6

Multiple-Group Growth Modeling

Building on the growth models with time-invariant covariates described in the previous chapter, we now work through an alternative framework for examining between-person differences in change—the multiple-groups framework (McArdle, 1989; McArdle & Hamagami, 1996). Growth models with time-invariant covariates are useful for studying differences in *average* growth trajectories but have limited utility for examining differences in other aspects of the within-person change process and the between-person differences in that process. Without extension, such time-invariant covariate models tell us nothing about differences in the variances and covariance among growth factors, residual variability, and the structure of within-person changes. In this chapter we illustrate how the multiple-groups framework can be used to examine differences in any aspect of the growth model. This flexibility can provide additional insight into how and why individuals differ in their development. Further extensions to search for developmental differences in unmeasured groups follow in Chapter 7 (i.e., growth mixture models).

The additional flexibility in understanding group differences provided by the multiple-groups framework is extremely useful when examining group differences in the amount of variation at both level-1 (measurement) and level-2 (structural). For example, in studying group differences in externalizing behaviors from age 2 through 7, we observed that boys and girls had different amounts of variance at the within (level-1) and between-person (level-2) levels of the growth model. That is, boys showed greater residual variance indicating that their trajectories were more difficult to capture—greater variance remained after accounting for the growth model. In addition, boys also showed greater variability in the intercept and slope, indicating that boys' trajectories were more different from one another than girls' trajectories. Furthermore, the structure of within-person changes varied across groups. That is, boys showed slow linear decline in externalizing behaviors through elementary school, whereas girls showed exponential decline in externalizing behaviors during that same time period. These aspects of group differences are difficult to study using the time-invariant covariate models discussed in Chapter 5, but they can be studied with the multiple-groups growth model discussed here. It must be noted, however, that the multiple-group approach is not recommended for all situations. Typically, it is reserved for inquiry into differences between a relatively small number of groups that are designated by a single categorical variable. In some situations the approach is useful when considering several grouping variables. These situations are accommodated by combining the grouping variables together (e.g., two two-group variables would create a four-group model), but as the net number of groups gets larger, multiple-group models become more and more unwieldy. Continuous variables can always be degraded into ordinal categorical grouping variables, but the reductions in precision and the arbitrary nature of cutoff choices can be difficult to justify. When continuous variables are available, we encourage using them in their originally intended form, if possible.

Example data for this chapter are the same data used in the previous chapters—longitudinal data on mathematics achievement collected from the NLSY-CYA. In Chapter 3, the within-person changes in mathematics from second grade through eighth grade and the between-person differences therein were represented using a linear growth model, with grade as the time metric and the intercept centered at second grade. Here, we use the multiple-groups framework to examine differences between groups defined by birthweight (0 = normal, 1 = low), a medical distinction assigned to children at birth based on whether or not their weight was less than 5.5 pounds.

MULTILEVEL MODELING FRAMEWORK

In the multilevel modeling framework, the multiple-group linear growth model is conceptualized as a collection of two-level models—one model for each group. Within each group, the level-1 model is similar to the linear growth model introduced in Chapter 3, but with an additional group-specific designation given by the superscript (g). Thus, for i = 1 to N individuals in g = 1 to G groups, the model is specified as

$$y_{ti}^{(g)} = b_{1i}^{(g)} + b_{2i}^{(g)} \cdot t + u_{ti}^{(g)} \tag{6.1}$$

where $y_{ti}^{(g)}$ is the repeatedly measured variable collected at time t for individual i in group g, $b_{1i}^{(g)}$ is the random intercept or predicted score for individual i in group g when $t = 0$, $b_{2i}^{(g)}$ is the random slope or rate of change for individual i in group g for a one-unit change in t, t represents the chosen time metric in an appropriate scale—that is, t could be replaced by $(t - k_1)/k_2$, and $u_{ti}^{(g)}$ is the time-specific residual score at time t for individual i in group g. The time-specific residual is assumed to be normally distributed with group-specific variance, $u_{ti}^{(g)} \sim N\left(0, \sigma_u^{2(g)}\right)$. Similarly, the group-specific level-2 equations are written as

$$\begin{aligned} b_{1i}^{(g)} &= \beta_1^{(g)} + d_{1i}^{(g)} \\ b_{2i}^{(g)} &= \beta_2^{(g)} + d_{2i}^{(g)} \end{aligned} \tag{6.2}$$

where $\beta_1^{(g)}$ and $\beta_2^{(g)}$ are group-specific means of the intercept and slope, and $d_{1i}^{(g)}$ and $d_{2i}^{(g)}$ are individual deviations from the respective group-specific mean for individual i in group g. The individual deviations are assumed to follow a multivariate normal distribution in each group with zero means, estimated variances, and a covariance; for example,

$$d_{1i}^{(g)}, d_{2i}^{(g)} \sim MVN\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\sigma_1^{2(g)} & \\ \sigma_{21}^{(g)} & \sigma_2^{2(g)}\end{bmatrix}\right)$$

In essence, the data are split by the grouping variable, and separate growth models are fit to each group simultaneously.

The multilevel model of Equations 6.1 and 6.2 is specified with each parameter of the linear growth model as group specific (each parameter is superscripted by (g)). In practice, variants of this model are fit to articulate and test specific types of group differences. Typically, four models, ordered to allow for more and more group specificity, are estimated and compared. The first model (M1) is the *invariance* model where all estimated parameters are invariant (identical) across groups. This model has the fewest parameters (simplest model) and maps onto the linear growth model described in Chapter 3. One set of parameters is used to describe change, without any group specificity. Parameter estimates from this model should be identical to those obtained from the linear growth model that does not include the grouping variable. With a linear growth trajectory, this model has six estimated parameters: intercept and slope means, intercept and slope variances, intercept–slope covariance, and residual variance—one set for all groups.

The second model (M2) is the *means* model where the latent variable means, β_1 and β_2, are estimated separately for each group, while all other parameters remain invariant across groups. This model follows the logic of commonly utilized statistical models, such as in the *analysis of variance* and *independent samples t-test*, which are used to examine differences in group-level means. Notably, the means model also maps directly onto the growth model with a single dichotomous time-invariant predictor of the intercept and slope (the type of model described in the previous chapter).

