The value of the likelihood function at iteration 8 = 1.931605E+002
    - The outcome variable is ATTIT
    - Final estimation of fixed effects:
    - Standard Approx.
    - Fixed Effect Coefficient Error T-ratio d.f. P-value
    - ----------------------------------------------------------------------------
    - For INTRCPT1, P0
    - INTRCPT2, B00 0.226829 0.015782 14.373 237 0.000
    - FEMALE, B01 -0.049811 0.022675 -2.197 237 0.028
    - For AGE11 slope, P1
    - INTRCPT2, B10 0.064633 0.006478 9.978 237 0.000
    - FEMALE, B11 -0.011009 0.009409 -1.170 237 0.242
    - *Random effects.* The log likelihood function (LL) printed above the fixed effects table is directly related to the deviance. Deviance = -2LL. The base deviance, -386.32097, is used later to compare with the -2LL's for later, more complex models in a likelihood ratio test of the significance of the difference between nested models.
    - Statistics for current covariance components model
    - --------------------------------------------------
    - Deviance = -386.32097
    - Number of estimated parameters = 19
  + **The homogenous variance multivariate linear mixed model**
    - *Example.* In this model, level 1 variances are constrained to be equal. The model generates results identical to what would be produced by univariate linear mixed model. Variances on the dependent variable, here ATTIT, are assumed to be the same between categories of the level 2 grouping variable, here FEMALE. Residuals are assumed to be independent with constant variance. By assuming the same variance for males and females, the model is more parsimonious, requiring only 8 rather than 19 parameters. The homogenous model is selected by rejecting the top default unrestricted option under the Basic Settings tab and instead selecting the second, homogenous variance setting as shown below.



* + - *The model.* The foregoing selections add a homogenous variance model to the default unrestricted model in HMLM, as shown below, unless the researcher has checked "Skip unrestricted" in the Basic Settings window above.



When this model is run, the same homogenous variance model is expressed as below in print format. The level 1 model is the same as in the previous unrestricted model. The level 2 model differs in that random error terms are added at level 2 to the equations estimating the level 1 intercept (P0) and level 1 slope of AGE11 (P1) respectively: R0 and R1.

Summary of the model specified (in equation format)

---------------------------------------------------

Level-1 Model

Y = IND1\*Y1\* + IND2\*Y2\* + IND3\*Y3\* + IND4\*Y4\* + IND5\*Y5\*

Y\* = P0 + P1\*(AGE11) + e

Level-2 Model

P0 = B00 + B01\*(FEMALE) + R0

P1 = B10 + B11\*(FEMALE) + R1

Var(E) = Var(A\*R + e) = D = A\*Tau\*A' + S

where S = sigma\_squared\*I

A

IND1 1.00000 0.00000

IND2 1.00000 1.00000

IND3 1.00000 2.00000

IND4 1.00000 3.00000

IND5 1.00000 4.00000

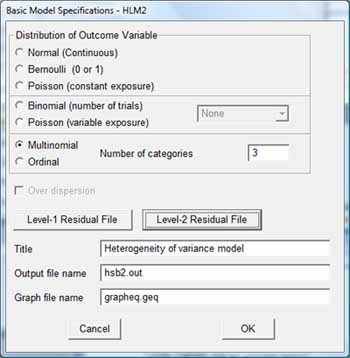
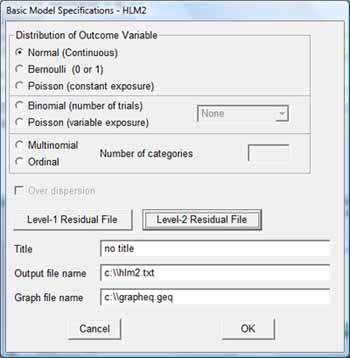
* + - *Fixed effects*. The output for the fixed effects table is similar to and interpreted the same as for the unrestricted model above. As before, only the B11 effect (the effect of FEMALE on the slope of AGE11) is non-significant. For this example, the researcher can conclude that the linear growth rate (0.070432) in the attitude toward deviant behaviors is significant, controlling for FEMALE (that is, for FEMALE = 0 = males). There is no significant gender effect on the linear growth rate (p = .222 for the B11 effect of FEMALE on the slope of AGE11).
    - Final estimation of fixed effects:
    - Standard Approx.
    - Fixed Effect Coefficient Error T-ratio d.f. P-value
    - ----------------------------------------------------------------------------
    - For INTRCPT1, P0
    - INTRCPT2, B00 0.221755 0.015961 13.894 237 0.000
    - FEMALE, B01 -0.048274 0.022926 -2.106 237 0.035
    - For AGE11 slope, P1
    - INTRCPT2, B10 0.070432 0.006781 10.386 237 0.000
    - FEMALE, B11 -0.012003 0.009826 -1.222 237 0.222
    - *"Summary of Model Fit" table, deviance, and likelihood ratio test of model difference*. HLM outputs a comparison of the unrestricted with the homogenous variance models. The likelihood ratio model comparison test in the bottom row is the likelihood ratio test of the significance of difference in deviance between two nested models. Lower deviance indicates better model fit. A finding of significance, such as p = 0.0 in this example, means that the homogenous variance model differs from the unrestricted model. Since sample size is large, even a small difference in deviance, as here, will be significant. The researcher might judge that the homogenous variance model was superior on parsimony grounds, having fewer parameters.
    - Summary of Model Fit
    - -------------------------------------------------------------------
    - Model Number of Deviance
    - Parameters
    - -------------------------------------------------------------------
    - 1. Unrestricted 19 -386.32097
    - 2. Homogeneous sigma\_squared 8 -345.90094
    - -------------------------------------------------------------------
    - Model Comparison Chi\_square df P-value
    - -------------------------------------------------------------------
    - Model 1 vs Model 2 40.42004 11 0.000

**Cross-Classified Models**

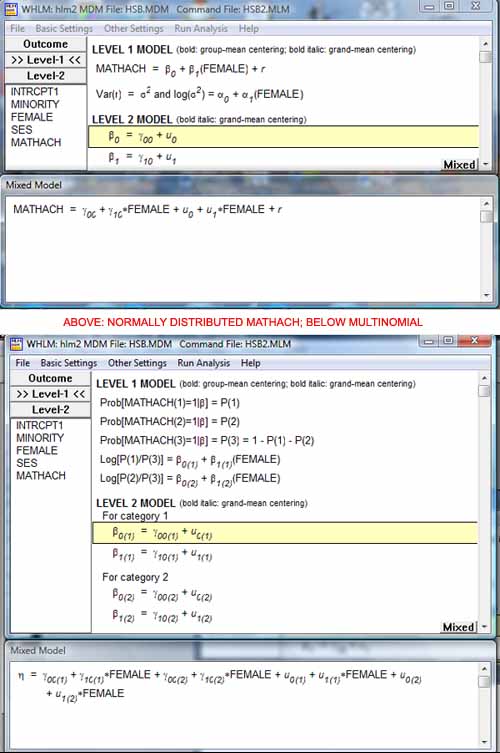
* + This additional coverage is found in Garson, ed. (2012).

**Generalized linear mixed models (GLMM)**

* + **GLMM** is primarily discussed in a [separate section](http://faculty.chass.ncsu.edu/garson/PA765/glmm.htm).
  + **Link functions.**
    - *Description.* In essence, GLMM is an extension of LMM to the case of non-normal dependent variables. As in generalized linear modeling ([GLM](http://faculty.chass.ncsu.edu/garson/PA765/glm.htm)), link function transformations of the dependent variable may be needed to fit dependent variables with binomial, multinomial, ordinal, Poisson, gamma, and other non-normal distributions. Whereas ordinary linear models maintain linearity in the identity function relating the independent to the dependent variables, generalized models maintain linearity in the link function relating independent to dependent variables.
    - *Link functions in HLM software*. Whereas SPSS and SAS conduct GLMM in separate modules, in HLM software, selection of link function is integrated with its several mixed model procedures. In HLM, clicking the "Basic Settings" menu choice allows the researcher to specify the distribution of the dependent variable, including normal, Bernoulli, Poisson, multinomial, and ordinal distributions, as illustrated below.



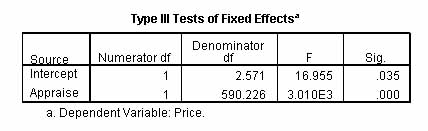
In the figure below, the researcher may also select "Multinomial" as the distribution of the outcome variable, causing a multinomial logit link to be employed, akin to multinomial logistic regression. There is also an option for hierarchical ordinal regression models. Hierarchical Poisson models employ a Poisson log link and require an exposure variable (time, for ex.).

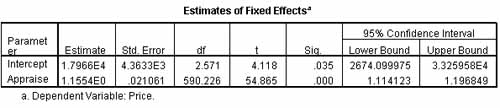


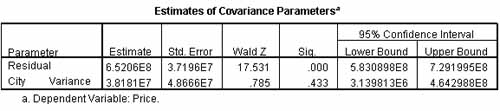
* + - *Example: binary outcome variables*. Selecting Bernoulli for a binary outcome variable applies a logistic link function and, as in logistic regression, making interpretations in terms of the log odds of the outcome rather than in terms of the raw outcome itself. The intercept ceases to be the mean value of the outcome but instead is the mean log odds of the outcom across level 2 units in a 2-level model. Also, there will be no level 1 variance component since the outcome is binary, and this means there is no simple intraclass correlation coefficient (discussed below for normally distributed outcome variables).

**LMM vs. GLM and VC Models**

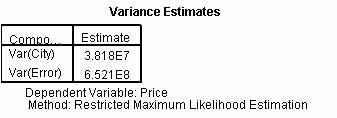
* + **Comparison with GLM or variance component procedures**.
    - *Example*. This section provides a brief example of a random intercept regression model in LMM for purposes of comparison with related output using the general linear model (GLM) and the variance components (VC) modules in SPSS. The example is a study of city real property appraisal values ("Appraise") to see how well they predict actual selling price of land parcels ("Price"), for each of three cities ("City"). The dependent variable in this example is Price. City is the grouping factor, conceptualized as a random effect. Appraise is a fixed covariate at level 1. There are no other level 2 predictors.
    - *SPSS*. In SPSS, select Analyze, Mixed Models, Linear; Enter the grouping variable, City, in the Subjects area; click Continue; in the next "Linear Mixed Models" dialog, enter Price as the Dependent Variable and Appraise as covariate; click the Fixed button and set Model to Appraise (so we are modeling one main fixed effect); click the Random button and enter City as the Subjects/Combinations variable; click the Statistics button to select output options such as Parameter Estimates and Tests for Covariance Parameters; click OK.
    - *Output*. Output below provides estimated coefficients for fixed and random effects. Interpretation is parallel to previous examples. The average selling price of a land parcel is estimated to be $17,967 when predicted from appraisal value nested within cities. Appraisal is confirmed to be a significant predictor. Increasing the appraisal value by $1.00 predicts increasing the parcel selling price by $1.15. In the covariance parameters table, the intercept, representing within-city variance in price, is 94% of the total of variance components and the between-city effect is 6% and is not significant. For these data we fail to conclude that there is a city effect.



