HMLM model for the NYS data

1. Introduction to HMLM models

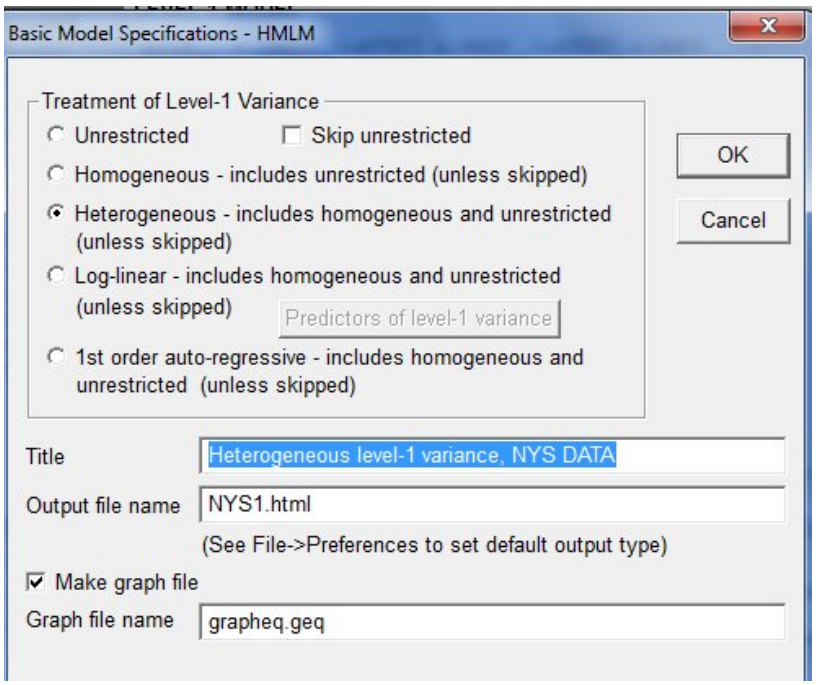
One of the most frequent applications of hierarchical models involves repeated observations (level 1) nested within persons (level 2). These are described in Chapter 6 of *Hierarchical Linear Models*. In these models, the outcome Y_{ij} for occasion i within person j is conceived as a univariate outcome, observed under different conditions or at different times. An advantage of viewing the repeated observations as nested within the person is that it allows each person to have a different repeated measures design. For example, in a longitudinal study, the number of time points may vary across persons, and the spacing between time points may be different for different persons. Such unbalanced designs would pose problems for standard methods of analysis such as the analysis of variance.

Suppose, however, that the aim of the study is to observe every participant according to a fixed design with, say, T observations per person. The design might involve T observation times or T different outcome variables or even T different experimental conditions. Given the fixed design, the analysis can be reconceived as a multivariate repeated measures analysis. The multivariate model is flexible in allowing a wide variety of assumptions about the variation and covariation of the T repeated measures (Bock, 1985). In the standard application of multivariate repeated measures, there can be no missing outcomes: every participant must have a full complement of T repeated observations.

Advances in statistical computation, beginning with the EM algorithm (Dempster, Laird, & Rubin, 1977; see also Jennrich & Schluchter, 1986), allow the estimation of multivariate normal models from incomplete data. In this case, the aim of the study was to collect T observations per person, but only n_j observations were collected ($n_j \leq T$). These n_j observations are indeed collected

according to a fixed design, but $T - n_j$ data points are missing at random.

Model specification for HMLM analysis are specified via the **Basic Settings** dialog box as shown below.



2. Description of the model

HMLM allows estimation of multivariate normal models from incomplete data. Within the framework of HMLM, it is possible to estimate models having

1. An unrestricted covariance structure, that is a full $T \times T$ covariance matrix.
2. A model with homogenous level-1 variance and random intercepts and/or slopes at level-2.
3. A model with heterogeneous variances at level 1 (a different variance for each occasion) and random intercepts and/or slopes at level 2.
4. A model that includes a log-linear structure for the level-1 variance and random intercepts and/or slopes at level 2.
5. A model with first-order auto-regressive level-1 random errors and random intercepts and/or slopes at level 2.

We note that applications 2 - 4 are available within the standard HLM2. However, within HMLM, models 2 - 4 can be compared to the unrestricted model (model 1), using a likelihood ratio test. No "unrestricted model" can be meaningfully defined within the standard HLM2; such a model is definable only within the confines of a fixed design with T measurements.

This model is appropriate when the aim of the study is to collect T observations per participant according to a fixed design. However, one or more observations may be missing at random. We assume a constant but otherwise arbitrary $T \times T$ covariance matrix for each person's "complete data."

The level-1 model relates the observed data, Y, to the complete data, Y^*:

$$Y_{hi} = \sum_{t=1}^{T} m_{thi} Y_{ti}^*$$

where Y_{hi} is the r-th outcome for person i associated with time h. Here Y_{ti}^* is the value that person i would have displayed if that person had been observed at time t, and m_{thi} is an indicator variable taking on a value of 1 if the h-th measurement for person i did occur at time t, 0 if not. Thus, Y_{ti}^*, $t = 1, ..., T$, represent the complete data for person i while Y_{hi}, $h = 1, ..., T_i$ are the observed data, and the indicators m_{thi} tell us the pattern of missing data for person i.

To make this clear, consider $T = 5$ and a person who has data at occasions 1,2, and 4, but not at occasions 3 and 5. Then Equation 9.1 expands to

$$\begin{pmatrix} Y_{1i} \\ Y_{2i} \\ Y_{3i} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} Y_{1i}^* \\ Y_{2i}^* \\ Y_{3i}^* \\ Y_{4i}^* \\ Y_{5i}^* \end{pmatrix}$$

or, in matrix notation,

$$Y_i = M_i Y_i^*$$

This model says simply that the three observed data points for person i were observed at times 1, 2, and 4, so that data were missing at times 3 and 5. Although these data were missing, they do exist, in principle. Thus, every participant has a full 5×1 vector of "complete data" even though the $T_i \times 1$ vector of observed data will vary in length across persons.