The third model (M3) is the *means and covariances* model and allows for group-specific latent variable means, variances, and covariances, while the remaining parameters (i.e., residual variance and linear structure) are held invariant across groups. Finally, the fourth model (M4) is the *means, covariances, and residual variances* model where all estimated parameters are allowed to be group specific. By comparing the relative fit of these four models, we can identify if and how the groups differ from one another with respect to (1) average trajectory, (2) the magnitude of between-person variability and covariability of the growth trajectories, and (3) the magnitude of unexplained within-person variability over time.

MULTILEVEL MODELING IMPLEMENTATION

Although multiple-group models are not often fit in the multilevel modeling framework, the flexibility of `NLMIXED` and `nlme` allows for the direct estimation of these models.

The longitudinal mathematics data are organized in the long format with multiple records per person. The repeatedly measured outcome is the child-level mathematics variable (`math`), and the grouping variable is `lb_wght`, coded 0 for normal-birthweight children and 1 for low-birthweight children.

The `NLMIXED` script for Model M4 (means, covariances, and residual variances) is shown in Script 6.1. Although inverted from the ordering (M1 to M4) outlined above, we present the M4 script because it is the most general. Models M1, M2, and M3 are obtained by making minor changes to this script (often removing elements). As in the previous chapters, the script begins by calling the `NLMIXED` procedure and `nlsy_math_long` dataset. Next, several `IF-THEN` statements are used to subset the data by the grouping variable, `lb_wght`. The statements invoke the general form of Equation 6.2, where the individual intercept (`b_1i`) and slope (`b_2i`) are a function of group-level means (`beta_1_N` and `beta_2_N`, or `beta_1_L` and `beta_2_L`) and deviations (`d_1i_N` and `d_2i_N`, or `d_1i_L` and `d_2i_L`). The letters `N` for *Normal* (`lb_wght = 0`) and `L` for *Low* birthweight (`lb_wght = 1`) are used to distinguish the group-specific parameters. After outlining the level-2 equations, two more `IF-THEN` statements are used to indicate that the residual variance, `v_u`, is group specific. Specifically, the residual variance, `v_u`, is set equal to `v_u_N` for the normal-birthweight

Script 6.1. `NLMIXED` Script for the Means, Covariances, and Residual Variances Model

```
PROC NLMIXED DATA = nlsy_math_long;
    IF lb_wght = 0 THEN b_1i = beta_1_N + d_1i_N;
    IF lb_wght = 0 THEN b_2i = beta_2_N + d_2i_N;

    IF lb_wght = 1 THEN b_1i = beta_1_L + d_1i_L;
    IF lb_wght = 1 THEN b_2i = beta_2_L + d_2i_L;

    IF lb_wght = 0 THEN v_u = v_u_N;
    IF lb_wght = 1 THEN v_u = v_u_L;

    traject = b_1i + b_2i * (grade - 2);

    MODEL math ~ NORMAL (traject, v_u);
    RANDOM d_1i_N d_2i_N d_1i_L d_2i_L ~
                NORMAL ([0,0,0,0], [v_1_N,
                                   c_21_N, v_2_N,
                                        0,      0,  v_1_L,
                                        0,      0, c_21_L, v_2_L])
    SUBJECT = id;
    PARMS
    beta_1_N = 40 beta_2_N = 5 beta_1_L = 4 beta_2_L = 5
    v_1_N = 70    v_2_N = 1    c_21_N = 0
    v_1_L = 70    v_2_L = 1    c_21_L = 0
    v_u_N = 40    v_u_L = 40;
RUN;
```

group and equal to `v_u_L` for the low-birthweight group. Next, the trajectory equation is written without the level-1 residual variance. The `MODEL` statement follows where the outcome, `math`, is assumed to follow a `NORMAL` distribution with mean `traject` and variance `v_u`. Both of these statements are identical to the statements needed for a single-group growth model.

The `RANDOM` line comes next, designating that the four level-2 individual deviations (`d_1i_N`, `d_2i_N` and `d_1i_L`, `d_2i_L`), two for each group, are assumed to follow `NORMAL` distributions with 0 means (`[0,0,0,0]`) and the specified variance–covariance matrix. The variance–covariance matrix contains parameters for the variance of the intercept for each group (`v_1_N` and `v_1_L`), variance of the slope for each group (`v_2_N` and `v_2_L`), and intercept–slope covariance for each group (`c_21_N` and `c_21_L`). Given that no child is in both groups (exclusive coding of `lb_wght`), between-group covariances are fixed at 0. That is, all the across-group random-effects parameters are fixed to 0 by placing zeroes in the appropriate locations (lower left block) of the variance–covariance matrix. Next, the statement `SUBJECT = id` is used to indicate that the random effects are random over individuals. Finally, the `PARMS` statement is used to provide starting values for the 12 (6 per group) estimated parameters.

To estimate Models M1, M2, and M3, this script is slightly modified. Working backwards, for Model M3, the residual variance, `v_u`, should be invariant across groups. Thus, the two `IF-THEN` statements that separate the residual variance are removed. In parallel, the `PARMS` statement is modified by removing the starting values for `v_u_N` and `v_u_L` and providing a starting value for `v_u` (for a total of 11 estimated parameters). Continuing to Model M2, we additionally modify the variance–covariance matrix in the `RANDOM` statement so that the variances and covariances of the individual deviations are invariant across groups. This is achieved by adjusting the parameter labels in the variance–covariance matrix so that they are no longer group-specific. The labels for the intercept variance (`v_1_N` and `v_1_L`), slope variance (`v_2_N` and `v_2_L`), and intercept–slope covariance (`c_21_N` and `c_21_L`) are changed to `v_1`, `v_2`, and `c_21`, respectively. Using the same labels for the normal- and low-birthweight groups forces the latent variable variances and covariance to be equal across groups. Again, the `PARMS` statement must be adjusted to provide starting values for `v_1`, `v_2`, and `c_21` and remove those for `v_1_N`, `v_2_N`, `c_21_N`, `v_1_L`, `v_2_L`, and `c_21_L` (now a total of eight estimated parameters). Finally, Model M1 is specified by modifying the first two sets of `IF-THEN` statements. These statements are modified by replacing the group-specific means, `beta_1_N`, `beta_2_N`, `beta_1_L`, and `beta_2_L`, with sample-level means, `beta_1` and `beta_2` (e.g., `IF lb_wght = 0 THEN b_1i = beta_1 + d_1i_N;` and `IF lb_wght = 1 THEN b_1i = beta_1 + d_1i_L;`). This forces the mean intercept and slope to be the same (invariant) across the two groups. As before, the `PARMS` statement is modified, now providing starting values for the six estimated parameters that remain.