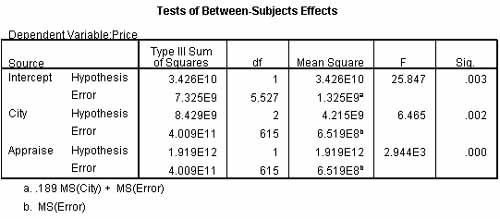




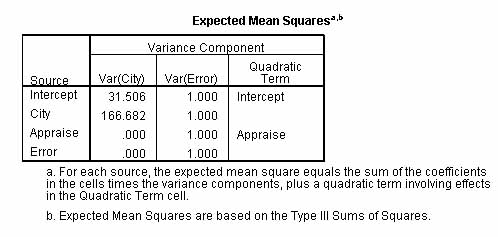
* + - *Comparison with Variance Components*. Linear mixed models in SPSS has the functionality of the [variance components procedure](http://faculty.chass.ncsu.edu/garson/PA765/variancecomponents.htm) described separately. That is, VARCOMP is mostly a subset of MIXED. For the same property appraisal example discussed above for linear mixed models, one may enter Price as dependent, Appraise as a covariate, and City as a random factor, specifying REML as the method so as to be comparable, to obtain the "Variance Estimates" below, which is identical (after rounding) to that in the "Estimates of Variance Components" table above for linear mixed modeling. However, variance components will not generate most of the additional output tables available for the linear mixed model procedure. While mostly a subset of MIXED, VARCOMP does have a few options not present in MIXED: it offers ANOVA and MINQUE estimation, not just ML and REML; and with ANOVA estimation comes sums of squares, mean squares, and expected mean squares not produced by the ML and REML estimation methods used by MIXED.

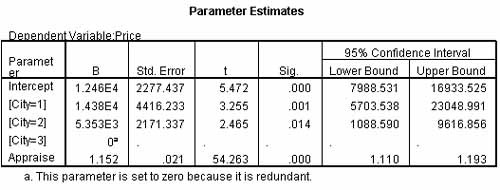


* + - *Comparison with GLM*. While one may obtain similar estimates using the general linear model (GLM), they will not be identical to the linear mixed model and variance components implementations illustrated above. When the design is unbalanced, as it is in this example (there are unequal numbers of observations in Cities 1, 2, and 3), the difference may be substantial. In GLM, one may enter Price as Dependent, Appraisal as a covariate, and City as a random factor, with the model set to main effects for Appraise and City, and selecting "Parameter estimates" under the Option button. GLM will generate these tables:



GLM produces Type III sums of squares for fixed effects only. Evem though City is entered as a random factor, the table above treats City as if it were a fixed effect for purposes of computing the sums of squares used to compute the F statistic. GLM estimates variance parameters for City (or any random effect) indirectly as described below, using expected mean squares. Linear mixed models and variance components analysis, in contrast, estimate variance parameters directly, using maximum likelihood (ML) or restricted maximum likelihood (REML) methods. For unbalanced designs (unequal n's in the groups formed by the random effect), as in this example, the GLM method will return estimates different from the methods used by linear mixed models or variance components analysis. Thus where in the linear mixed model run, the F value for the main effect Appraisal was 3.010E3, for the GLM run it is 2.994E3 in the table above.



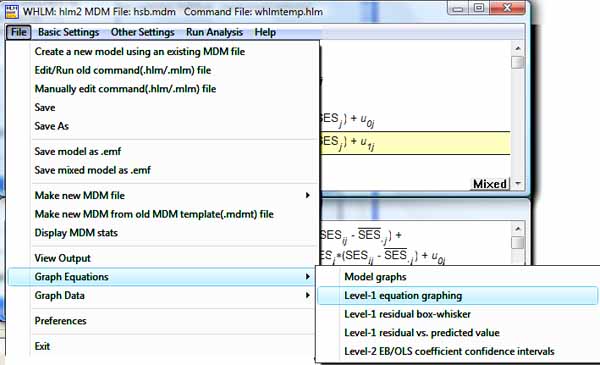


Above it is seen that the GLM method generates coefficient estimates for the fixed effect Appraise similar to that for linear mixed model. It also generates coefficients for the random effect City, which is not part of linear mixed model output due to the LMM not being based on sums of squares methods of estimation. However, the variance estimate for City, which was 38,180,000 in linear mixed modeling and in variance components analysis, is only 21,376,633 in GLM. The GLM variance estimate is computed as Var(City)=[MS(City)-MS(Error)]/EMS(City), where MS(City) = 4.215E9 and MS(Error)=6.519E8 (both from the "Between Subjects Effects" table in GLM) and EMS(City)=166.682 (from the "Expected Mean Squares" table in GLM). Even when the LMM and GLM variance estimates are the same as they will be in balanced designs, GLM has the drawback that the standard error of estimate for the variance of random factor(s) (ex., City) that appear in the "Estimates of Covariance Parameters" table in LMM, cannot be computed in GLM.

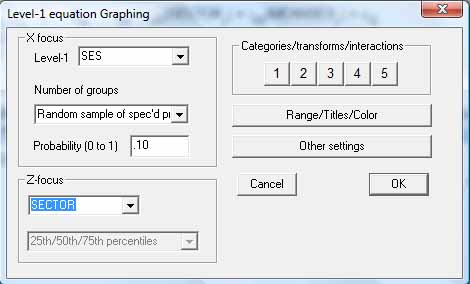
See further discussion in the [FAQ section](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#glmvslmm).

**Other Topics**

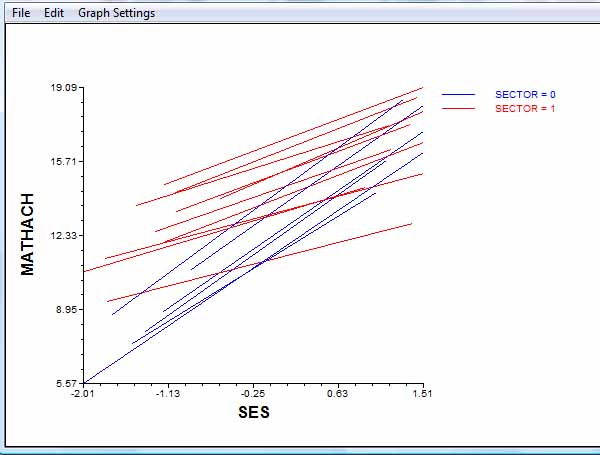
* + **Graphical analysis in HLM software.**
    - *Introduction.* HLM software provides a useful array of tools for graphical analysis. For the two-level RC model with Level 1 and Level 2 covariates (intercepts-and-slopes-as-outcomes model) discussed above, some forms of graphical analysis using HLM software is illustrated below.
    - Level 1 equation graphing.



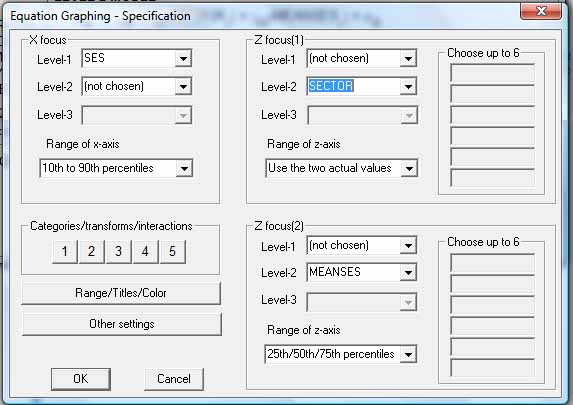
Below, the the "Level 1 equation graphing" dialog, we choose a random sample of 10% of the 160 schools.



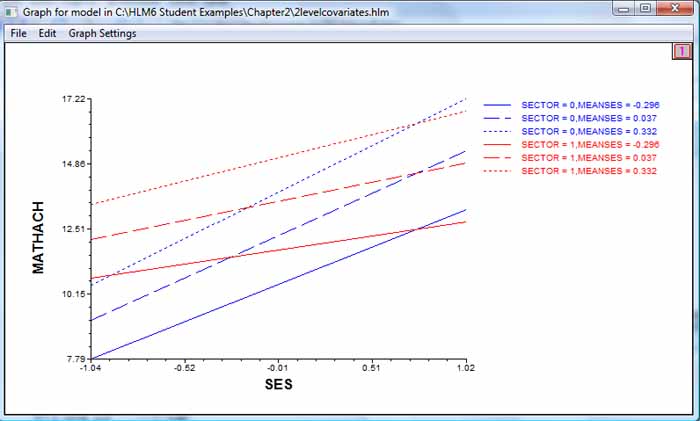
Below, lines in blue are SECTOR = 0 = public schools, while those in red are SECTOR = 1 = parochial schools. In this random sample, there are 6 public and 10 parochial schools, hence a total of 16 lines in the graph. The lines depict graphically how, as the level 1 variable centered SES increases within any given school, math achievement goes up. The slopes of the blue public school lines are steeper than the red line parochial school slopes, showing as discussed above that the within-school relation of SES to MATHACH is stronger for public schools.



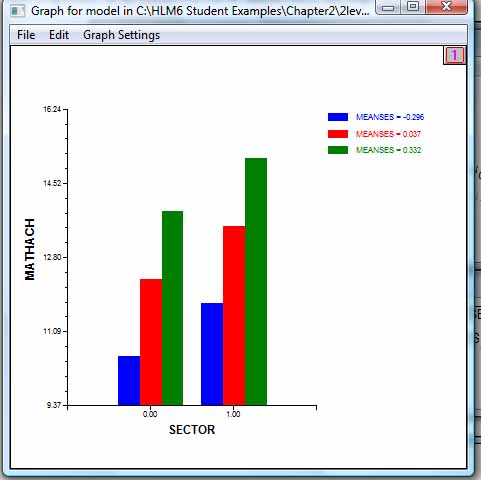
* + - General equation modeling. Choosing File, Graph equations, Model graphs from the menus brings up a more generalized dialog. In the example below we choose to graph level 1 centered SES by both level 2 variables - SECTOR and MEANSES. MEANSES can be represented at various levels but here we ask for the 15th, 50th, and 75th percentiles of MEANSES.



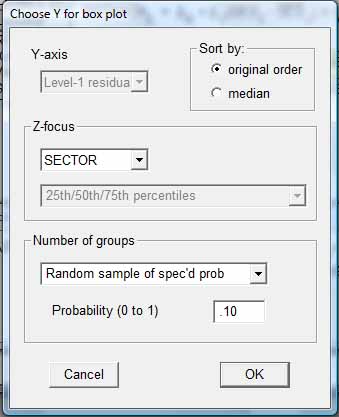
Below, the resulting graph shows that the same pattern of stronger relation (slope) of SES to MATHACH persists: at all selected levels of MEANSES, the slope of SES is greater for public schools (the blue lines).



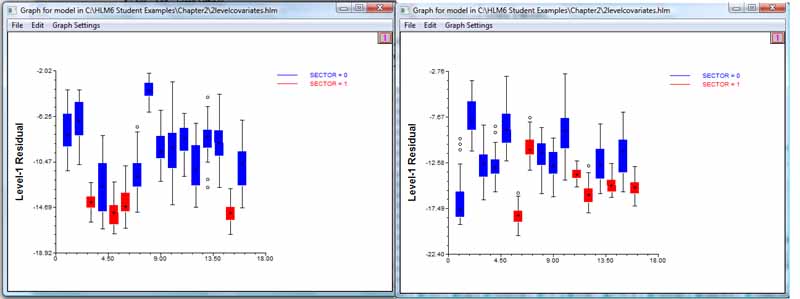
From the same "Model graphs" dialog, we can set the X focus as the level 2 variable SECTOR and the Z focus as the level 2 variable MEANSES, generating the graph below. This bar chart shows that at any percentile level of MEANSES (25th, 50th, 75th), math achievement is higher for parochial schools (SECTOR = 1) than for public schools.