We now pose a structural model for the within-person variation in Y^*:

$$Y_{ti}^* = \pi_{0i} + \sum_{p=1}^{P} \pi_{pi} a_{pt} + \varepsilon_{ti}$$

or, in matrix notation

$$Y_i^* = A\pi_i + \varepsilon_i,$$

where we assume that ε_i is multivariate normal in distribution with a mean vector of 0 and an arbitrary $T \times T$ covariance matrix Δ. In fact, Δ is not a "within-person" covariance. Rather, it captures all variation and covariation among the T repeated observations. The level-2 model includes covariates, X_i, that vary between persons:

$$\pi_{pi} = \beta_{p0} + \sum_{q=1}^{Q} \beta_{pq} X_{qi}$$

or in matrix notation

$$\pi_i = X_i \beta$$

Note there is no random variation between persons in the regression coefficients π_{pi} because all random variation has been absorbed into Δ.

Substituting the level-2 model into the level-1 model gives the combined model for the complete data, in matrix form:

$$Y_i^* = AX_i\beta + \varepsilon_i, \qquad \varepsilon_i \sim N(0, \Delta)$$

Here the design matrix captures main effects of within-person covariates (the *a*s), main effects of person-level covariates (*X*s), and two-way interaction effects between them ($a \times X$ terms).

In sum, our reformulation poses a "multiple measures" model that relates the observed data Y_i to the "complete data" Y_i^*, that is, the data that would have been observed if the researcher had been successful in obtaining outcome data at every time point. Our combined model is a standard multivariate normal regression model for the complete data.

Algebraically substituting the combined model expression for Y_i^* into the model for the observed data yields the combined model

$$Y_i = M_i AX_i\beta + M_i\varepsilon_i.$$

Under the unrestricted model, the number of parameters estimated is $f + T(T+1)/2$, where *f* is the number of fixed effects and *T* is the number of observations intended for each person. The models below impose constraints on the unrestricted model, and therefore include fewer parameters. The fit of these simpler models to the data can be compared to the fit of the unrestricted model using a likelihood ratio test.

This is the first of a set of three HMLM models fitted to the NYS data.

6. Creating the MDM file

The range of options for data input are the same as for HLM2 and HLM3. We will use SPSS file input in our example.

The level-1 file, NYS1.SAV, has 1,079 observations collected from interviewing annually 239 eleven-year-old youths beginning at 1976 for five consecutive years. Therefore, *T* = 5. The variables and the *T* indicator variables are:

ATTIT	a 9-item scale assessing attitudes favorable to deviant behavior. Subjects were asked how wrong (very wrong, wrong, a little bit wrong, not wrong at all) they believe it is for someone their age to, for example, damage and destroy property, use marijuana, use alcohol, sell hard drugs, or steal. The measure was positively skewed, so a logarithmic transformation was performed to reduce the skewness.
EXPO	Exposure to deviant peers. Subjects were asked how wrong their best friends thought the nine deviant behaviors surveyed in the ATTIT scale were.
AGE	age of the participant
AGE11	age of participant at a specific time minus 11
AGE13	age of participant at a specific time minus 13
AGE11s	AGE11* AGE11
AGE13s	AGE13* AGE13
IND1	indicator for measure at time 1
IND2	indicator for measure at time 2
IND3	indicator for measure at time 3
IND4	indicator for measure at time 4
IND5	indicator for measure at time 5

The five indicators were created to facilitate use of HMLM. Data for the first two children are shown below. Child 15 had data at all five years. Child 33, however, did not have data for the fourth year.

While the structure of HMLM input files is almost the same as in HLM2, there is one important difference: the indicator variables. In order to create these, one first needs to know *the maximum number of level-1 records per level-2 group; this determines the number of indicators*. We shall call them the number of "occasions." (This is the number of time points in a repeated measures study or the number of outcome variables in a cross-sectional multivariate study. Also note that each person does not need to have this number of occasions.) Then create the indicator variables so that a given variable takes on the value of 1.0 if the given occasion is at this time point, 0.0 otherwise. Looking at the level-1 data file shown below, we see that IND1 is 1 if AGE11 is 0, IND2 is 1 if AGE11 is 1, IND3 is 1 if AGE11 is 2, and so on.

Indicators for the repeated measures

	id	attit	age	age11	age13	age11s	age13s	ind1	ind2	ind3	ind4	ind5
16	15	.44	11.00	.00	-2.00	.00	4.00	1.0	.00	.00	.00	.00
17	15	.44	12.00	1.00	-1.00	1.00	1.00	.00	1.0	.00	.00	.00
18	15	.89	13.00	2.00	.00	4.00	.00	.00	.00	1.0	.00	.00
19	15	.75	14.00	3.00	1.00	9.00	1.00	.00	.00	.00	1.0	.00
20	15	.80	15.00	4.00	2.00	16.00	4.00	.00	.00	.00	.00	1.0
21	33	.20	11.00	.00	-2.00	.00	4.00	1.0	.00	.00	.00	.00
22	33	.64	12.00	1.00	-1.00	1.00	1.00	.00	1.0	.00	.00	.00
23	33	.69	13.00	2.00	.00	4.00	.00	.00	.00	1.0	.00	.00
24	33	.11	15.00	4.00	2.00	16.00	4.00	.00	.00	.00	.00	1.0