The flexibility of `nlme` also allows for the direct estimation of multiple-group growth models; however, compared with `NLMIXED`, it does have some limitations. As before, the longitudinal data are organized in the long format with multiple records per person.

The outcome variable is the child-level mathematics score (math). The grouping variable is lb_wght, coded 0 for normal-birthweight children and coded 1 for low-birthweight children. Additionally, we introduce another grouping variable nb_wght that is coded in the opposite way of lb_wght. That is, nb_wght is coded 0 for the low-birthweight group and coded 1 for the normal-birthweight group. Thus, lb_wght identifies if the child is in the low-birthweight group, and nb_wght identifies if the child is from the normal-birthweight group and these variables are used to invoke multiple-groups within the nlme package.

The nlme script for Model M4 is shown in Script 6.2. First, we create the nb_wght variable in the nlsy_math_long dataset by reverse coding lb_wght. With the dummy variables now in place, we can proceed to specify the model. An object named mg4.math.nlme is created that will hold the output from the nlme function. The equation for the multiple-group model is written in two parts, each of which is invoked when one of the dummy variables equals 1. In the first part nb_wght is multiplied by a linear growth model with beta_1_N and beta_2_N as fixed effects and d_1i_N and d_2i_N as random coefficients. Since nb_wght is coded 1 for the normal-birthweight group, this part of the model is for the normal-birthweight children. Thus, beta_1_N and beta_2_N are the means of the intercept and slope, and d_1i_N and d_2i_N are the random coefficients for the normal-birthweight group. In the second part of the model, lb_wght is multiplied by a linear growth model, with beta_1_L and beta_2_L as fixed effects and d_1i_L and d_2i_L as random coefficients. Since lb_wght is coded 1 for the low-birthweight group, this part of the model is invoked for low-birthweight children. Thus, beta_1_L and beta_2_L are the means of the intercept and slope, and d_1i_L and d_2i_L are the random coefficients for the low-birthweight group.

The next lines of script specifies the datafile for analysis and lists the fixed and random effects. The fixed command specifies all four fixed effects (mean intercept and slope for each group), and the random command specifies a Blocked structure for

Script 6.2. nlme Script for the Means, Covariances, and Residual Variances Model

```
nlsy_math_long$nb_wght <- 1 - nlsy_math_long$lb_wght

mg4.math.nlme<-nlme(math~nb_wght*((beta_1_N+d_1i_N)+
                              (beta_2_N+d_2i_N)*(grade-2))+
                        lb_wght*((beta_1_L+d_1i_L)+
                              (beta_2_L+d_2i_L)*(grade-2)),
                      data=nlsy_math_long,
                      fixed=beta_1_N+beta_2_N+beta_1_L+beta_2_L~1,
                      random=pdBlocked(list(d_1i_N+d_2i_N~1,
                                            d_1i_L+d_2i_L~1)),
                      group=~id,
                      start=c(35, 4, 35, 4),
                      weights=varIdent(form=~1|factor(lb_wght)),
                      control=(list(returnObject=TRUE)))
summary (mg4.math.nlme)
```

the latent variable covariance matrix. The Blocked structure estimates correlations (or covariances) for random effects within the same block and will fix the across-block correlations to zero. Using a comma-separated list, the first block contains d_1i_N and d_2i_N, the random effects for the normal-birthweight group, and the second block contains d_1i_L and d_2i_L, the random effects for the low-birthweight group. As was done in NLMIXED, the Blocked structure invokes estimation of the correlation between the intercept and slope within each group but does not estimate correlations among the random coefficients from different groups. The group command follows and indicates that the random effects apply over participants, and the start command provides the starting values for the four fixed effects. The weights command is then used to allow for group-specific level-1 residual variances. The varIdent option specifies that the residual variances should be organized as a matrix with no off-diagonal elements and form is used to specify a variance covariate, which allows for a heteroskedastic error structure that differs by lb_wght. Finally, the control command is used to obtain additional ouput that is useful when diagnosing problems with model convergence. Here the command indicates that the parameter estimates obtained at the last iteration should be output, even if convergence issues are encountered. The summary function is used to obtain the output from fitting the model.

Script 6.2 is easily modified to fit Models M1, M2, and M3. For Model M3, the weights command is removed, which forces the residual variance to be equal across the two groups. For Model M2, we make two additional changes. In the model equation, the group-specific deviations (d_1i_N, d_2i_N and d_1i_L, d_2i_L) are replaced by d_1i and d_2i, and the random statement is changed to d_1i+d_2i~1. These changes force the latent variable covariance matrix (i.e., intercept and slope variances and covariance) to be invariant over groups. Continuing to Model M1, two more changes are needed. The model equation is further adjusted so that the group-specific means (beta_1_N, beta_2_N and beta_1_L, beta_2_L) are replaced by beta_1 and beta_2, and the fixed command is changed to beta_1+beta_2~1. Thus, the mean intercept and slope are invariant over groups. The estimated parameters should be the same as those obtained in the single-group model fit in Chapter 3.

Fit statistics and parameter estimates from Model M4 are shown in Output 6.1 and 6.2 for NLMIXED and nlme, respectively. The fit statistics will be used for model comparisons discussed at the end of the chapter. Here we walk through the parameter estimates from Model M4. In both the NLMIXED and nlme output, estimates labeled with an _N are for the normal-birthweight group, and estimates labeled with an _L are for the low-birthweight group. In a few instances, interpretation of the nlme output requires a bit of additional calculation. Overall, parameter estimates describing individual changes in mathematics growth are interpreted in the same manner as discussed in Chapter 3; however, these estimates are now group specific. Thus, the normal-birthweight group had a mean intercept of 35.48 (beta_1_N; Fixed effect of beta_1_N) and a mean linear slope of 4.29 (beta_2_N; Fixed effect of beta_2_N) points per year. The variance of the intercept was 62.13 (v_1_N; Random effect of d_1i_N is given as a standard devation and must be squared), the variance of the slope was 0.78

Output 6.1. `NLMIXED` Output for the Means, Covariances, and Residual Variances Model

Fit Statistics

-2 Log Likelihood	15926
AIC (smaller is better)	15950
AICC (smaller is better)	15950
BIC (smaller is better)	16008