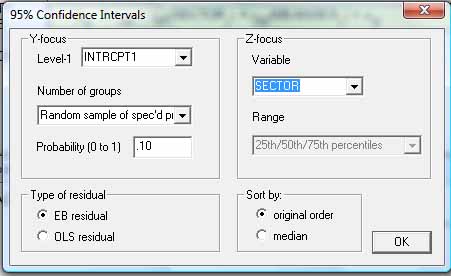
* + - Graphs for residual analysis. Choosing File, Graph equations, Model graphs from the menus brings up the following dialog. The Y axis is the residual of MATHACH when predicted by level 1 SES, adjusting for level 2 SECTOR and MEANSES. We ask for a 10% random sample of the 160 schools.



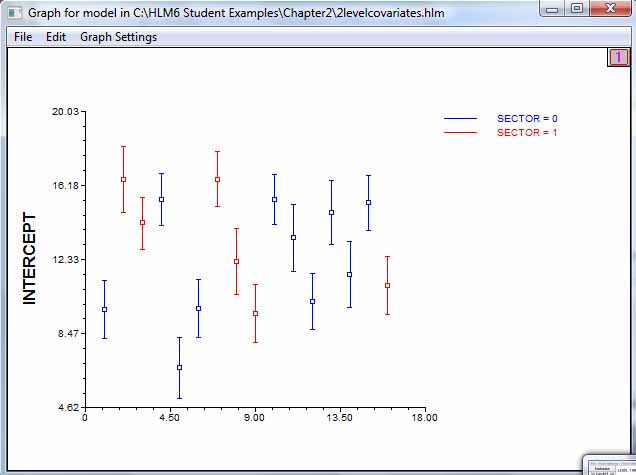
In the resulting graph, the Y axis represents the MATHACH residuals. The X axis is not meaningful, only representing the 16 random schools selected in each of two runs of this graph. In the box-and-whisker plot for any school, the box is the interquartile range and the whiskers are the distance to the minimum and maximum residual values. These box-and-whisker plots show than for this model, residuals (error) are generally greater for public than for parochial schools, and dispersion of error is also greater.



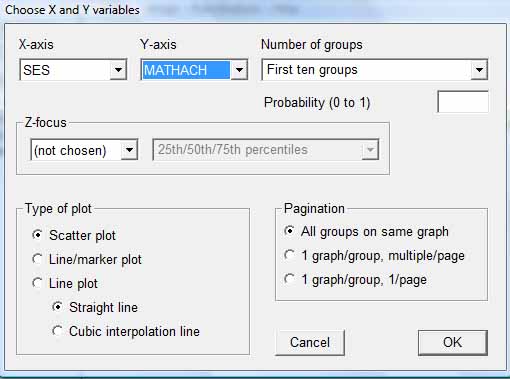
* + - Graphing parameter confidence intervals. Selecting "EB/OLS coefficient confidence intervals" from the File, Graph equations menu leads to the dialog below. Here we graph the confidence limits on B0, the intercept, but could equally have graphed the confidence limits on B1 by setting the level 1 Y focus as SES.



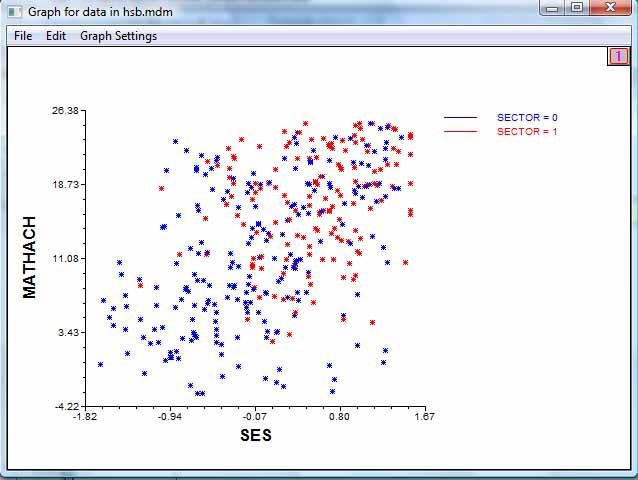
The resulting graph below shows, for the 16 randomly selected schools, than confidence limits on B0 do not differ markedly by sector.



* + - Graphing from data rather than models. Selecting File, Graph data, Line plots/scatterplots from the HLM menu leads to the dialog below. "Graph data" means that raw data are graphed, without benefit of fixed or random effects as controls. In this example we ask to graph MATHACH by SES.



The resulting graph below shows the individual-level relation of SES to MATHACH, by SECTOR. Points thus represent students, with blue points for public school and red points for parochial school students. It is possible to ask for one such graph per school, bhut here the first 10 are combined in one graph. The graph shows that math achievement scores tend to be higher for parochial schools. The relationship of SES to MATHACH is not very strong (as would be reflected by points being on a line rather than forming a cloud) for either sector, but is stronger for the public than parochial sector - leading to the expectation of high B1 slopes for SES in the public than the parochial sector.



* + - HLM supports many more graph options than covered here.

**Assumptions**

* + **Proper model specification**. Also, as in all models, the researcher should be aware that changes in what variables and effects are specified may well change coefficient estimates substantially.
  + **No high multicollinearity**. As with other forms of regression, high multicollinearity will make parameter estimates inefficient. Prior to running multilevel models, the researcher should rule out multicollinearity among the level 2 (or higher) predictors.
  + **Few outliers**. As with other forms of regression, presence of outliers will bias parameter estimates. Outliers may be tested in the SPSS regression module using Mahalanobis distance, leverage, and Cook's D (it is not a statistics option in the LMM module).
  + **Linearity**. Linear mixed models discussed in this module assume linearity between the independents and the dependent. Like regression, nonlinear terms may be added, however (ex., time-squared as well as time). Generalized linear mixed models (GLMM) supports nonlinear link functions (ex., logistic rather than linear (identity) linking of the dependent variable. Testing for linearity is discussed in the [module on testing for assumptions](http://faculty.chass.ncsu.edu/garson/PA765/assumpt.htm#linearity).
  + **Random grouping**. For levels above the bottom level of individuals as units of analysis, the groups (ex., schools, where students are the bottom level) are assumed to be a random sample of all such groups (ex., all schools) in random coefficients models. This is a critical assumption in multilevel modeling. I
  + **Independent observations** are not assumed, which is why multi-level modeling is recommended when intraclass correlation exists. That is, OLS regression and GLM in general assume error terms are independent and have equal error variances, whereas when one has hierarchical data, individual-level observations from the same upper level group will not be independent but rather will be more similar due to such factors as shared group history and group selection processes. This clustering by group increases Type I error if not taken into account, as LMM does. While random effects (such as upper-level effects, which LMM models as random effects) do not affect lower-level population means they do affect the covariance structure of the data and, indeed, adjusting for this is a central point of LMM models and why they are used instead of GLM, which assumes independence.

Repeated measures designs, if not adjusted by specifying repeated measures in LMM, involve autocorrelation, which will inflate the F and t tests as well as R-squared value.

To check for lack of independence, meaning some form of mixed modeling and/or repeated measures analysis is needed, the researcher can run an OLS regression and save the residuals. An ANOVA of residuals by group (ex., agency, where agency is the level 2 grouping for level 1 individual data) can be run. If the ANOVA F-test is significant, the researcher rejects the null hypothesis that residuals are independent by group. That is, a significant F means data are correlated, not independent, and LMM should be used instead of OLS. It is also possible to employ a [runs test](http://faculty.chass.ncsu.edu/garson/PA765/runs.htm), discussed separately.

* + - [**Intraclass correlation**](http://faculty.chass.ncsu.edu/garson/PA765/anova.htm#intra) is a measure of the extent to which observations are not independent of a grouping variable (ex., schools). The presence of a significant intraclass correlation is an indicator of the need to employ multi-level modeling rather than conventional regression, as discussed [above](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#icctest). To pursue OLS regression modeling anyway in the face of lack of independence and lack of homoscedastic error variance will mean that significance tests will not be accurate. Put another way, OLS significance tests (and standard errors and confidence limits) are not at all robust when the assumption of independence is violated.
  + **Independent blocks**. While observations are not assumed to be independent, the groups (blocks) formed by the subject variable(s) are assumed to be independent and to have the same covariance structures, for models that involve random effects or repeated measures.
  + **Continuous outcome variables**. Linear mixed modeling in SPSS, including HLM modeling, assumes the dependent variable is continuous. Binary, multinomial, and even ordinal dependent variables should not be employed with the default ML and REML estimation methods, but ordinal data may be used with Bayesian estimation.
  + **Properly specified covariance structures**. For random effects and separately for repeated measures, the researcher must specify the assumed covariance structure. Changing the specified covariance structure will change the covariance parameter estimates and tests of fixed effects.
  + **Random sampling**. As true for other procedures, significance tests assume random sampling so the sample is representative. For enumerations, all effects are significant since chance of sampling is ruled out. For convenience and other non-random samples, unknown bias is introduced into significance testing.
  + **Adequate sample size**. The efficiency and power of multi-level tests rests on pooled data across the units comprising two or more levels, which implies large datasets. The REML and ML estimation methods used by LMM give asymptotically efficient estimates, meaning efficiency depends on large samples. Hox (1995) suggests that for MLM regression models, the higher level sample size be at least 20, preferably 50, and if assessing variance components is important (not just comparing regression coefficients), preferably 100. For structural equation modeling approaches, Hox recommended sample size should be at least 50, preferably 100. More recently Maas & Hox (2005: 90), based on simulation studies, concluded, "The standard errors of the second-level variances are estimated too small when the number of groups is substantially lower than 100. With 30 groups, the standard errors are estimated about 15% too small..." In view of this, Maas & Hox (2005: 91-92) suggest that with small samples (as small as 10), that bootstrapping be used to estimate standard errors. However, bootstrapping was only available for multilevel modeling in MLwIN software, at the time of this writing.

Note that for purposes of improving power and precision of parameter estimates, increasing the number of level 2 groups is more important than increasing the number of level 1 individuals. For instance, simulation studies by Kreft (1996) found there was adequate statistical power with 30 groups of 30 observations each; 60 groups with 25 observations each; 150 groups with 5 observations each. There is a rapid fall-off in statistical power as the number of groups/observations falls below the threshhold needed. With less than adequate power there is an unacceptable risk of not detecting cross-level interactions (ex., between schools and students). However, both adequate number of individual observations and adequate number of groups are needed. Power for individual-level estimates depends on number of individuals observed, and power for second level estimates depends on number of groups.

Specifically with regard to MSEM, Hox & Maas (2001) used simulation studies to show for small group-level sample sizes, coefficient estimates were not stable. They recommended group-level samples of at least 100. However, Cheung & Au (2005: 612) used resampling to test sample size effects and found sample size "can be as small as 50, yet the results are still comparable with other larger sample size conditions." Unbalanced individual-level samples within groups may require larger group samples. Cheung & Au's experiments also disconfirmed the assertion of some that larger individual-level samples could compensate for small group-level samples.

Selecting Bayesian estimation rather than REML or ML estimation is one approach to small sample research. When Bayesian estimation is selected, smaller level 2 samples may be tolerated (Raudenbush and Bryk, 2002: 14).