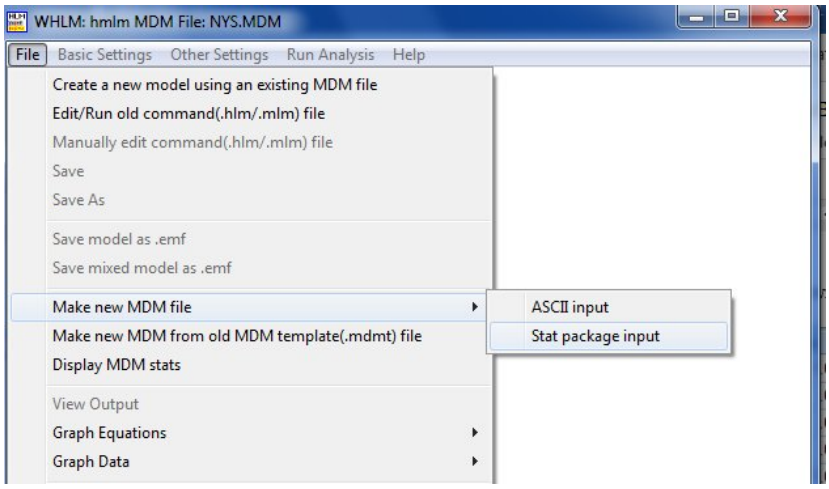
The level-2 data file, NYSB.SAV, consists of three variables on 239 youths. The file has the same structure as that for HLM2. The variables are:

FEMALE	an indicator for gender (1 = female, 0 = male)
MINORITY	an indicator for ethnicity (1 = minority, 0 = other)
INCOME	income

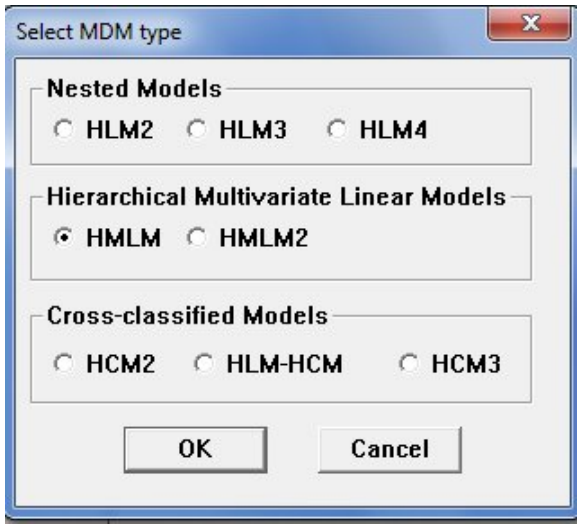
The first few lines of the level-2 data file are shown below.

	id	female	minority	income
1	3	1.00	.00	3.00
2	8	.00	.00	4.00
3	9	.00	.00	3.00
4	15	.00	.00	4.00
5	33	1.00	.00	4.00
6	45	1.00	.00	4.00
7	52	.00	.00	3.00
8	77	.00	.00	4.00
9	79	1.00	.00	6.00
10	83	.00	1.00	3.00

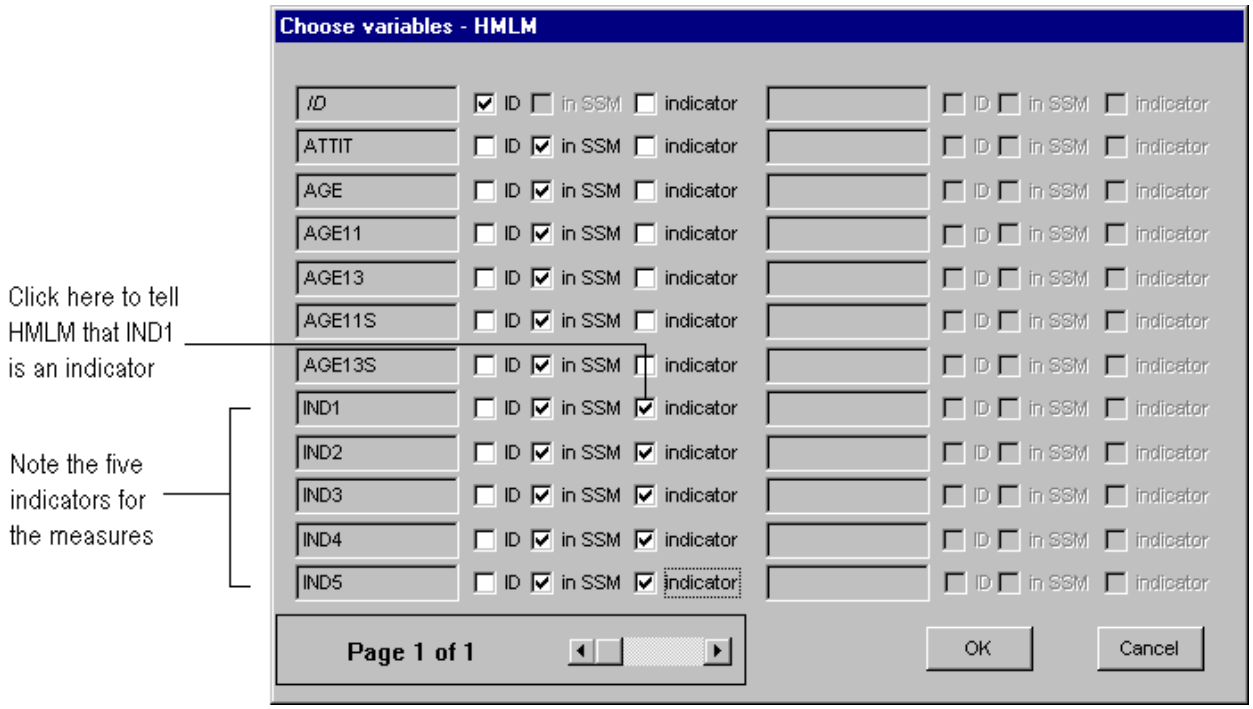
Start by selecting the **File**, **Make new MDM file**, **Stat package data** option from the main menu bar.