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper	Gradient
beta_1_N	35.4818	0.3644	928	97.38	<.0001	0.05	34.7667	36.1968	0.007959
beta_2_N	4.2976	0.09111	928	47.17	<.0001	0.05	4.1188	4.4764	0.000736
beta_1_L	32.7313	1.3755	928	23.80	<.0001	0.05	30.0319	35.4307	-0.01019
beta_2_L	4.8971	0.3435	928	14.26	<.0001	0.05	4.2229	5.5713	0.019715
v_1_N	62.1335	5.7001	928	10.90	<.0001	0.05	50.9470	73.3201	-0.01076
v_2_N	0.7801	0.3338	928	2.34	0.0196	0.05	0.1250	1.4352	-0.02959
c_21_N	-0.06246	1.1564	928	-0.05	0.9569	0.05	-2.3319	2.2070	-0.01321
v_1_L	73.0535	23.0337	928	3.17	0.0016	0.05	27.8493	118.26	-0.01856
v_2_L	0.009963	0.9575	928	0.01	0.9917	0.05	-1.8691	1.8890	0.139719
c_21_L	0.8507	4.5391	928	0.19	0.8514	0.05	-8.0574	9.7589	-0.0216
v_u_N	35.0704	1.8944	928	18.51	<.0001	0.05	31.3527	38.7882	-0.03962
v_u_L	48.5132	7.5707	928	6.41	<.0001	0.05	33.6556	63.3708	-0.00261

Output 6.2. `nlme` Output for the Means, Covariances, and Residual Variances Model

```
Nonlinear mixed-effects model fit by maximum likelihood
 Model: math~nb_wght*((beta_1_N+d_1i_N)+(beta_2_N+d_2i_N)*
                    (grade-2))+lb_wght*((beta_1_L+d_1i_L)+
                    (beta_2_L+d_2i_L)*(grade-2))
 Data: nlsy_math_long
       AIC       BIC     logLik
  15950.12  16018.59  -7963.061

Random effects:
 Composite Structure: Blocked

 Block 1: d_1i_N, d_2i_N
 Formula: list(d_1i_N ~ 1, d_2i_N ~ 1)
 Level: id
 Structure: General positive-definite
        StdDev     Corr
 d_1i_N 7.8922007  d_1i_N
 d_2i_N 0.8795757  -0.009

Block 2: d_1i_L, d_2i_L
 Formula: list(d_1i_L ~ 1, d_2i_L ~ 1)
 Level: id
 Structure: General positive-definite
          StdDev      Corr
 d_1i_L   8.99408209  d_1i_L
 d_2i_L   0.02847274  0.988
 Residual 5.93062209

Variance function:
 Structure: Different standard deviations per stratum
 Formula: ~1 | factor (lb_wght)
  Parameter estimates:
        0        1
  1.000000  1.170216

Fixed effects: beta_1_N + beta_2_N + beta_1_L + beta_2_L ~ 1
              Value  Std.Error    DF    t-value    p-value
 beta_1_N  35.48131  0.3647427   930   97.27763          0
 beta_2_N   4.29731  0.0901739  1288   47.65582          0
 beta_1_L  32.79780  1.4062468   930   23.32293          0
 beta_2_L   4.87806  0.3382249  1288   14.42253          0

Number of Observations: 2221
Number of Groups: 932
```

(v_2_N; Random effect of d_2i_N), and the intercept–slope covariance was –0.06 (c_21_N; Corr between d_1i_N and d_2i_N given as a correlation). Finally, the level-1 residual variance was 35.07 (v_u_N; Random effect of Residual). In nlme, the residual standard deviation for the normal-birthweight group is calculated by multiplying the reported residual standard deviation (5.93) by the parameter estimate listed under Variance function for the group coded 0 on the lb_wght variable. This parameter was 1.00, which yields a residual standard deviation equal to 5.93 for the normal-birthweight group.

The low-birthweight group had a mean intercept of 32.73 (beta_1_L; Fixed effect of beta_1_L) and a mean linear slope of 4.90 (beta_2_L; Fixed effect of beta_2_L) points per year. In this group, the variance of the intercept was 73.05 (v_1_L; Random effect of d_1i_L), the variance of the slope was 0.01 (v_2_L; Random effect of d_2i_L), the intercept–slope covariance was 0.85 (c_21_L; Corr between d_1i_L and d_2i_L), and the level-1 residual variance was 48.51 (v_u_L; Random effect of Residual multiplied by 1.17). As before, the residual standard deviation for the low-birthweight group is calculated by multiplying the reported residual standard deviation (5.93) by the parameter estimate listed under Variance function for the group coded 1 on the lb_wght variable. This estimate was 1.17, and multiplying these values yields a residual standard deviation equal to 6.94 for the low-birthweight group.

STRUCTURAL EQUATION MODELING FRAMEWORK

There is a long history of modeling multiple-group data in the structural modeling framework (Jöreskog, 1971). In brief, a linear growth model is specified for each group (group-specific model specification), and parameter labels are used to constrain parameters to be equal across groups to test hypotheses regarding group differences in specific aspects of the linear growth model. The multiple-group linear growth model can be written as

$$\mathbf{y}_i^{(g)} = \mathbf{\Lambda}\eta_i^{(g)} + \mathbf{u}_i^{(g)} \tag{6.3}$$

where $\mathbf{y}_i^{(g)}$ is a $T \times 1$ vector of the repeatedly measured observed scores for individual i in group g, T represents the number of repeated assessments based on the chosen time metric, $\mathbf{\Lambda}$ is a $T \times 2$ matrix of factor loadings defining the latent growth factors, $\eta_i^{(g)}$ is a 2×1 vector of latent factor scores for individual i in group g, and $\mathbf{u}_i^{(g)}$ is a $T \times 1$ vector of residual or unique scores for individual i in group g. Note that the $\mathbf{\Lambda}$ matrix is *not* group specific (not superscript by g). Rather, both groups are assumed to follow a linear growth trajectory. In the structural equation modeling framework, it is possible to accommodate different shapes or structures of change by making the $\mathbf{\Lambda}$ matrix group specific (similarly, different functions of change can be specified in the multilevel modeling framework). Because of our focus on the linear growth model, we assume the change pattern is linear and invariant across groups.