* + **Similar group sizes**. The REML and ML estimation methods used by LMM give asymptotically efficient estimates for unbalanced as well as balanced designs (whereas the ANOVA methods in GLM are optimum only for balanced designs). Nonetheless, when sample sizes within groups are unbalanced, tests of parameter estimates and of overall fit will have inflated Type I error (Hox & Maas, 2001). However, FIML estimators are more robust for unbalanced designs (du Toit and du Toit, 2005).
  + **Normal distribution of variables**. RC models assume a normal distribution for purposes of empirical Bayes maximum likelihood estimation. However, REML and ML estimates may be assumed to display asymptotic normality for large samples. Also, extensions have been developed for non-normal data (Wong and Mason, 1985; Goldstein, 1991; Morris, 1995).
  + **Normal distribution of residuals**. Residual normality is required to properly define significance tests (for ex., to define the alpha region = .05). This can be tested in any statistical package by saving the residuals, then requesting a Q-Q plot: the more points lie on a straight line in this plot, the more residuals are normally distributed. In HLM 7, select Basic Settings, check Create Residual File, select SPSS (or SAS) as the Residual File Type, assign a filename, click OK. In SPSS, for instance, select Analyze, Descriptive Statistics, Q-Q plot. Use robust standard errors whenever there is departure from normality.
  + **Homogeneity of variance of residuals**. Error variance is not required to be constant across groups, unlike OLS regression. By default, statistical packages run homogenous variance models in which level 1 error terms are assumed to be homoscedastic (display homogeneity of variance). While the great majority of LMM modeling assumes this default, heterogenous variance options are available. Heterogeneity may arise for a number of reasons, including important omitted predictor variables, outliers, a skewed dependent variable, treating a random effect as fixed, and nonlinearity in the data.
    - *HLM*. HLM supports allows the variance of the level 1 random error term in LMM models to be modeled using a grouping variable which is not otherwise a predictor variable. This grouping variable need not be one defining a higher level in the model. An [earlier example](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#heterogenous) in this section illustrates heterogenous variance modeling in HLM, which includes a test for whether a heterogenous model is called for. From the HLM menu, select Other Settings, Hypotheses testing, Test homogeneity of level-1 variance. If the p value in this test is significant, then there is significant heterogeneity.
    - *SPSS* The SPSS REGWGT option is akin to [weighted least squares](http://faculty.chass.ncsu.edu/garson/PA765/wls.htm) analysis in ordinary regression models. Optionally the researcher can specify the name of a variable containing regression weights. These are applied (and only applied) to the residual covariance matrix. Residual weights are used for models with unequal variances between groups formed by the subject variable(s).
    - *SAS*. Heterogenous variance models in SAS may be implemented by treating the categorical variable associated with heterogeneity as a repeated measure in the REPEATED clause of PROC MIXED. See Hedeker & Mermelstein (2007).
  + **Dropped cases due to missing values** should by less than 5% of the total sample. While some researchers may use data imputation to eliminate missing values (ex., by using the [SPSS "Missing Data" module](http://faculty.chass.ncsu.edu/garson/PA765/missing.htm)), imputation introduces biases of its own.

**Frequently Asked Questions**

* + **Why use LMM instead of OLS regression?**

Data with multiple levels involve group effects on individuals which may be assessed invalidly by traditional statistical techniques. When grouping is present (ex., students in schools), observations within a group are often more similar than would be predicted on a pooled-data basis. That is, simple regression ignores grouping effects and violates the assumption of independence of observations. Mixed (multi-level) modeling handles this by using variables at upper levels (ex., school-level budgets at level 2) to adjust the regression of base level dependent variables on base level independent variables (ex., predicting student-level performance from student-level socioeconomic status scores).

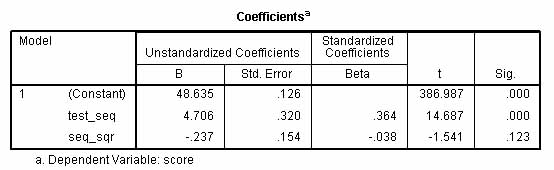
Multi-level modeling in LMM is particularly helpful in the analysis of covariance when data are sparse. For instance, in a study of a Social Security agency office, there may be too few minority employees to enable valid statistical inferences on performance evaluations, using traditional regression models. However, if multi-level data are available on employees and multiple SSA offices, then multi-level models can use not only the individual data in the SSA office but also information in the pooled data for all offices. The resulting prediction equation applied to the given SSA office will use coefficients reflecting both their own and also pooled data. For agencies with a large number of minorities, the multi-level and ordinary regression models will be similar. For agencies with sparse data -- few minorities -- it is true their estimate will rely considerably on the pooled data, but the advantage is that the pooling involved in multi-level models affords a "borrowing of strength" that supports statistical inference in a situation where no inference would be possible using traditional methods.

*Traditional regression models vs. LMM analysis*. There were three traditional approaches to regression modeling of multilevel data:

* + - 1. *Simple regression*, also called naive regression, simply ignored higher level effects (ex., ignored class or school effects in a study of students). This is appropriate, of course, only when the researcher can be sure there are no higher level effects. More often, there are such effects and simple regression leads to too low estimates of standard error, a higher rate of Type I errors and too-narrow confidence limits compared to multilevel modeling of the same data.
      2. *Fixed effects regression*. A popular traditional approach was to disaggregate data to the base level (ex., each student is assigned various school-level variables such as funding level per student, and all students in a given school have the same value on these *contextual variables*, and students are used as the unit of analysis). In fixed effects regression, sampling error is taken into account only for level 1 (the base) level, and sampling error at level 2 (or higher) is ignored. That is, information from fewer units at the upper level is wrongly treated as if it were independent data for the many units at the base level (the number of higher level observations is exaggerated), and this error in treating sample size led to over-optimistic estimates of significance. Also, there was the danger of the *ecological fallacy*: there is no necessary correspondence between individual-level and group-level variable relationships (ex., race and literacy correlate little at the individual level but correlate well at the state level, since Southern states have many African-Americans and many illiterates of all races). Finally, under fixed effects regression, the number of dummy variables increases as the number of clusters increases, making estimation inefficient. While adding crosslevel interaction terms is possible and does represent an improvement, it is inferior to multilevel modeling in LMM, which will model separate intercepts and slopes for individuals in each level 2 (or higher level) group, where the grouping variable is treated as a random effect.
      3. *Summary measures regression*. Another traditional approach to multi-level problems was to aggregate data to a higher level (ex., student performance scores are averaged to the school level and schools are used at the unit of analysis). Aggregated data was often centered (the mean was subtracted so the average value was zero). Ordinary OLS regression or another traditional technique was then performed on the unit of analysis chosen. A problem with summary measures regression is that under aggregation, fewer units of analysis at the upper level replace many units at the base level, resulting in loss of statistical power. As with simple regression, summary measures regression regression leads to too low estimates of standard error, a higher rate of Type I errors and too-narrow confidence limits compared to multilevel modeling of the same data. Summary measures regression also suffers from the possibility of ecological fallacy. Finally, summary measures regression prevents the valid analysis of covariate interactions due to loss of individual-level information.

Based on a review of the literature and on simulation studies, Ita G. G. Kreft (1996) concluded, "for researchers specifically interested in variance components, and posterior means, RC modeling provides them with separate estimates for separate contexts, and the iteration procedure improves the estimates of the variance components." That is, although effect size as revealed through regression is apt to be similar to effect size in multi-level modeling (see discussion below), multi-level modeling is more helpful in revealing differences in variance among units of analysis in different groups which comprise the levels. An empirical comparison of OLS regression with multilevel modeling by Moerbeek, van Breukelen, & Berger (2003) found that "The treatment effect and especially its standard error, are generally incorrectly estimated by traditional methods, which should, therefore, not in general be used as an alternative to multilevel regression" (p. 341). Also, multi-level modeling may be a preferred method when data are sparse, including studies (ex., twin studies) where groups are sparse.

* + - *Longitudinal example*. In this how-not-to example, imagine that a researcher wished to use OLS regression to see if "score" was related to time (test-seq) or time-squared (seq\_sq). Score could be regressed on test\_seq and seq\_sq to get the following output:



There are two problems with assessing growth in scores over time in this manner: (1) observations are not independent because one would predict they are clustered by employee and therefore there is likely an employee effect (tests are more similar within subjects than between subjects); and (2) observations may not be independent because they may be clustered by test occasion and there may be a time effect (with time measured by test\_seq). Having violated fundamental assumptions of OLS regression, we cannot make inferences on the resulting significance tests for the effects of time and time-squared.

* + - *In defense of OLS*. Based on a review of the literature and on simulation studies, Ita G. G. Kreft (1996) concluded, "if researchers in the social sciences are interested in the estimates of the regression parameters, the results of multi-level analysis will be close to the results obtained with more traditional regression techniques. In both cases the fixed effects estimates are unbiased. The main difference is in the standard errors of these parameters, which are estimated too small if [intra-class correlation](http://faculty.chass.ncsu.edu/garson/PA765/anova.htm#intra) is present in traditional [regression analyses](http://faculty.chass.ncsu.edu/garson/PA765/regress.htm). This fact makes the random coefficient model more conservative than the traditional regression." Kreft also raised questions about whether multi-level models are as generalizable as ordinary regression models. Because multi-level models rely on the complex, particular distribution of relationships across and within levels, Kreft concluded, "Outcomes are less general, since each best fitting model may be very specific for that dataset collected at that time and place." The smaller or less random the sample of higher-level units, the more true this would be. See also Kreft and de Leeuw (1991); Kreft, de Leeuw, and van der Leeden. (1994).
  + **Why use LMM instead of GLM?**

Many research problems might be modeled as general linear models (GLM option in SPSS) rather than as linear mixed models (Mixed option in SPSS). LMM provides a number of advantages, however:

* + - 1. GLM assumes independence but LMM does not, as discussed in the ["Assumptions" section](http://faculty.chass.ncsu.edu/garson/PA765/multilevel.htm#assumptions).
      2. Handling missing data: GLM will apply listwise deletion to drop cases with missing values, whereas Mixed (LMM) will include incomplete cases in the analysis.
      3. GLM repeated measures assumes all subjects are measured at the same points in time, whereas LMM allows subjects to be measured at different points in time.
      4. Repeated measures GLM requires subjects to have equal numbers of repeated measurements, but LMM allows unequal repetititons. That is, LMM is asymptotically efficient for both balanced and unbalanced designs, but GLM is optimally efficient only for balanced designs.
      5. GLM requires all interactions of within-subjects (repeated measures) and between-subjects factors be included in the model, whereas LMM allows the researcher to just include the interactions of interest.
      6. GLM makes certain assumptions about the covariance matrix and thus data must meet the sphericity test, whereas LMM allows for a wide variety of assumptions about the covariance matrix.
      7. LMM supports hierarchical data (data at one level nested within a higher level), but GLM does not.
      8. LMM estimates are based on maximum likelihood (ML) or restricted maximum likelihood (REML) methods, whereas GLM is based on ANOVA methods.
  + **What are the steps in multi-level modeling?**.

Hox (1995) suggests the following procedure for multi-level analysis:

* + - 1. Compute deviance for the baseline (null) model which includes only the intercept.
      2. Compute deviance for the model with the base level independent variables included and the variance components of the slopes constrained to zero (that is, for the fixed model).
      3. Use a chi-square difference test to see if the fixed model has a significantly better fit than the baseline model. If it does, then the researcher proceeds to investigate higher level modifier variables using random effects and/or RC models. Also, at this second step the researcher can assess the relative contribution of the base level independent variables. Drop non-significant base level independents and covariances from the model.
      4. Identify which base-level regression slopes have significant variance across upper level groups. Compute -2LL for the model with the variance components of the base-level slopes constrained to zero only for the slopes which do not have significant variance across upper level groups.
      5. Add upper level modifier variables, determining which improve model fit. Drop modifier variables which do not improve model fit.
      6. Add cross-level interactions between explanatory modifier variables and base level independent variables that had slope variance (in step 3). Drop interactions which do not improve model fit.
  + **Can I have no subject variable in LMM? What if I just click "Continue" on the SPSS opening LMM screen?**

It is conceptually clearer to have a subjects variable, which is a grouping variable used to define level 2 (or higher), and then in SPSS to enter it as the subjects/combinations variable under the Random button. This customary method corresponds to syntax (1) below. However, one will get identical results if one has no subjects variable, continues on in SPSS leaving the initial dialog screen blank, but then enters the grouping variable as a random effect to be modeled under the random button. This alternative method corresponds to syntax (2) below.