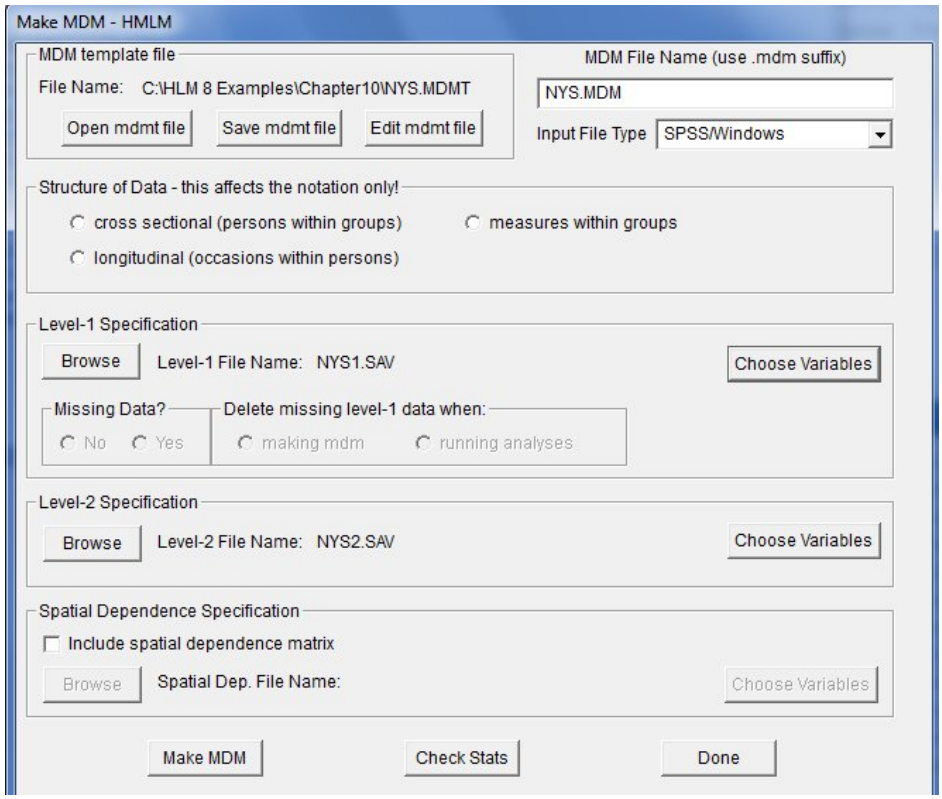
Indicate that the MDM file to be made is for a HMLM model on the **Select MDM type** dialog box.



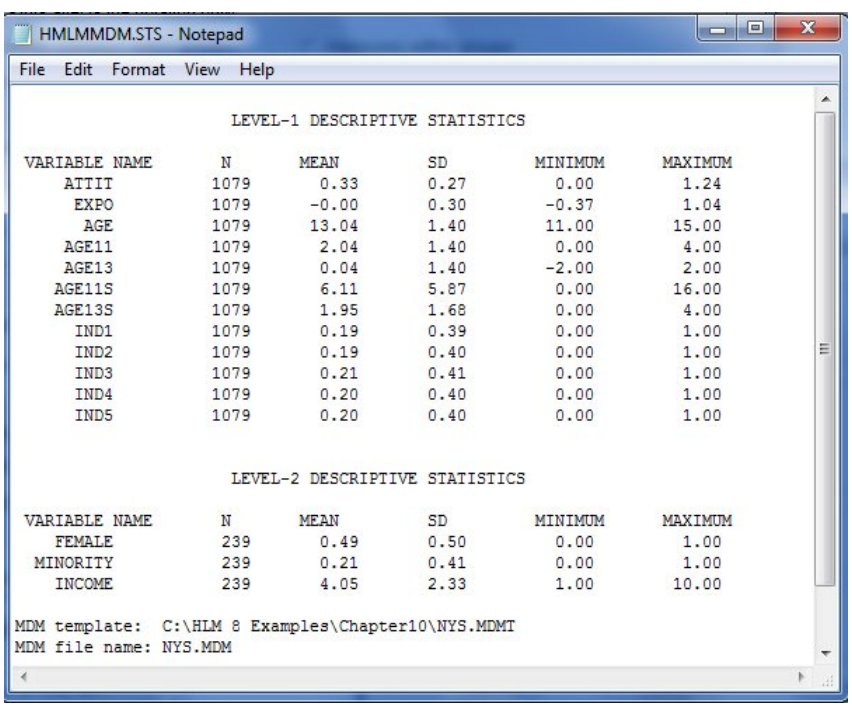
After selecting the type of input to be used (SPSS in this case) browse for the level-1 data file and complete the **Choose variables – HMLM** dialog box as shown below. Note that the 5 indicator variables created prior to importing the data are selected in the **Indicator** column.



Next, browse for the level-2 data file and complete the selection of variables for inclusion at this level. The completed **Make MDM -HMLM** dialog box is shown below. Note that for HMLM models there is no missing data options. As the model itself can be seen as a way to deal with missing data, no provision is made in either HMLM or HMLM2 for missing data in the data files.

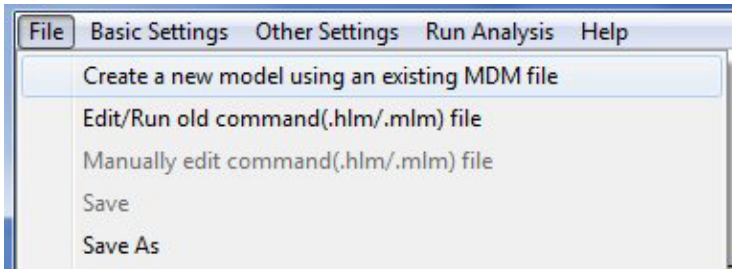


After assigning a name for the MDM file and saving the MDM template (*.**mdmt**) file, click **Make MDM.** After the MDM has been created, the descriptive statistics for the contents will be displayed automatically.