As in the previous chapters, the latent factor scores are written as deviations from group-specific means, such that

$$\eta_i^{(g)} = \alpha^{(g)} + \xi_i^{(g)} \tag{6.4}$$

where $\alpha^{(g)}$ is a 2 × 1 vector of latent factor means for group g and $\xi_i^{(g)}$ is a 2 × 1 vector of residual deviations for individual *i* in group g. In this multiple-group setting, the implied population mean vector ($\mu^{(g)}$) and covariance matrix ($\Sigma^{(g)}$) are also group specific. These expectations are

$$\begin{aligned} \mu^{(g)} &= \Lambda\alpha^{(g)} \\ \Sigma^{(g)} &= \Lambda\Psi^{(g)}\Lambda' + \Theta^{(g)} \end{aligned} \tag{6.5}$$

where $\Psi^{(g)}$ is a 2 × 2 latent variable covariance matrix for group g and $\Theta^{(g)}$ is a $T \times T$ residual diagonal covariance matrix for group g.

As in the specification presented for the multilevel modeling framework, all matrices containing estimated parameters of the multiple-group linear growth model are group specific (i.e., $\alpha^{(g)}$, $\Psi^{(g)}$, and $\Theta^{(g)}$). The four models discussed above, M1 to M4, are specified by imposing equality constraints on specific matrices. Model M1 is specified by placing equality constraints on all three matrices containing estimated parameters (i.e., $\alpha^{(g)} = \alpha$, $\Psi^{(g)} = \Psi$ and $\Theta^{(g)} = \Theta$). Model M2 is specified by placing equality constraints on the Ψ and Θ matrices (i.e., $\alpha^{(g)}$, $\Psi^{(g)} = \Psi$, and $\Theta^{(g)} = \Theta$). Model M3 is specified by only placing equality constraints on Θ (i.e., $\alpha^{(g)}$, $\Psi^{(g)}$, and $\Theta^{(g)} = \Theta$). Finally, Model M4 is specified with no equality constraints across groups—all of the estimated parameters of the multiple-group linear growth model are estimated separately for each group (i.e., $\alpha^{(g)}$, $\Psi^{(g)}$, and $\Theta^{(g)}$).

Figure 6.1 is a path diagram of a multiple-group linear growth model with four repeated measurements in $\mathbf{y}_i$. As seen in this figure, there are separate path diagrams for each group (g = 1 and g = 2), each of which contains a linear growth model. Across the group-specific path diagrams, parameter labels that are identical indicate where an equality constraint has been imposed. In contrast, parameter labels that are superscripted by the group number indicate that those parameters are estimated separately for each group. Given the equality constraints presented in this path diagram, this figure represents Model M2 because only the labels for the means of the intercept and linear slope are superscripted by the group number.

STRUCTURAL EQUATION MODELING IMPLEMENTATION

For the structural equation modeling implementation, the longitudinal data are organized in the wide format, with one record per person. In our example, the repeated outcomes of interest are the child-level mathematics scores (`math2` through `math8`) and

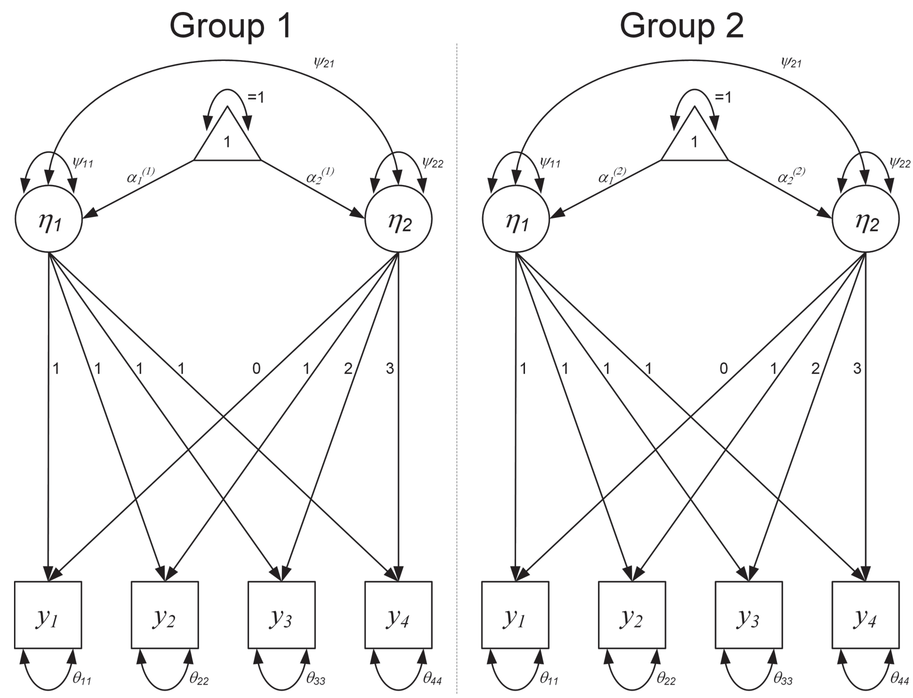

FIGURE 6.1. Path diagram of the multiple-group linear growth model.

the grouping variable is `lb_wght`, which is coded 0 for normal-birthweight children and 1 for low-birthweight children.

The `VARIABLE`, `ANALYSIS`, and `MODEL` statements from the *Mplus* script for fitting Model M4 are shown in Script 6.3. In the `VARIABLE` statement, the repeatedly measured variables are listed on the `USEVAR` line. Additionally, `lb_wght` is specified as the `GROUPING` variable, along with specific mapping between the values of the grouping variable and group-specific labels (0 = `normal`, 1 = `low`). These labels are used in the script to designate the group-specific model statements. As in previous chapters, in the `ANALYSIS` line we specify `TYPE=MEANSTRUCTURE` and the minimum covariance coverage.