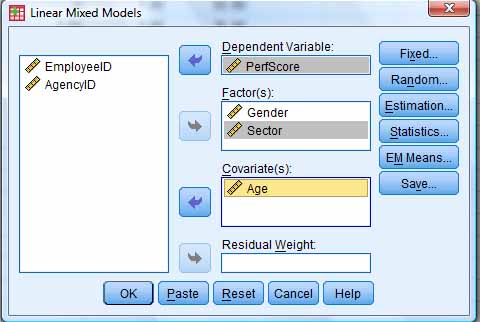
In the alternative two SPSS syntax segments below, Actual sale price of housing is predicted from appraisal value, with data nested within cities as the grouping variable.

* + - 1. Usual syntax with City as a Subjects variable.
      2. MIXED Price WITH Appraise
      3. /CRITERIA=CIN(95) MXITER(100) MXSTEP(10) SCORING(1) SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE) LCONVERGE(0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
      4. /FIXED=Appraise | SSTYPE(3)
      5. /METHOD=REML
      6. /PRINT=SOLUTION TESTCOV
      7. /RANDOM=INTERCEPT | SUBJECT(City) COVTYPE(VC).
      8. Alternative syntax with City as a random effect and no Subjects variable.
      9. MIXED Price BY City WITH Appraise
      10. /CRITERIA=CIN(95) MXITER(100) MXSTEP(10) SCORING(1) SINGULAR(0.000000000001) HCONVERGE(0, ABSOLUTE) LCONVERGE(0, ABSOLUTE) PCONVERGE(0.000001, ABSOLUTE)
      11. /FIXED=Appraise | SSTYPE(3)
      12. /METHOD=REML
      13. /PRINT=SOLUTION TESTCOV
      14. /RANDOM=INTERCEPT City | COVTYPE(VC).
  + **What is the multi-level equivalent to R-squared in OLS regression?**

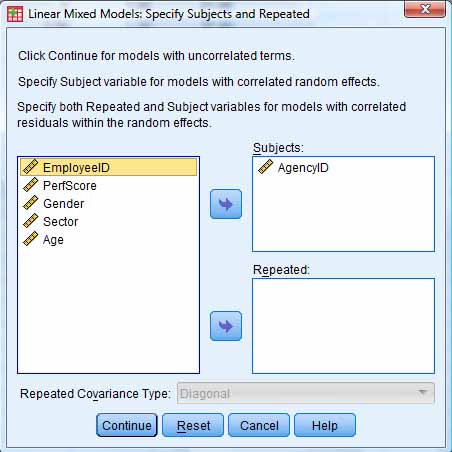
There is no equivalent to R-squared. Goodness of fit measures (ex., AICC, BIC) are used instead in multi-level modeling. Analogues to R-square have been proposed but are not widely used.

* + **How are data entered in SPSS software?**
    - *SAS and SPSS*. In SAS and SPSS, data are entered into a single file with rows representing observations. The observations are ordered by level 1 id within ordered groups of level 2 id, and so on for higher levels. For instance, student ids may be ordered within school ids ordered from 1 to n.
    - *SPSS overview*. Variables may be entered in SPSS linear mixed models in three locations:
      1. On the initial "Subjects and Repeated" LMM screen, the grouping variable is entered as the "Subjects" variable. For example, where agency-id is the level 2 grouping variable and empl\_id is the level 1 employee id variable, agency\_id would be entered as the "Subjects" variable. In a repeated measures study of test scores, where "time" was test occasion (0, 1, 2, etc.), each student would have one data row per test occasion and student\_id would the the grouping variable entered as "Subjects" and time would be entered as "Repeated".
      2. On the dialog screen for the "Fixed" button, the level 1 dependent variable (ex., score) would be entered as "Dependent" and any level 1 or level 2 variables would be entered as "Factors" if categorical or "Covariates" if continuous. If level 2 covariates are entered, their interactions with level 1 covariates should be entered as well. For repeated measures, time would be entered as a covariate (assuming metric intervals). Depending on the model, the grouping variable may or may not be entered as a factor or covariate (usually not).
      3. On the dialog screen for the "Random" button the grouping variable is entered in the ""Subjects"-"Combinations" area. One regression will be calculated for each level of the grouping variable, making the intercepts of the level 1 dependent variable a random effect of the grouping (Subjects-Combinations) variable. In the "Model" area one may enter level 1 variables whose slopes are hypothesized to be random effects of the level 2 grouping variable. Level 2 covariates are not entered in a two-level model as there are no higher levels for which their slopes might be a random effect (instead they are entered as fixed effects).
    - *The dependent variable* in univariate multilevel modeling is assumed to be a level 1 (individual level) normally distributed quantitative variable which is linearly related to the fixed and random factors and covariates in the model. Do not use a binomial or multinomial variable as a dependent, for example. In multivariate multilevel modeling, there may be more than one dependent variable (this is analogous to the difference between ANOVA and MANOVA in GLM).

As illustrated below, the dependent variable, factors, and covariates are entered in the SPSS dialog which follows when "Continue" is clicked on the initial SPSS LMM page. Factors and covariates are then declared by clicking on the "Fixed" or "Random" buttons as appropriate (see below).

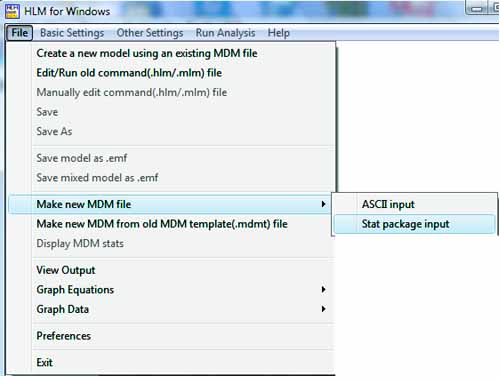


* + - *Grouping variables*
      1. *Subject variables*. Subject variables are variables which define the groupings. That is, in a two-level study, there are one or more level 2 id variables (ex., schoolid; raceid, religionid). The subject variable is used to define groups such that each group is independent of the others. For instance, in a study of individual performance test scores grouped by agency, if "EmployeeID" is the id number for the individual employee and "AgencyID" is the id number for the agency in which the person works, then, as illustrated below for SPSS, it would be the AgencyID variable which would be entered as the subject variable. An exception would be repeated measures studies (ex., where any given employee is tested for performance, say, in each of 6 months). In repeated measures, the six monthly scores are grouped by EmployeeID. Since the individual employee would then be the grouping variable, "EmployeeID" would be the subject variable in a repeated measures study and "Month" would be the repeated measures variable.



* + - * + *The level 2 grouping variable*. The subject variable, for ex. AgencyID, is seen as a random sample of all possible grouping entities at level 2 (ex., all possible agencies). This means the subject variable identifies a set of randomly selected level 2 entities. This implies that the level 2 grouping (subject) variable is associated with a random error term. In study of public administration, each agency will have a different effect on the level 1 variable being predicted, such as employee performance score. That is, performance score is a random effect of the subject variable (Agency) plus possibly other agency-level or higher-level predictors that may be in the model, plus possibly other level 1 predictors. By making AgencyID a subject grouping variable, one regression will be run for each agency. The intercept of the dependent (performance score) is modeled as the mean of all the intercepts in the separate regressions for each agency (each subject).
        + *Multiple identifiers*. In some cases more than one variable is needed as the subject variable. That is, more than one variable may define the groups. For instance, if subject variables are Gender and JobStatus, then the groups might be Male-Working, Male-NotWorking, Female-Working, and Female-NotWorking. The researcher would be hypothesizing that these two subject variables explained variance in whatever dependent variable was measured. Observations would be similiar within groups but each group is assumed independent of the others.
        + *SPSS:* The initial "Linear Mixed Models: Specify Subjects and Repeated" dialog screen allows the researcher specify one or more subject variables. If there are repeated measures or random effects a subject variable is usually entered. Note that if an observation has a missing value on any of the subject variables, it will be dropped from analysis.
      1. *Repeated measures variables*. Repeated measures variables are observation (time) variables for repeated measures studies. Repeated measures variables are usually metric, reflecting equal intervals between observations. For instance, one might have a single repeated variable "Year" to indicate each of four years in which a measurement was taken. Alternatively, there could be two repeated measures variables, Month and Day, which together would constitute the measure for the repeated measurements. This, too, is specified on the initial "Linear Mixed Models: Specify Subjects and Repeated" dialog screen in SPSS.
  + **How are data entered in HLM software?**

HLM software stores data in its own multivariate data matrix (MDM) format, which may be created from raw data or created from datafiles imported from SPSS, SAS, Stata, or other packages. MDM format files come in flavors keyed to the several types of HLM modules noted above. File creation options are accessed from the HLM File menu, illustrated below. The example below illustrates data entry from SPSS .sav file for models of type HLM2, but similar procedures are followed for other model types.



* + - 1. *Stat file input* is the most common method of creating MDM files. In the following example, level 1 is students and level 2 is schools, as from the Singer (1998) "High School and Beyond" study (provided as an HLM example file) illustrated below.
         * *Input method 1: Separate files for each level.* This method results in faster processing but requires more time to set up the data. It requires that separate files be created for each level of HLM analysis. For SPSS, these are .sav files. For SAS, these are SAS 5 transport files. Separate SYSTAT and Stata files are also acceptable.

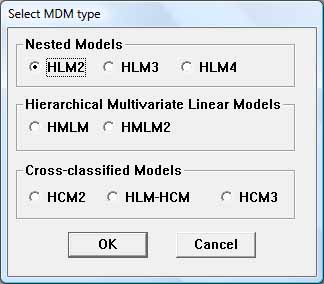
For instance, in the SPSS example here, the HSB1.SAV file contains the level 2 link field (id is school id) and any student-level variables. There are multiple rows per school, one row per student, with each student measured on the variables minority, female, ses (standardized), and mathach. In the HSB2.SAV file there is the same level 2 link field (id is still school id) with additional school-level variables. There is one row per school, with the school-level variables size, sector (public = 0 or parochial = 1), pracad (binary, reflecting proportion in academic track), disclim (ordinal rating on disciplinary climate, from -2 to +2), himinty (binary, with 0 = less than 40% minority), and meanses (coded -1, 0, +1, reflecting mean standardized SES of students in a school).

Thus the level 1 and level 2 variables are in separate files, with the school id variable as the link. In the file for level 1 (HSB1.SAV) the data must be sorted such that all students for a given school id are adjacent.