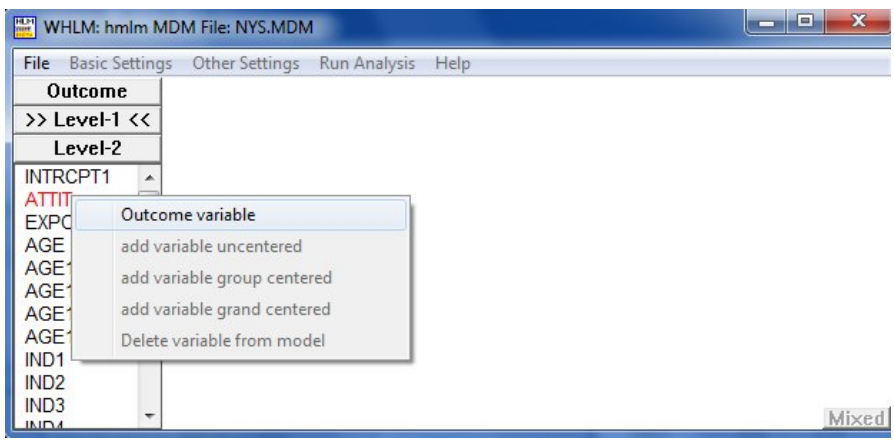


7. Creating the command file

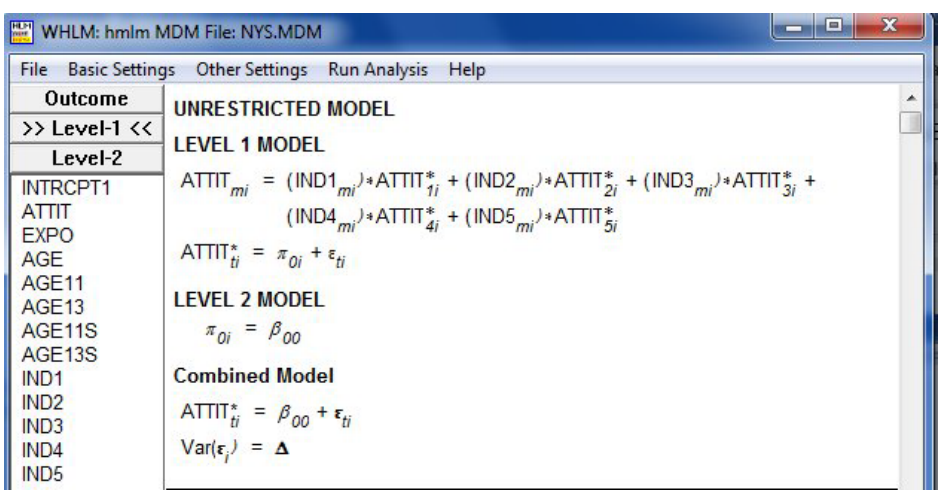
Start by selecting the **Create a new model using an existing MDM file** option from the **File** menu and open the MDM file **NYS.MDM**.



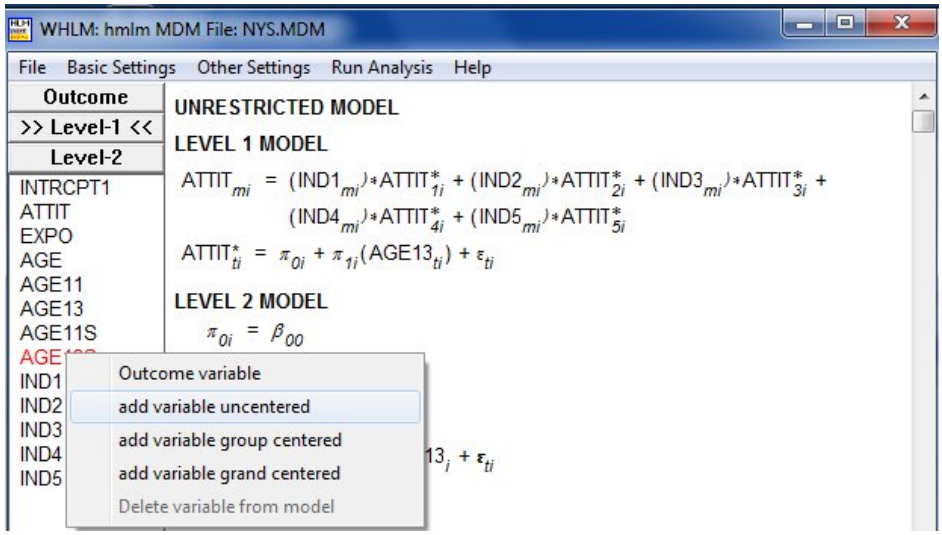
Select the outcome variable ATTIT by clicking on the variable name at left and selecting **Outcome variable** from the pop-up menu.



The unconditional HMLM model appears, as shown below.



Next, add the variables AGE13 and AGE13S to the model by selecting the un**centered** option from the pop-up menu.



After specifying the basic model, click the **Outcome** button at the top of the variable list box to the left of the main HLM window to open the **Basic Model Specifications – HMLM** dialog box. In addition to the unsaturated model described previously in this example, HMLM offers a selection of other ways in which the level-1 variance may be modelled. A brief description of each is given below.

HLM with varying level-1 variance

One can model heterogeneity of level-1 variance as a function of the occasion of measurement. Such a model is suitable when we suspect that the level-1 residual variance varies across occasions. The models that can be estimated are a subset of the models that can be estimated within the standard HLM2 (see Section 2.8.8.2 on the option for heterogeneity of level-1 variance).The level-1 model is the same as in the case of homogenous variances except that now

$$Var(e_i) = \Sigma = diag\left\{\sigma_t^2\right\},$$

that is, Σ is now diagonal with elements σ_t^2, the variance associated with occasion t, $t = 1, \ldots, T$.

The number of parameters estimated is $f + r(r+1)/2 + T$. Now r must be no larger than $T-1$. When $r = T-1$, the results will duplicate those based on the unrestricted model.