The `MODEL` statement is used to specify the linear growth model. This is done in the usual manner, except that labels are included for all estimated parameters. This is done because *Mplus* has several defaults for multiple-group models, some of which are inappropriate for growth models. The labels override some of these defaults by initially imposing equality constraints on all estimated parameters, some of which are subsequently relaxed. The general model is followed by two group-specific model statements utilizing the labels defined in the `GROUPING` statement (`MODEL normal:` and `MODEL low:`). In `MODEL normal:`, the latent variable means (`[eta_1 eta_2];`), latent variable variances (`eta_1 eta_2;`) and covariance (`eta_1 WITH eta_2;`), and the residual variance of observed variables (`math2-math8 (theta1);`) are respecified

Script 6.3. `Mplus` Script for the Means, Covariances, and Residual Variances Model

```
VARIABLE: NAMES = id female lb_wght anti_k1
                  math2-math8 age2-age8 men2-men8
                  spring2-spring8 anti2-anti8;
    MISSING = .;
    USEVAR = math2-math8;
    GROUPING = lb_wght (0 = normal, 1 = low);

ANALYSIS: TYPE=MEANSTRUCTURE; COVERAGE=0;

MODEL:
    eta_1 BY math2-math8@1;
    eta_2 BY math2@0 math3@1 math4@2 math5@3 math6@4 math7@5 math8@6;
    [eta_1](alpha_1); [eta_2] (alpha_2);
    eta_1 (psi_11); eta_2 (psi_22);
    eta_1 WITH eta_2 (psi_21);
    [math2-math8@0];
    math2-math8 (theta);

MODEL normal:
    [eta_1 eta_2];
    eta_1 eta_2;
    eta_1 WITH eta_2;
    math2-math8 (theta1);

MODEL low:
    [eta_1 eta_2];
    eta_1 eta_2;
    eta_1 WITH eta_2;
    math2-math8 (theta2);
```

to make the estimation of these parameters specific to the normal-birthweight group. Similar statements are contained in `MODEL low:` to ensure that the model parameters are estimated separately for each group.[1] The parameters in the group-specific model statements will be estimated separately for each group, unless they are given common labels. Additionally, to constrain the residual variances to be equal across time, but to remain group specific, we give these parameters different labels (i.e., `theta1` and `theta2`) in the group-specific model statements.

Models M1 through M3 are specified by making small adjustments to the M4 script, either by removing select statements from the group-specific models or by adding common labels to parameters in the group-specific models. For example, to obtain Model M3, we can remove the `math2-math8 (theta1);` and `math2-math8 (theta2);`

[1]In two-group models, respecifying the model specification for the second group is unnecessary. Moving to models with more than two groups, having a model specification for each group is important to be able to constrain or free parameters over groups.

statements or we can add a common label to both statements so that the final lines read `math2-math8 (theta);`. To specify Model M2, we remove the lines or add labels to the group-specific designations of the latent variable variances (`eta_1 eta_2;`) and covariance (`eta_1 WITH eta_2;`). Finally, for Model M1, common labels are applied to the means of the intercept and slope in the group-specific model statements. (Removing the statements for the latent variable means in the group-specific model will result in having the first group's mean intercept and slope to be fixed at 0 due to the *Mplus* default specification.)

The `OpenMx` code for specifying Model M4 is shown in Script 6.4. The script begins by creating two datasets with exclusive content, one for the normal-birthweight group and one for the low-birthweight group. These datasets are created using the `subset` command in `R`. The `mathdata_N` dataset contains the observations that belong to the normal-birthweight group (`lb_wght==0`) and `mathdata_L` contains the subset of observations that belong to the low-birthweight group (`lb_wght==1`). Separate datasets are needed to conduct multiple-group analyses in `OpenMx` (as opposed to having all data in one file that includes a group membership variable).

After the two datasets are created, the models are specified separately for each group. First, the linear growth model is specified for the normal-birthweight group using the `mathdata_N` dataset. The model specification follows the single-group specification of the linear growth model described in Chapter 3 using multiple `mxPath` statements to specify the model. `mxPath` statements for the residual variances of the observed scores, the intercept–slope variances and covariance, factor loadings for the intercept and slope, and means for the intercept and slope are given. Importantly, the labels given to each parameter are group specific. For the normal-birthweight group all parameter labels end with `_N` (e.g., `th_N` for the residual variance). After closing the first group-specific model, the same linear growth model is specified for the low-birthweight group using the `mathdata_L` dataset and group-specific labels that end with `_L`. The group specificity of the labels is important because it is this designation that allows parameter estimates to be separately estimated for each group. After closing the second group-specific model, `mxAlgebra` and `mxFitFunctionAlgebra` statements are used to specify the overall objective function as the sum of the group-specific objective functions, and then the full model is closed. Finally, `mxRun` and `summary` functions are used to fit the model and display the results.

As with the other implementations, Models M1 to M3 are specified by making small adjustments to this script. Here, these modifications are implemented by using common labels for specific parameters to constrain them to be invariant across groups. For example, to specify Model M3, the labels for the group-specific residual variances (`th_N` and `th_L`) are replaced by a common label (i.e., `th`). Similarly, Model M2 is specified by also replacing the group-specific labels for the latent variable variances and covariance by a set of common labels (i.e., `psi_11`, `psi_22`, and `psi_21`). Finally, Model M1 is specified by making the labels for the latent variable means the same across groups (i.e., `alpha_1` and `alpha_2`).

The *Mplus* and `OpenMx` results from Model M4 are shown in Output 6.3 and 6.4, respectively. Before discussing parameter estimates and model fit, we note that the *Mplus*

Script 6.4. OpenMx Script for the Means, Covariances, and Residual Variances Model

```
mathdata_N <- subset(nlsy_math_wide,lb_wght==0)
mathdata_L <- subset(nlsy_math_wide,lb_wght==1)

mg4.math.omx <- mxModel('Multiple-group Growth Model, Means, Covariances, Residuals, Path Specification',
 mxModel('group1',type='RAM', mxData(observed=mathdata_N, type='raw'),
 manifestVars=c('math2','math3','math4','math5','math6','math7','math8'),
 latentVars=c('eta_1','eta_2'),

# residual variance paths
mxPath(from=c('math2','math3','math4','math5','math6','math7','math8'),
 arrows=2, free=TRUE, values=60, labels='th_N'),

# latent variable variances and covariance paths
mxPath(from=c('eta_1','eta_2'), connect='unique.pairs',
 arrows=2, free=TRUE, values=c(60,0,5), labels=c('psi_11_N','psi_21_N','psi_22_N')),

# Factor Loadings
mxPath(from='eta_1', to=c('math2','math3','math4','math5','math6','math7','math8'),
 arrows=1, free=FALSE, values=1),

mxPath(from='eta_2', to=c('math2','math3','math4','math5','math6','math7','math8'),
 arrows=1, free=FALSE, values=c(0, 1, 2, 3, 4, 5, 6)),

# means and intercepts
mxPath(from='one', to=c('eta_1','eta_2'),
 arrows=1, free=c(TRUE,TRUE), values=c(100, 15), labels=c('alpha_1_N','alpha_2_N'))

), # close Normal Birthweight Model
```

```
mxModel('group2', type='RAM', mxData(observed=mathdata_L, type ='raw'),
 manifestVars=c('math2','math3','math4','math5','math6','math7','math8'),
 latentVars=c('eta_1','eta_2'),

# residual variance paths
mxPath(from=c('math2','math3','math4','math5','math6','math7','math8'),
 arrows=2, free=TRUE, values=60, labels='th_L'),

# latent variable variances and covariance paths
mxPath(from=c('eta_1','eta_2'), connect='unique.pairs',
 arrows=2, free=TRUE, values=c(60,0,5), labels=c('psi_11_L','psi_21_L','psi_22_L')),

# Factor Loadings
mxPath(from='eta_1', to=c('math2','math3','math4','math5','math6','math7','math8'),
 arrows=1, free=FALSE, values=1),

mxPath(from='eta_2', to=c('math2','math3','math4','math5','math6','math7','math8'),
 arrows=1, free=FALSE, values=c(0, 1, 2, 3, 4, 5, 6)),

# means and intercepts
mxPath(from='one', to=c('eta_1','eta_2'),
 arrows=1, free=c(TRUE,TRUE), values=c(100, 15), labels=c('alpha_1_L','alpha_2_L'))

), # Close Low Birthweight Model

mxAlgebra(group1.objective + group2.objective, name='mg_objective'), mxFitFunctionAlgebra('mg_objective')

) # Close Model

mg4.math.fit <- mxRun(mg4.math.omx)

summary(mg4.math.fit)
```