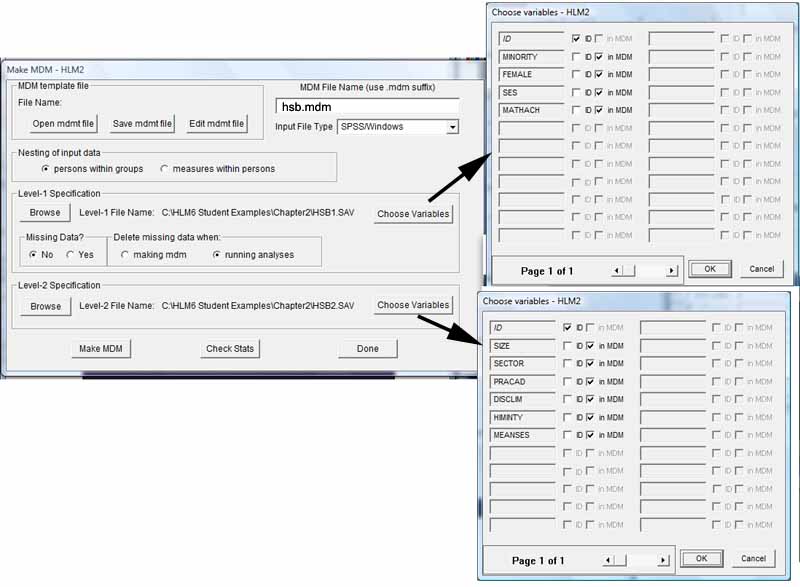


* + - * + *Input method 2: Using a single statistics program data file.* This method is easier in terms of data management. The same stat file formats as for Method 1 may be used. The single data file must be sorted such that all students for a given school id are adjacent. (This method is supported for HLM 6 and higher).
        + *Making the MDM file.* The next step is to create the .MDM file, which is HLM software's native data format (after it is created, the input data files are not needed). After the researcher has created the input data file(s) in SPSS, SAS, SYSTAT, Stata (or other packages supporting the generic HLM input format), the researcher runs HLM. In the initial "HLM for Windows" window, the researcher selects "File", then "Make new MDM file", then "Stat package input."

After selecting "Stat package input," the "Select MDM type" window illustrated below appears. The researcher chooses the HLM model type wanted. For a simple two-level hierarchical linear model, the selection would be HLM2, for instance.



After selecting HLM2, the "Make MDM - HLM2" dialog box appears, illustrated below.



Here the following steps are necessary:

Set the "Input File Type" to "SPSS/Windows" or to the selection for another statistical package.

In the level 1 specification area, click the "Browse" button and browse to the input file for level 1. Then click the "Choose variables" button and click the checkbox indicating the level 2 link variable (id in the example) and click the checkboxes of any other level 1 variables in the analysis.

In the level 2 specification area, click the "Browse" button and browse to the input file for level 2. This may be the same file as for level 1 (following Method 2 above). Again click the "Choose variables" button and click the checkbox indicating the level 2 link variable (id) and click the checkboxes of any other level 2 variables in the analysis.

Save the MDM template file by clicking the "Save mdmt file" button, making sure the file location window points to the desired folder and giving a filename (add the .mdmt extension), then clicking the "Save" button.

To complete the process, the researcher clicks the "Make MDM" button, giving a filename (here, hsb.mdm). The .mdm file is created and the descriptive statistics module runs. Alternatively, one may click the "Check Stats" button. This output should be examined to verify the results. For instance, it is prudent to examine the reported sample size, which, if low, flags that the researcher has not sorted the Level 1 file to assure individual rows for the same level 2 id (school id in this example) are adjacent.



Click the "Done" button to exit to the WHLM model construction screen discussed below. At this point, the researcher will have saved three files to the disk: the newly created HLM-compatible data file, HSB.MDM in this example; the default template creatmdm.mdmt (the researcher may override the default name); and the output file above, HLM2MDM.STS (if desired, use File, Save As, to save output under a different name as this default file may get re-used with new content if there are multiple runs).

* + **How do I make a single SPSS file from two .SAV files used to create the HLM .MDM file, as described above?**

The researcher might want to do this to use hierarchical linear modeling in SPSS rather than in HLM software. Even if the .MDM files for each level were separate SPSS .SAV files, the linear mixed models module in SPSS wants a single file in which each level 1 case for a given level 2 entity will have the same level 2 variable values as any other case in that entity.

In SPSS, open both the level 1 and level 2 .SAV files used to create the .MDM file. Make the level 2 file the active one. In SPSS, select Data, Merge, Add Variable. In the ensuing dialog, select the level 1 file as the one to merge with. Then check "Match cases on key variables". Then select the radio button reading "Active dataset is keyed table." Set the key variable to the id variable. Note that the data must have been sorted ascending on the key (id) variable.

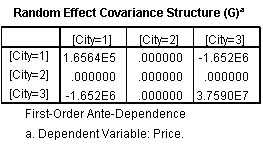
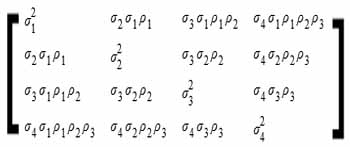
* + **Why does my GLM model give the same parameter estimates and corresponding significance levels as LMM for the same data, but LMM does not print out sums of squares?**

For some models, such as fixed effects with uncorrelated residuals, GLM and LMM will give the same parameter estimates and significance levels (LMM in the "Tests of Fixed Effects" and "Estimates of Fixed Effects" tables; GLM in the "Tests of Between-Subjects Effects" and "Parameter Estimates" tables). However, LMM can fit a wider variety of models than GLM and in some LMM models, test effects cannot be expressed in terms of ratios of sums of squares and therefore there is no "Sum of Squares" column.

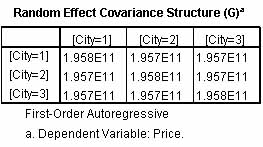
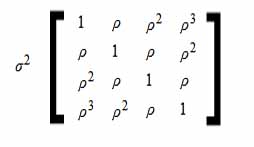
* + **Besides variance components and diagonal (simple) covariance structure assumptions, what other assumptions might be made?**

SPSS supports the following, listed alphabetically:

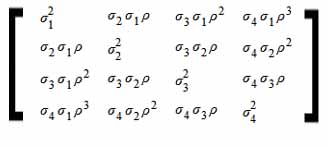
* + - *Ante-Dependence: First Order*, AD(1). This covariance structure has heterogenous variances and heterogenous correlations between adjacent elements. The correlation between two nonadjacent elements is the product of the correlations between the elements that lie between the elements of interest.



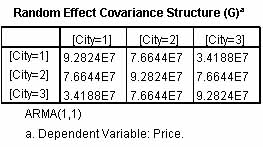
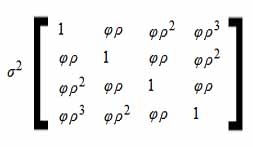
* + - *AR(1)*. This is first-order autoregressive structure with homogenous variances. AR(1) is the most frequent assumption for data where there is a common trend, such as where the correlation of any pair of repeated measurements is assumed to decrease exponentially according to how far apart they are in time. It is an assumption common to time series analysis and requires that for all subjects the intervals between any two adjacent time periods are the same. That is, AR(1) assumes the time variable is metric and equally spaced. The residuals at time 1 are less similar to residuals at time 4 than they would be with the residuals at time 2. The larger the time lag, the lower the correlation of residuals. In the residual covariance matrix, the diagonal variances will be roughly equal, and the off-diagonal covariances will show a pattern, normally decreasing over time. (The particular data for the random effects variable city, illustrated below, do not follow this pattern). An AR(1) pattern would exist, for example, if the correlation between any two elements (ex., time periods) is equal to r for time-adjacent elements , then was r2 for elements that are separated by one other element, r3 for elements that are separated by two other elements, etc., within the constraint that the bounds of +/- 1 are not exceeded.



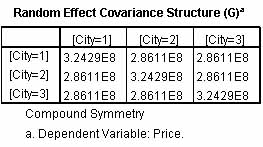
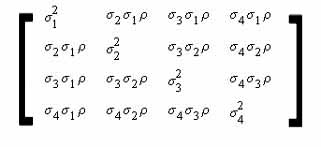
* + - *AR(1): Heterogeneous*. This is a first-order autoregressive structure with heterogenous variances. The correlation between any two elements is equal to r for adjacent elements, r2 for two elements separated by a third, and so on.



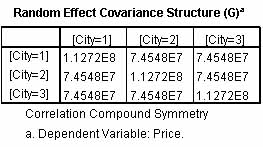
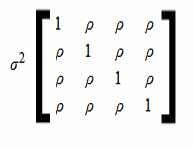
* + - *ARMA(1,1)*. This is a first-order autoregressive moving average structure. It has homogenous variances. The correlation between two elements is equal to f\*r for adjacent elements, f\*(r2) for elements separated by a third, and so on. r and f are the autoregressive and moving average parameters, respectively, and their values are constrained to lie between –1 and 1, inclusive.



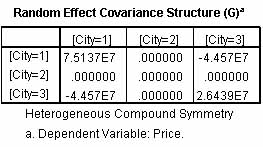
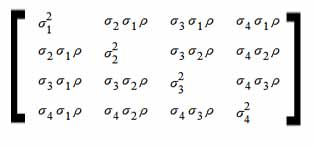
* + - *Compound Symmetry*. Also called "Exhangeable" covariance structure, this is a common repeated measures covariance structure assumption. Within-subject correlation of error terms is assumed to be equal. That is, if elements are time periods, the correlation of residuals for measurements nearby in time should be the same as for measurements with a large time-distance. Compound symmetry may not be plausible because it assumes measures close in time are no more correlated than measures far apart in time. This covariance structure has homogenous variances and homogenous covariances between elements (ex., time periods for repeated measures, cities for random effects). This means that the residuals have the same covariance for any pair of time periods (repeated measures) or cities (random effects). Also, the variance of residuals is the same for any time period or city. Thus, in the residual covariance matrix, the diagonal variances will be roughly equal, and the off-diagonal variances will also be roughly equal to each other. This is the type assumed in univariate Anova models and is the classical approach to repeated measures. The assumption of compound symmetry is more likely to be met with classical experimental data than with longitudinal data where autoregressive or Toeplitz assumptions may be more appropriate.



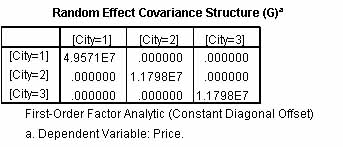
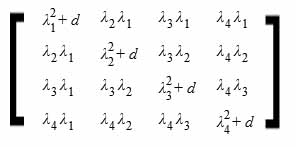
* + - *Compound Symmetry: Correlation Metric*. This covariance structure has homogenous variances and homogenous correlations between elements.



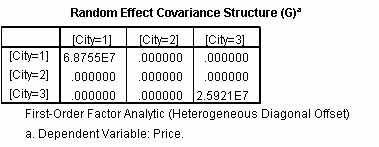
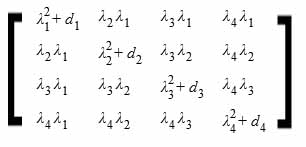
* + - *Compound Symmetry: Heterogeneous*. This covariance structure has heterogenous variances and constant correlation between elements.



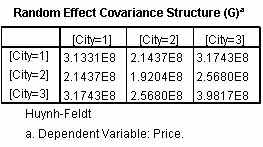
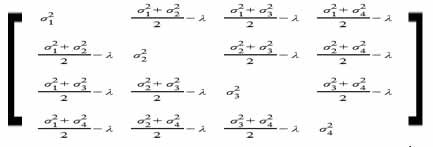
* + - *Factor Analytic: First Order*. This covariance structure has heterogenous variances that are composed of a term that is heterogenous across elements and a term that is homogenous across elements. The covariance between any two elements is the square root of the product of their heterogenous variance terms.