HLM with a log-linear model for the level-1 variance

The model with varying level-1 variance, described above, assumes a unique level-1 variance for every occasion. A more parsimonious model would specify a functional relationship between aspects of the occasion (*e.g.* time or age) and the variance. We would again have $\Sigma = diag\left\{\sigma_t^2\right\}$, but now

$$\log\left(\sigma_t^2\right) = \alpha_0 + \sum_{l=1}^{L} \alpha_l c_{lt}.$$

Thus, the natural log of the level-1 variance may be a linear or quadratic function of age. If the explanatory variables c_l are $T-1$ dummy variables, each indicating the occasion of measurement, the results will duplicate those of the previous section.

The number of parameters estimated is now $f + r(r+1)/2 + L + 1$. Again, r must be no larger than $T-1$ and L must be no larger than $T-1$.

First-order auto-regressive model for the level-1 residuals

This model allows the level-1 residuals to be correlated under Markov assumptions (a level-1 residual depends on previous level-1 residuals only through the immediately preceding level-1 residuals). This leads to the level-1 covariance structure

$$Cov(e_{ti}, e_{t'i}) = \sigma^2 \rho^{|t-t'|}.$$

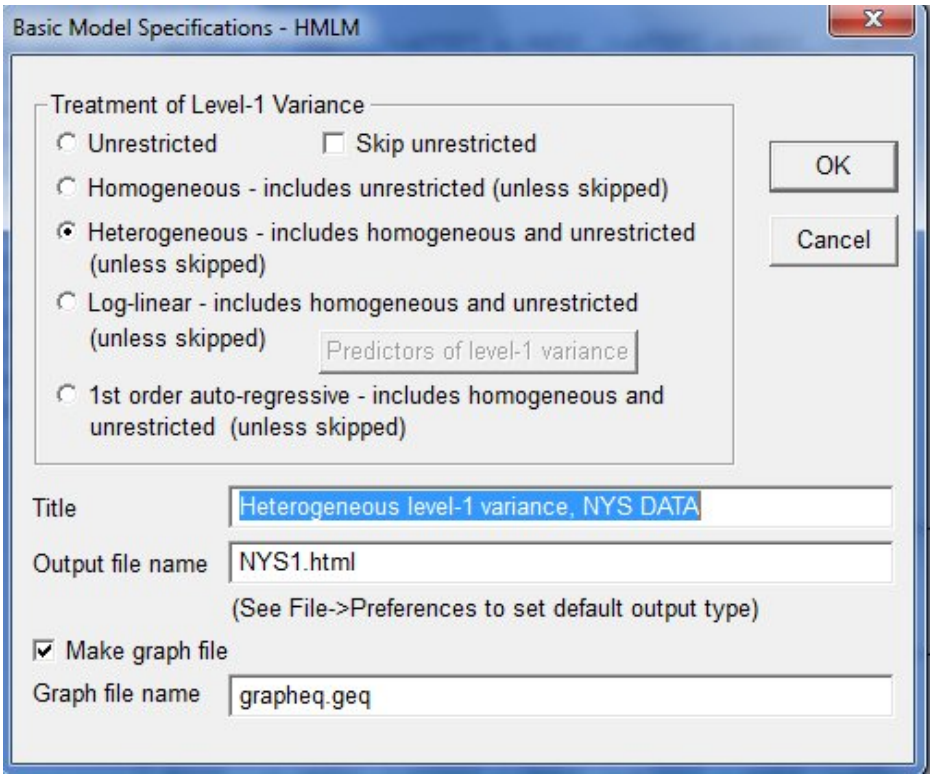
Thus, the variance at each time point is σ^2 and each correlation diminishes with the distance between time points, so that the correlations are $\rho, \rho^2, \rho^3, \ldots$ as the distance between occasions is 1, 2, 3, ... The number of parameters estimated is now $f + r(r+1)/2 + 2$. Again, r must be no larger than $T-1$.

Note that level-1 predictors are assumed to have the same values for all level-2 units of the complete data. This assumption can be relaxed. However, if the design for a_{pti} varies over i, its coefficient cannot vary randomly at level 2. In this regard, the standard 2-level model is more flexible than HMLM.

In this example we specify AGE13 and AGE13S as predictors at level 1. At level 2, the model is unconditional. We shall compare three alternative covariance structures:

- an unrestricted model,
- the homogeneous model, $\sigma_t^2 = \sigma^2$ for all t, and
- the heterogeneous model, which allows σ_t^2 to vary over time.

These three models are requested simply by checking the **Heterogeneous** option in the **Basic Model Specifications – HMLM** dialog box.



Click **OK** to return to the main window and remember to save the command file prior to running the analysis.

8. Interpreting the output

The data source for this run = NYS.MDM
The command file for this run = nys1.hlm
Output file name = nys1.html
The maximum number of level-1 units = 1079
The maximum number of level-2 units = 239
The maximum number of iterations = 100

The outcome variable is ATTIT

The model specified for the fixed effects was:

Level-1 Coefficients	Level-2 Predictors
INTRCPT1, π0	INTRCPT2, β00
# AGE13 slope, π1	INTRCPT2, β10
# AGE13S slope, π2	INTRCPT2, β20

'#' - The residual parameter variance for this level-1 coefficient has been set to zero.

Level-1 Model

$ATTIT_{mi} = (IND1_{mi})*ATTIT_{1i}^* + (IND2_{mi})*ATTIT_{2i}^* + (IND3_{mi})*ATTIT_{3i}^* + (IND4_{mi})*ATTIT_{4i}^* + (IND5_{mi})*ATTIT_{5i}^*$

$ATTIT_{ti}^* = \pi_{0i} + \pi_{1i}*(AGE13_{ti}) + \pi_{2i}*(AGE13S_{ti}) + \varepsilon_{ti}$

The level-1 model relates the observed data, Y, to the complete data, Y^*.

Level-2 Model

$\pi_{0i} = \beta_{00}$
$\pi_{1i} = \beta_{10}$
$\pi_{2i} = \beta_{20}$

For the restricted model, there is no random variation between persons in regression coefficient β_0, β_1, and β_2 because all random variation has been absorbed into Δ.