Output 6.3. Mplus Output for the Means, Covariances, and Residual Variances Model

```
THE MODEL ESTIMATION TERMINATED NORMALLY
      WARNING: THE LATENT VARIABLE COVARIANCE MATRIX (PSI) IN GROUP
      LOW IS NOT POSITIVE DEFINITE. THIS COULD INDICATE A NEGATIVE
      VARIANCE/RESIDUAL VARIANCE FOR A LATENT VARIABLE, A CORRELATION
      GREATER OR EQUAL TO ONE BETWEEN TWO LATENT VARIABLES, OR A LINEAR
      DEPENDENCY AMONG MORE THAN TWO LATENT VARIABLES. CHECK THE TECH4
      OUTPUT FOR MORE INFORMATION. PROBLEM INVOLVING VARIABLE ETA_2.

MODEL FIT INFORMATION
Number of Free Parameters                    12
Loglikelihood
          H0 Value                    -7963.056
          H1 Value                    -7840.549
Information Criteria
          Akaike (AIC)                15950.111
          Bayesian (BIC)              16008.159
          Sample-Size Adjusted BIC    15970.048
          (n* = (n + 2) / 24)
Chi-Square Test of Model Fit
          Value                         245.013
          Degrees of Freedom                 57
          P-Value                        0.0000
Chi-Square Contributions From Each Group
          NORMAL                        190.639
          LOW                            54.374
RMSEA (Root Mean Square Error Of Approximation)
          Estimate                        0.084
          90 Percent C.I.                 0.073     0.095
          Probability RMSEA <= .05        0.000
CFI/TLI
          CFI                             0.780
          TLI                             0.842

MODEL RESULTS

                                                      Two-Tailed
           Estimate     S.E.   Est./S.E.       P-Value
Group NORMAL
ETA_1    BY
    MATH2     1.000    0.000     999.000       999.000
    ...
    MATH8     1.000    0.000     999.000       999.000
ETA_2    BY
    MATH2     0.000    0.000     999.000       999.000
    ...
    MATH8     6.000    0.000     999.000       999.000
```

```
ETA_1    WITH
    ETA_2    -0.063    1.161      -0.054         0.957
Means
    ETA_1    35.481    0.365      97.257         0.000
    ETA_2     4.297    0.091      47.145         0.000
Variances
    ETA_1    62.287    5.729      10.873         0.000
    ETA_2     0.774    0.334       2.314         0.021
Residual Variances
    MATH2    35.173    1.904      18.473         0.000
    ...
    MATH8    35.173    1.904      18.473         0.000

Group LOW
ETA_1    BY
    MATH2     1.000    0.000     999.000       999.000
    ...
    MATH8     1.000    0.000     999.000       999.000
ETA_2    BY
    MATH2     0.000    0.000     999.000       999.000
    ...
    MATH8     6.000    0.000     999.000       999.000
ETA_1    WITH
    ETA_2     0.746    5.522       0.135         0.892
Means
    ETA_1    32.800    1.407      23.314         0.000
    ETA_2     4.873    0.341      14.298         0.000
Variances
    ETA_1    79.626   25.630       3.107         0.002
    ETA_2    -0.158    1.477      -0.107         0.915
Residual Variances
    MATH2    48.689    8.445       5.766         0.000
    ...
    MATH8    48.689    8.445       5.766         0.000
```

output begins with a warning about the estimation of the latent variable covariance matrix (`PSI`) for the low-birthweight group and highlights variable `ETA_2` as the problem. As we will see when parsing through the output, the estimated variance of the linear slope for the low-birthweight group was small and negative (–0.16). Given that variances cannot be less than zero, the error message and this parameter estimate indicate that the model was not viable. Thus, in practice, the parameter estimates should not be interpreted. We do describe the output here for pedagogical reasons, given that readers may encounter either similar error messages or viable output.

Parameter estimates from M*plus* and `OpenMx` are organized according to group. M*plus* reports all fixed and estimated parameters, first for `Group NORMAL` and then for `Group LOW`. Output from `OpenMx` includes only the estimated parameters with designations starting with `group1.` or `group2.` in the `matrix` column to indicate their group association. We note that parameter estimates constrained to be equal across