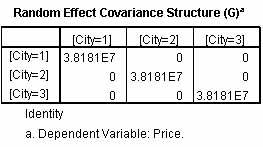
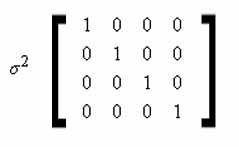
* + - *Factor Analytic: First Order, Heterogeneous*. his covariance structure has heterogenous variances that are composed of two terms that are heterogenous across elements. The covariance between any two elements is the square root of the product of the first of their heterogenous variance terms.



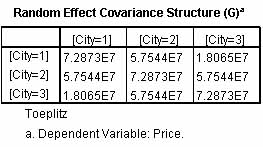
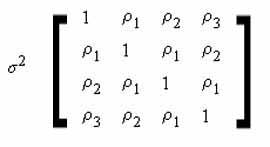
* + - *Huynh-Feldt*. This is a "circular" matrix in which the covariance between any two elements is equal to the average of their variances minus a constant. Neither the variances nor the covariances are constant.



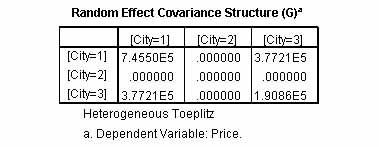
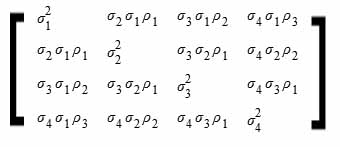
* + - *Scaled Identity*. This structure has constant variance. There is assumed to be no correlation between any elements. This is a common assumption when modeling the interaction of a random factor (ex., City) with a fixed grouping factor (ex., Agency), where it is assumed that the City\*Agency interaction effect is normally distributed around a mean of zero, with unknown variance to be estimated.



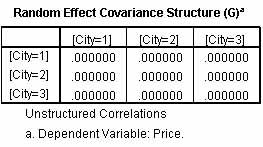
* + - *Toeplitz*. This structure is a generalization of the AR(1) type. Like AR(1), the Toeplitz model assumes metric, equal time intervals and assumes pairs of within-subject correlations are equal for the same time-distance apart. Unlike the special case of AR(1), however, in a Toeplitz model the pattern sequence for off-diagonal covariances does not have to step by some common multiple but may step by some unique multiple associated with that time step. That is, there is a trend as in AR(1) models but there is no common function describing that trend for all time intervals. This covariance structure has homogenous variances and heterogenous correlations between elements. The correlation between adjacent elements is homogenous across pairs of adjacent elements. The correlation between elements separated by a third is again homogenous, and so on.



* + - *Toeplitz: Heterogeneous*. This covariance structure has heterogenous variances and heterogenous correlations between elements. The correlation between adjacent elements is homogenous across pairs of adjacent elements. The correlation between elements separated by a third is again homogenous, and so on.



* + - *Unstructured: Correlations*. This covariance structure can have heterogenous variances and heterogenous correlations.



* + **What about multilevel modeling in structural equation models rather than linear mixed models?**

Multilevel structural equation models (MSEM) are generalizations of path analysis and as such are part of the [structural equation modeling](http://faculty.chass.ncsu.edu/garson/PA765/structur.htm) (SEM) family. MSEM models can have multiple dependent variables or latent variables (constructs measured by indicator variables) as well as multiple levels of measurement. See McDonald (1994) and Muthén (1994). While MLM regression models (MRM) and MSEM will perform similarly, compared to MRM, the SEM approach can more easily incorporate complex path models and multiple group models. Multilevel regression modeling (MRM) is compared with multilevel structural equation modeling (MSEM) by Tomarken & Waller (2005: 38). They note the advantages of MSEM to be more and more interpretable measures of goodness of fit, better modeling of residuals, and better capacity to model latent variables. They note the advantages of MRM to be easier model specification, fewer estimation problems, and ability to handle certain types of analysis difficult to handle within MSEM. For instance, in MSEM you cannot model between-group variablility in factor loadings or path coefficients (Hox, 2002). They note, however, that the two approaches are more similar than different, and as techniques and software evolve, functional similarity is increasing. On the relation of multilevel modeling to SEM, see Curran (2003).

Multilevel structural equation modeling is implemented in Mplus and to a lesser extent in LISREL but not AMOS. See Schumacker & Lomax (2010: 307-320).

* + **By modeling a grouping factor (ex., cities) as a random effect, may the researcher generalize conclusions to all cities, not just those in the sample?**

While one sometimes encounters this statement in literature on random effects models, it is incorrect if stated in an unqualified manner. There is no way for, say, results from a convenience sample of five cities in Utah to be generalized to all U. S. cities. The statement would be true only if the cities were randomly selected from all U. S. cities, and the number of these level 2 groups (cities) was sufficient (>20 is a common rule of thumb).

* + **What are split plot experiments in LMM?**

Split plot experiments are a common type of linear mixed model. Arising out of agricultural applications, there would be one factor such as levels of irrigation applied randomly to whole plots of land within farms, then another factor such as seed types applied randomly to subplots (the "split plots") nested within the plots. The whole plots reflect diffent randomly-selected levels of irrigation and contain a set of subplots with randomly selected seed types. There may also be control variables, called "blocking variables," such as the farm on which the plots are located. There is, of course, also a dependent variable such as plant height, called the "response variable." In this example, irrigation level, seed type, and farm would all be "classification variables" for plant height as dependent. A variety of mixed models could be constructed using these factors as fixed and random effects. For instance, one could specify irrigation, seed type, and irrigation\*seed type as fixed effects and farm and irrigation\*farm as random effects. Output of LMM would include covariance parameter estimates for Farm, Farm\*Irrigation interaction, and Residual, along with AIC and other fit statistics for the model, interpreted as discussed above.

* + **What software is available for multilevel modeling?**

Several software packages for multi-level modeling have emerged in the last decade. Leading packages are listed below.

* + - 1. [MPlus](http://www.statmodel.com) supports multi-level modeling with latent variables. It is more comprehensive and makes multilevel structural equation modeling easier than other packages as of this writing.
      2. [SPSS's "Linear Mixed Models" module](http://www.spss.ch/upload/1126184451_Linear%20Mixed%20Effects%20Modeling%20in%20SPSS.pdf), part of its SPSS Advanced Models extension, handles hierarchical linear models (HLM) as well as related models for random or mixed ANOVA and ANCOVA, repeated measures ANOVA and MANOVA, and variance component estimation (VARCOMP).
      3. [*AMOS*](http://www.SPSS.com/amos/) *and* [*LISREL*](http://www.ssicentral.com). It is possible to implement multi-level models in structural equation modeling programs like AMOS and LISREL. LISREL's MLM module is called *MULTILEV*.
      4. [*SAS's PROC MIXED* procedure](http://cc.uoregon.edu/cnews/summer2004/procmixed.htm) can implement several models: simple random-effect only, simple mixed with a single fixed and random effect, split-plot, multilocation, repeated measures, analysis of covariance, random coefficients, and spatial correlation See Littell et al. (1999).
      5. [HLM](http://www.ssicentral.com/hlm/index.html), authored Steve Raudenbush and Tony Bryk. Raudenbush headed the longitudinal and [multi-level methods project](http://www.googlesyndicatedsearch.com/u/MichiganStCOE?q=multilevel) at Michigan State University. *HLM* can read data from a variety of statistical packages, including SPSS, SAS, SYSTAT, and STATA, and it covers nonlinear as well as linear models. A [free student version](http://www.ssicentral.com/hlm/student.html) is available. This was perhaps the leading package during the development of multi-level modeling in the 1990s. HLM does not have a bulit-in data editor: data preparation must be done in SPSS (which HLM imports) or another program. HLM does not read ordinal variables, which must be converted to a series of dummy variables in the data preparation stage. Cross-level interaction terms are created automatically by HLM, and there is an option for automatic centering of variables (group mean centered or grand mean centered).
      6. [MLWin](http://www.cmm.bristol.ac.uk/MLwiN/index.shtml), a Windows program produced by the UK/Canada Multilevel Models Project, for models with any number of levels. It is the Windows version of the earlier MLn multi-level modeling software package.
  + **Can HLM models be replicated in structural equation modeling (SEM)?**

EQS and MPlus support certain types of multilevel structural equation modeling. LISREL supports two-level general structural equation modeling. Rabe-Hesketh, Skrondal, & Zheng (2007) have described how to implement multilevel SEM using GLLAMM within Stata.

*Latent growth models (LGM)* are a type of multilevel model suitable for clustered data with repeated observations of variables at multiple levels. While multiple group SEM is a conventional approach discussed in the [SEM section](http://faculty.chass.ncsu.edu/garson/PA765/structur.htm) of *Statnotes*, multilevel SEM (MSEM) is a newer methodological development which handles large numbers of groups (ex, 100 - 200). Strictly speaking, an LGM model adapts SEM to analyze change over time when both individual and group-level variables are in the model. However, LGM has been extended and generalized to other types of multi-level models and may refer simply to multilevel models based on SEM software as opposed to regression-based multilevel models.. As such, LGM is a more versatile alternative to repeated measures ANOVA (see Tomarken & Waller, 2005: 36 for a list of nine ways LGM is more versatile). For an introduction to LGM, see Duncan, Duncan, Strycker, et al. (1999) and Hox (2002). See also du Toit & du Toit (2005).

* + **What is the SPSS syntax for linear mixed modeling (MIXED)?**
  + MIXED dependent varname [BY factor list] [WITH covariate list]
  + [/CRITERIA = [CIN({95\*\* })] [HCONVERGE({0\*\* } {ABSOLUTE\*\*})
  + {value} {value} {RELATIVE }
  + [LCONVERGE({0\*\* } {ABSOLUTE\*\*})] [MXITER({100\*\*})]
  + {value} {RELATIVE } {n }
  + [MXSTEP({5\*\*})] [PCONVERGE({1E-6\*\*},{ABSOLUTE\*\*})] [SCORING({1\*\*})]
  + {n } {value } {RELATIVE } {n }
  + [SINGULAR({1E-12\*\*})] ]
  + {value }
  + [/EMMEANS = TABLES ({OVERALL })]
  + {factor }
  + {factor\*factor ...}
  + [WITH (covariate=value [covariate = value ...])
  + [COMPARE [({factor})] [REFCAT({value})] [ADJ({LSD\*\* })] ]
  + {FIRST} {BONFERRONI}
  + {LAST } {SIDAK }
  + [/FIXED = [effect [effect ...]] [| [NOINT] [SSTYPE({1 })] ] ]
  + {3\*\*}
  + [/METHOD = {ML }]
  + {REML\*\*}
  + [/MISSING = {EXCLUDE\*\*}]
  + {INCLUDE }
  + [/PRINT = [CORB] [COVB] [CPS] [DESCRIPTIVES] [G] [HISTORY(1\*\*)] [LMATRIX] [R]
  + (n )
  + [SOLUTION] [TESTCOV]]
  + [/RANDOM = effect [effect ...]
  + [| [SUBJECT(varname[\*varname[\*...]])] [COVTYPE({VC\*\* })]]]
  + {covstruct+}
  + [/REGWGT = varname]
  + [/REPEATED = varname[\*varname[\*...]] | SUBJECT(varname[\*varname[\*...]])
  + [COVTYPE({DIAG\*\* })]]
  + {covstruct†}
  + [/SAVE = [tempvar [(name)] [tempvar [(name)]] ...]
  + [/TEST[(valuelist)] =
  + ['label'] effect valuelist ... [| effect valuelist ...] [divisor=value]]
  + [; effect valuelist ... [| effect valuelist ...] [divisor=value]]
  + [/TEST[(valuelist)] = ['label'] ALL list [| list] [divisor=value]]
  + [; ALL list [| list] [divisor=value]]
  + \*\* Default if the subcommand is omitted.
  + † covstruct can take the following values: AD1, AR1, ARH1, ARMA11, CS, CSH, CSR, DIAG, FA1, FAH1, HF, ID, TP, TPH, UN, UNR, VC.