Var(εi) = Δ

Δ(0)

IND1	0.03507	0.01671	0.01889	0.02149	0.02486
IND2	0.01671	0.04458	0.02779	0.02468	0.02714
IND3	0.01889	0.02779	0.07272	0.05303	0.04801
IND4	0.02149	0.02468	0.05303	0.08574	0.06636
IND5	0.02486	0.02714	0.04801	0.06636	0.08985

The 5×5 matrix Δ contains the maximum likelihood estimates of the five variances (one for each time point) and ten covariances (one for each pair of time points). The associated correlation matrix is printed below.

Standard errors of Δ

IND1	0.00347	0.00304	0.00375	0.00413	0.00429
IND2	0.00304	0.00434	0.00430	0.00457	0.00473
IND3	0.00375	0.00430	0.00678	0.00631	0.00625
IND4	0.00413	0.00457	0.00631	0.00811	0.00736
IND5	0.00429	0.00473	0.00625	0.00736	0.00853

Δ (as correlations)

IND1	1.000	0.423	0.374	0.392	0.443
IND2	0.423	1.000	0.488	0.399	0.429
IND3	0.374	0.488	1.000	0.672	0.594
IND4	0.392	0.399	0.672	1.000	0.756
IND5	0.443	0.429	0.594	0.756	1.000

The 5×5 matrix above contains estimated standard errors for each element of Δ.

The value of the log-likelihood function at iteration 8 = 1.891335E+002

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, π0					
INTRCPT2, β00	0.320244	0.014981	21.377	238	<0.001
For AGE13 slope, π1					
INTRCPT2, β10	0.059335	0.004710	12.598	238	<0.001
For AGE13S slope, π2					
INTRCPT2, β20	0.000330	0.003146	0.105	238	0.917

The expected log attitude at age 13 is 0.320244. The mean linear growth rate of increase is estimated to be 0.059335, t = 12.598, indicating a highly significantly positive average rate of increase in deviant attitude at age 13. The quadratic rate is not statistically significant.

Statistics for the current model

Deviance = -378.266936
Number of estimated parameters = 18

There are 3 fixed effects (f = 3) and five observations in the "complete data" for each person (T = 5). Thus, there are a total of $f + T(T+1)/2 = 3 + 5(5+1)/2 = 18$ parameters. This is the end of the unrestricted model output.

Next follows the results for the homogeneous level-1 variance.

Summary of the model specified

Level-1 Model

$ATTIT_{mi} = (IND1_{mi})*ATTIT_{1i}^{*} + (IND2_{mi})*ATTIT_{2i}^{*} + (IND3_{mi})*ATTIT_{3i}^{*} + (IND4_{mi})*ATTIT_{4i}^{*} + IND5_{mi})*ATTIT_{5i}^{*}$

$ATTIT_{ti}^{*} = \pi_{0i} + \pi_{1i}*(AGE13_{ti}) + \pi_{2i}*(AGE13S_{ti}) + \varepsilon_{ti}$

Level-2 Model

$\pi_{0i} = \beta_{00}$
$\pi_{1i} = \beta_{10}$
$\pi_{2i} = \beta_{20}$

The above equation, written with subscripts and Greek letters, is

$$Var(Y^{*}) = \Delta = ATA' + \Sigma$$

where $\Sigma = \sigma^2 I_T$.

A

IND1	1.00000	-2.00000	4.00000
IND2	1.00000	-1.00000	1.00000
IND3	1.00000	0.00000	0.00000
IND4	1.00000	1.00000	1.00000
IND5	1.00000	2.00000	4.00000

The above matrix describes the design matrix on occasions one through five.

Note: The results below duplicate exactly the results produced by a standard HLM2 run using homogeneous level-1 variance.

Final Results - Iteration 5

	Parameter	Standard Error
σ2	0.02421	0.001672

τ

INTRCPT1,r0	0.04200	0.00808	-0.00242
AGE13,r1	0.00808	0.00277	-0.00012
AGE13S,r2	-0.00242	-0.00012	0.00049

Standard errors of τ

INTRCPT1,r0	0.00513	0.00127	0.00089
AGE13,r1	0.00127	0.00054	0.00024
AGE13S,r2	0.00089	0.00024	0.00025

τ (as correlations)

INTRCPT1,r0	1.000	0.749	-0.532
AGE13,r1	0.749	1.000	-0.101
AGE13S,r2	-0.532	-0.101	1.000

Δ

IND1	0.03536	0.01388	0.01616	0.01801	0.01943
IND2	0.01388	0.04870	0.03150	0.03488	0.03464
IND3	0.01616	0.03150	0.06620	0.04766	0.04849
IND4	0.01801	0.03488	0.04766	0.08056	0.06095
IND5	0.01943	0.03464	0.04849	0.06095	0.09625

The 5×5 matrix above contains the five variance and ten covariance estimates implied by the "homogeneous level-1 variance" model.

Δ (as correlations)

IND1	1.000	0.334	0.334	0.338	0.333
IND2	0.334	1.000	0.555	0.557	0.506
IND3	0.334	0.555	1.000	0.653	0.607
IND4	0.338	0.557	0.653	1.000	0.692
IND5	0.333	0.506	0.607	0.692	1.000

The value of the log-likelihood function at iteration 5 = 1.741132E+002

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, π0					
INTRCPT2, β00	0.327231	0.015306	21.379	238	<0.001
For AGE13 slope, π1					
INTRCPT2, β10	0.064704	0.004926	13.135	238	<0.001
For AGE13S slope, π2					
INTRCPT2, β20	0.000171	0.003218	0.053	238	0.958

Statistics for the current model

Deviance = -348.226421
Number of estimated parameters = 10

There are 3 fixed effects (f = 3); the dimension of τ is 3, and a common σ^2 is estimated at level-1. Thus, there are a total of $f + r(r+1)/2 + 1 = 3 + 3(3+1)/2 = 10$ parameters.

This is the end of the output for the "homogeneous level-1 variance" model. Finally, the heterogeneous level-1 variance solution is listed.

Output for Random Effects Model with Heterogeneous Level-1 Variance

Summary of the model specified

Level-1 Model

$ATTIT_{mi} = (IND1_{mi})*ATTIT_{1i}^* + (IND2_{mi})*ATTIT_{2i}^* + (IND3_{mi})*ATTIT_{3i}^* + (IND4_{mi})*ATTIT_{4i}^* + (IND5_{mi})*ATTIT_{5i}^*$

$ATTIT_{ti}^* = \pi_{0i} + \pi_{1i}*(AGE13_{ti}) + \pi_{2i}*(AGE13S_{ti}) + \varepsilon_{ti}$

Level-2 Model

$\pi_{0i} = \beta_{00} + r_{0i}$
$\pi_{1i} = \beta_{10} + r_{1i}$
$\pi_{2i} = \beta_{20} + r_{2i}$

$Var(\varepsilon_i) = Var(\mathbf{A}\mathbf{r}_i + e_i) = \mathbf{\Delta} = \mathbf{A}\tau\mathbf{A}' + \sigma^2\mathbf{I}$

The above equation, written with subscripts and Greek letters, is

$$Var(Y^*) = ATA' + \Sigma$$

where $\Sigma = diag\left\{\sigma_t^2\right\}$, *i.e.* that is, Σ is now a diagonal matrix with diagonal elements σ_t^2, the variance associated with occasion t, $t = 1, 2, \ldots, T$.

A

IND1	1.00000	-2.00000	4.00000
IND2	1.00000	-1.00000	1.00000
IND3	1.00000	0.00000	0.00000
IND4	1.00000	1.00000	1.00000
IND5	1.00000	2.00000	4.00000

Iterations stopped due to small change in likelihood function

Final Results - Iteration 8

	σ2	Standard Error
IND1	0.01373	0.005672
IND2	0.02600	0.003296
IND3	0.02685	0.003658
IND4	0.02602	0.003633
IND5	0.00275	0.007377

The five estimates above are the estimates of the level-1 variance for each time point.

τ

INTRCPT1,r0	0.04079	0.00736	-0.00241
AGE13,r1	0.00736	0.00382	0.00025
AGE13S,r2	-0.00241	0.00025	0.00106

Standard errors of τ

INTRCPT1,r0	0.00512	0.00124	0.00088
AGE13,r1	0.00124	0.00066	0.00042
AGE13S,r2	0.00088	0.00042	0.00030

τ (as correlations)

INTRCPT1,r0	1.000	0.590	-0.366
AGE13,r1	0.590	1.000	0.124
AGE13S,r2	-0.366	0.124	1.000

Δ

IND1	0.03410	0.01707	0.01646	0.01851	0.02325
IND2	0.01707	0.05165	0.03103	0.03322	0.03223
IND3	0.01646	0.03103	0.06764	0.04574	0.04588
IND4	0.01851	0.03322	0.04574	0.08208	0.06421
IND5	0.02325	0.03223	0.04588	0.06421	0.08996

The 5×5 matrix above contains the estimates of five variances and ten covariances implied by the "heterogeneous level-1 variance" model.

Δ (as correlations)

IND1	1.000	0.407	0.343	0.350	0.420
IND2	0.407	1.000	0.525	0.510	0.473
IND3	0.343	0.525	1.000	0.614	0.588
IND4	0.350	0.510	0.614	1.000	0.747
IND5	0.420	0.473	0.588	0.747	1.000

The value of the log-likelihood function at iteration 8 = 1.816074E+002

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, π0					
INTRCPT2, β00	0.327646	0.015252	21.482	238	<0.001
For AGE13 slope, π1					
INTRCPT2, β10	0.060864	0.004737	12.849	238	<0.001
For AGE13S slope, π2					
INTRCPT2, β20	-0.000541	0.003178	-0.170	238	0.865

Statistics for the current model

Deviance = -363.214879
Number of estimated parameters = 14

There are 3 fixed effects (f = 3), the dimension of τ is 3, and there are five observations intended for each person, each associated with a unique level-1 variance. Thus, there are a total of $f + r(r+1)/2 + T = 3 + 3(4)/2 + 5 = 14$ parameters.

Summary of Model Fit

Model	Number of Parameters	Deviance
1. Unrestricted	18	-378.26694
2. Homogeneous σ2	10	-348.22642
3. Heterogeneous σ2	14	-363.21488

Model Comparison	χ2	d.f.	p-value
Model 1 vs Model 2	30.04052	8	<0.001
Model 1 vs Model 3	15.05206	4	0.005
Model 2 vs Model 3	14.98846	4	0.005

The model deviances are employed to evaluate the fits of the three models (unrestricted, homogeneous σ^2, and heterogeneous σ^2). Differences between deviances are distributed asymptotically as chi-square variates under the null hypothesis that the simpler model fits the data as well as the more complex model does. The results show that Model 1 fits better than does the homogeneous sigma squared model χ^2 = 30.04052, df = 8; it also fits better than does the heterogeneous sigma squared model χ^2 = 15.05206, df = 4.

In addition to the evaluation of models based on their fit to the data, the above results can be used to check the sensitivity of key inferences to alternative specifications of the variance-covariance structure. For instance, one could compare the mean and variance in the rate of change at age 13 obtained in Model 2 and Model 3 to assess how robust the results are to alternative plausible covariance specifications. The mean rate, γ_{10}, for Model 2 is 0.064704 (s.e. = 0.004926), and the variance, τ_{22}, is 0.00277 (s.e. = 0.00054). The mean rate, G10, for Model 3 is 0.060864 (s.e. = 0.004737), and the variance, τ_{22}, is 0.00382 (s.e. = 0.00066). The results are basically similar. See Raudenbush (2001) for a more detailed analysis of alternative covariance structures for polynomial models of individual growth and change using the same NYS data sets employed here for the illustrations.