**Bibliography**

* + Booth, J.G. & Hobert, J.P. (1998). Standard errors of prediction in generalized linear mixed models. *Journal of the American Statistical Association* 93, 262 -272.
  + Breslow, N. E., & Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association* 88, 9-25.
  + Bickel, Robert (2007). *Multilevel analysis for applied research: It's just regression!* Thousand Oaks, CA: Sage.
  + Bryk, A.S., and Raudenbush, S.W. (2002). *Hierarchical linear models: Applications and data analysis methods, Second ed.* Newbury Park, Sage Publications.
  + Bryk, A. S., S. W. Raudenbush, and R. Congdon (1996). *Hierarchical Linear and nonlinear modeling with the HLM/2L and HLM/3L Programs*. Chicago: Scientific Software International.
  + Burton, B. (1993). *Some observations on the effect of centering on the results obtained from hierarchical linear modeling*. Washington, DC: National Center for Education Statistics, U. S. Department of Education.
  + Cheung, M. W.-L. & Au, K. (2005). Applications of multilevel structural equation modeling to cross-cultural research. *Structural Equation Modeling* 12(4), 598-619.
  + Curran, P. J. (2003). Have multilevel models been structural equation models all along? *Multivariate Behavioral Research* 38(4): 529–569.
  + de Leeuw, Jan & Kreft, I. G. G. (1986). Random coefficient models for multi-level analysis. *Journal of Educational Statistics* 11: 57-86.
  + de Leeuw, Jan; Meijer,Erik ; & Goldstein, H. (2007). *Handbook of multilevel analysis.* NY: Springer.
  + Diggle, P. J. (1988). An approach to the analysis of repeated measurements. *Biometrics* 44, 959-971.
  + Duncan, T. E., Duncan, S. C., Strycker, L. A., Li. F, & Alpert, A. (1999). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications.* Mahwah, NJ: Erlbaum
  + du Toit, S. H. C. & du Toit, M. (2005). Multilevel structural equation modeling. In de Jeeuw, J. & Kreft, I. G. G., eds., .*Handbook of quantitative multilevel analysis.* Boston: Kluwer Acad.
  + Ehlers, Margaret (2004). Assessment of covariance selection strategies with repeated measures data. 34th Annual SCASA Meeting, American Statistical Association. Retrieved 4/223/07 from http://www.stat.sc.edu/scasa/Apr04t.html.
  + Garson, G. David, ed. (2012). *Hierarchical linear modeling: Guide and applications.* Thousand Oaks, CA: Sage Publications.
  + Gelman, Andrew & Hill, Jennifer (2006). *Data analysis using regression and multilevel/hierarchical models.* NY: Cambridge University Press.
  + Goldstein, H. (1991). Nonlinear multi-level models with an application to discrete response data. *Biometrika* 78: 45-51.
  + Goldstein, H. (1995). *Multilevel statistical models.* London, Edward Arnold; New York, Halstead Press.
  + Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46(6): 1251-1271. Basis of the Hausman test, whereby the random-effects estimator is tested for consistency vis-à-vis the fixed effects estimator.
  + Heck, Ronald H.; Thomas, Scott L. ; and Tabata, Lynn N. (2010). *Multilevel and longitudinal modeling with IBM SPSS.* . NY: Routledge.
  + Hedeker, D. & Mermelstein, R.J. (2007). Mixed-effects regression models with heterogeneous variance: Analyzing ecological momentary assessment data of smoking. Ch. 8 in T.D. Little, J.A. Bovaird, & N.A. Card, eds., *Modeling contextual effects in longitudinal studies*. Erlbaum: Mahwah, NJ. Available at this url: <http://www.uic.edu/classes/bstt/bstt513/Hedeker_Mermelstein_07.pdf>.
  + Hoffman, D. A. & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research organizations. *Journal of Management* 24(5), 623-641.
  + Hox, J. J. (1995). *Applied multi-level analysis, 2nd Ed.*. Amsterdam: TT-Publikaties. A basic, non-technical introductory text.
  + Hox J. J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, NJ: Erlbaum.
  + Hox, J. J. & Maas, C. J. M. (2001). The accuracy of mulitlevel structural equation modeling with pseudobalanced groups and small samples. *Structural Equation Modeling* 8, 157-174.
  + Jiang, Jiming (2007). *Linear and generalized linear mixed models and their applications* Springer Series in Statistics. New York: Springer.
  + Julian, M. W. (2001). The consequences of ignoring multilevel data structures in nonhierarchical covariance modeling. *Structural Equation Modeling* 8, 325-352.
  + Kreft, Ita G. G. (1996). Are multi-level techniques necessary? An overview, including simulation studies. Obtained online on Nov. 22, 1999, at http://www.calstatela.edu/faculty/ikreft/quarterly/quarterly.html.
  + Kreft, Ita G. G., Jan de Leeuw, J., and Aiken, Leona S. (1995). The effect of different rorms of centering in hierarchical linear models. *Multivariate Behavioral Research* 30(1), 1-21.
  + Kreft, Ita G. G., Jan de Leeuw, J., and R. van der Leeden (1994). Review of five multi-level analysis programs: BMDP-5V, GENMOD, HLM, ML3, VARCL. *American Statistician* 48(4): 324-335.
  + Kreft, Ita G. G. and Leeuw, Jan de (1991). Model-based ranking of schools. *International Journal of Education* 15: 45-59.
  + Kreft, Ita G. G. and Leeuw, Jan de (1998). *Introducing multi-level modeling*. Thousand Oaks, CA: Sage Publications. Introducing Statistical Methods Series. A user-oriented introductory text with a minimum of formal mathematics. Includes worked examples using real datasets using *MLn*, the predecessor of *MLWin*.
  + Littell, Ramon C. ; Milliken, George A.; Stroup, Walter W.; Wolfinger, Russell D. ; & Schabenberber, Oliver (2006). ) *SAS for mixed models, second edition*. Cary, NC: SAS Publishing.
  + Longford, N.T. and Muthen,B.O. (1992). Factor analysis for clustered populations. *Psychometrika*, 57, 581-97.
  + Longford, N.T. (1993). *Random coefficient models.* Oxford, Clarendon Press.
  + Luke, Douglas (2004). *Multilevel modeling*. Thousand Oaks, CA: Sage Publications. Vol. 142 in Quantitative Applications in the Social Sciences Series.
  + Maas, CoraJ. M. & Hox, Joop J. (2005) Sufficient sample sizes for multilevel modeling. *Methodology* 1(3), 86-92.
  + McCulloch, Charles E. (1997). Maximum likelihood algorithms for generalized linear mixed models. *Journal of the American Statistical Association* 92, 162-170.
  + McCulloch, Charles E. & and Searle, Shayle R.(2001). *Generalized, linear, and mixed models*. NY: Wiley-Interscience.
  + McCulloch, Charles E. (2003). *Generalized linear mixed models*. Regional Conference Series. Bethesda, MD: Institute of Mathematical Statistics
  + McDonald, R.P. (1994). The bilevel reticular action model for path analysis with latent variables. *Sociological Methods and Research*, 22, 399-413.
  + McDonald, R.P. and Goldstein, H. (1988). Balanced versus unbalanced designs for linear structural relations in two level data. *British Journal of Mathematical and Statistical Psychology,* 42, 215-32.
  + Moerbeek, Mirjam; van Breukelen, Gerard J. P.; & Berger, Martijn P. F. (2003). A comparison between traditional methods and multilevel regression for the analysis of multicenter intervention studies. *Journal of Clinical Epidemiology* 56, 341-350.
  + Morris, C. (1995). Hierarchical models for educational data - an overview. *Journal of Educational and Behavioral Statistics* 20(2).
  + Muthén, B.O. (1989). Latent variable modelling in heterogeneous populations. *Psychometrika,* 54, 557-85.
  + Muthén, B.O. (1994). Multilevel covariance structure analysis. *Sociological Methods and Research* 22: 376-399.
  + Peugh, James L. & Enders. Craig K. (2005). Using the SPSS Mixed procedure to fit cross-sectional and longitudinal multilevel models. *Educational and Pcyhological Measurement* 65(5), 717-741. Replicates Singer (1998) using SPSS Mixed instead of SAS. Enders has made supplementary instructional material available at <http://www.asu.edu/clas/psych/people/documents/FittingmultilevelmodelsusingSPSSpull-downs.pdf>.
  + Rabe-Hesketh, Sophia; Skrondal, Anders ; & Zheng, Xiaohui (2007). "Multilevel structural equation modeling." Pp. 277-301. in S. Y. Lee, ed/. *Handbook on structural equation models*. Amsterdam: Elsevier.
  + Raudenbush. Stephen W. (2010). O*utlines & highlights for hierarchical linear models: Applications and data analysis methods, Vol. 1.* Ventura, CA: Academic Internet Publishers Inc. (AIPI).
  + Raudenbush, Stephen W. and Anthony S. Bryk (2002). *Hierarchical linear models: Applications and data analysis methods, Second edition*. Thousand Oaks, CA: Sage Publications. Advanced Quantative Techniques in the Social Sciences Series No. 1.
  + Schumacker, Randall E. & Lomax, Richard G. (2010). *A beginner's guide to structural equation modeling, Third edition*. NY: Routledge.
  + Singer, J. (1998). Using SAS Proc Mixed to fit multilevel models, hierarchical models, and individual growth curves. *Journal of Educational and Behavioral Statistics* 24(4): 323-355.
  + Singer, J. & Willett, J. (2003). *Applied logitudinal data analysis*. Oxford, UK: Oxford University Press.
  + Snijders, T. & Bosker, R. (2000). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: Sage.
  + SPSS, Inc. (2005). *Linear mixed-effects modeling in SPSS: An introduction to the MIXED procedure*. SPSS Technical Report. Chicago, IL: SPSS, Inc.
  + Sullivan, Lisa M.; Dukes, Kimberly A.; & Losina, Elena (1999). Tutorial in biostatistics: An introduction to hierarchical linear modelling. *Statistics in Medicine* 18: 855-888. Illustrated using HLM and SAS.
  + Thompson, R. (1985). A note on restricted maximum likelihood estimation with an alternative outlier model, *Journal of the Royal Statistical Society*. Series B (Methodological). 47(1), 53-55.
  + Tomarken, A. J. & Waller, N. G. (2005). Structural equation modeling: Strengths, limitations, and misconceptions. *Annual Review of Clinical Psychology* 1, 31-65.
  + Verbeke, Geert & Molenberghs, Geert (1997). *Linear mixed models in practice: A SAS-oriented approach.* Lecture Notes in Statistics. New York: Springer.
  + Verbeke Geert & Molenberghs, Geert (2000). *Linear mixed models for longitudinal data*. NY: Springer-Verlag.
  + Verbeke, G., & Molenberghs, G. (2000). *Linear mixed models for longitudinal data.* New York: Springer.
  + West, Brady; Welch, Kathleen B.; & Galecki, Andrzej T. (2006). *Linear mixed models: A practical guide using statistical software.* Boca Raton, FL: Chapman & Hall/CRC.
  + Wong, G. and W. Mason (1985). The hierarchical logistic regression model for mulitlevel analysis. *Journal of the American Statistical Association* 80: 513-524.

@c 2006, 2008, 2009, 2010, 2011 G. David Garson  
Do not post on other servers, even for educational use.  
Last update: 8/14/2011.

Top of Form

Bottom of Form