Output 6.4. OpenMx Output for the Means, Covariances, and Residual Variances Model

```
free parameters:
         name    matrix    row     col      Estimate     Std.Error
1        th_N  group1.S  math2   math2   35.17197839    1.87799257
2    psi_11_N  group1.S  eta_1   eta_1   62.28975865    5.58685757
3    psi_21_N  group1.S  eta_1   eta_2   -0.06320182    1.09799941
4    psi_22_N  group1.S  eta_2   eta_2    0.77373272    0.32227432
5   alpha_1_N  group1.M      1   eta_1   35.48131985    0.36482093
6   alpha_2_N  group1.M      1   eta_2    4.29730928    0.09112965
7        th_L  group2.S  math2   math2   48.68704569    8.39569515
8    psi_11_L  group2.S  eta_1   eta_1   79.62008119   25.35468150
9    psi_21_L  group2.S  eta_1   eta_2    0.74648044    5.41729561
10   psi_22_L  group2.S  eta_2   eta_2   -0.15728435    1.45729012
11  alpha_1_L  group2.M      1   eta_1   32.79999194    1.40680570
12  alpha_2_L  group2.M      1   eta_2    4.87300524    0.34077354

observed statistics: 2221
estimated parameters: 12
degrees of freedom: 2209
-2 log likelihood: 15926.11
saturated -2 log likelihood: NA
number of observations: 933
chi-square: NA
p: NA
Information Criteria:
      df Penalty Parameters Penalty   Sample-Size Adjusted
AIC:   11508.1112          15950.11                     NA
BIC:     820.0741          16008.17               15970.06
```

groups will only appear once, with a `group1.` designation. Parameter estimates describing the growth process in mathematics from second through eighth grade for the normal-birthweight group (`Group NORMAL`; `group1.`) include the mean intercept (35.48; `Means of ETA_1`; `alpha_1_N`) and slope (4.30; `Means of ETA_2`; `alpha_2_N`), variance of the intercept (62.29; `Variances of ETA_1`; `psi_11_N`) and slope (0.77; `Variances of ETA_2`; `psi_22_N`), intercept–slope covariance (–0.06; `ETA_1 WITH ETA_2`; `psi_21_N`), and residual variance (35.17; `Residual Variances of MATH2` through `MATH8`; `th_N`). That is, on average, normal-birthweight children had a mathematics score of 35.48 in second grade and increased 4.30 points per year from second through eighth grades. Although there were significant between-child differences in both the intercept and slope, these individual differences were not significantly associated with one another.

In parallel, parameter estimates describing the growth process in mathematics from second through eighth grades for the low-birthweight group (`Group_LOW`; `group2.`) include the mean intercept (32.80; `Means of ETA_1`; `alpha_1_L`) and slope (4.87; `Means of ETA_2`; `alpha_2_L`), variance of the intercept (79.63; `Variances of ETA_1`; `psi_11_L`) and slope (–0.16; `Variances of ETA_2`; `psi_22_L`), intercept–slope covariance (0.75; `ETA_1 WITH ETA_2`; `psi_21_L`), and residual variance (48.69; `Residual Variances` of `MATH2` through `MATH8`; `th_L`). On average, low-birthweight children had a mathematics score of 32.80 in second grade, and their scores increased 4.87 points per year from second through eighth grade. Keeping in mind the problem with model convergence noted earlier, one can see significant between-child differences in mathematics scores in second grade but not in the annual rate of change.

IMPORTANT CONSIDERATIONS

Model Comparisons

Contrasted with the inclusion of time-invariant predictors covered in the previous chapter, the multiple-groups framework provides some additional opportunities for examining between-person differences in change. By comparing the relative fit of the four fitted models, M1 to M4, we can identify if and how the groups differed from one another with respect to their means (M2), means and covariances (M3), and means, covariances, and residuals (M4). Because the models are *nested* (e.g., M1 is a constrained version of M4), model comparisons can be done using likelihood ratio tests (LRTs). To aid our comparisons, likelihood-based fit indices, along with the $-2LL$, are organized in Table 6.1. First, we compare the fit of Model M2 to that of Model M1. This comparison examines whether the groups differ in their average trajectories. Changes in the $-2LL$ (or χ^2) are χ^2 distributed, with degrees of freedom equal to the difference in the number of estimated parameters. Thus, examining the M1 and M2 columns of Table 6.1, we find that the two additional parameters (group-specific vs. group-invariant means) in Model M2 led to a reduction in the $-2LL$ of 5(16,937 − 16,932), a nonsignificant ($\chi^2(2) = 5$, $p = .08$) improvement in model fit. The increase in the BIC and the minor change in the AIC confirm the conclusion that Model M2 did not fit significantly better than Model M1. That is, the groups did not significantly differ in their average trajectories. Second, we compared the fit of Model M3 to that of Model M2. Again, the LRT was nonsignificant ($\chi^2(3) = 3$, $p = .39$) indicating that the latent variable covariance matrix was invariant over groups. Finally, Model M4 was compared with Model M3 to determine whether the groups differed in their residual variability. Again, the LRT was nonsignificant ($\chi^2(1) = 3$, $p = .08$) indicating that the residual variance was not significantly different over groups. In summary, the sequential model fitting and comparison indicated that the four models fit similarly, which suggests that the invariance model (Model M1) was the most parsimonious model that represents the observed data with nearly the same precision as the three multiple-group models that allowed for differences in the means, variances, and covariances, and residual variance.[2] Thus, we conclude that the normal- and low-birthweight

[2]In a similar way, Model M3 and Model M4 can also be compared directly to Model M1. Comparing Model M3 to Model M1, we conclude that, together, the differences in means and covariances are nonsignificant ($\chi^2(5) = 8$, $p = .16$), and comparing Model M4 to Model M1 we conclude that all the differences in means, covariances, and residual variance are nonsignificant ($\chi^2(6) = 11$, $p = .09$).

TABLE 6.1. Fit Statistics for the Multiple-Group Linear Growth Models

	Model M1	Model M2	Model M3	Model M4
Parameters	6	8	11	12
$-2LL$	15937	15932	15929	15926
BIC	15978	15987	16005	16008
AIC	15949	15948	15951	15950
Δ parameters	–	2	3	1
$\Delta -2LL$	–	5	3	3

children did not differ in their average growth trajectories of mathematics, the extent of between-person differences in those trajectories, or the extent of fluctuation around their individual trajectories.

It may be noted that the conclusions about birthweight-related differences in the growth trajectories for mathematics reached here, in the multiple-groups framework, were not congruent with the conclusions reached in the previous chapter when birthweight group was used as a time-invariant covariate. In the time-invariant covariate analysis, low-birthweight children were found to have lower mathematics scores in second grade, while in the multiple-group analysis no differences were noted. This is because of two reasons. First, in the multiple-group models we did not control for antisocial behaviors. Second, the statistical tests were conducted in different ways. Here, in the multiple-groups framework, we tested for differences in the average trajectory, operationalized as a two degrees of freedom test of differences in the mean of both the intercept and the slope. In the time-invariant covariate models, we conducted separate one degree of freedom tests of the difference in the mean of the intercept and difference in the mean of the slope. The discrepancy between a global test of all latent factor means versus a series of one-parameter-at-a-time tests might be considered a limitation of the multiple-group approach. Single-parameter tests can be implemented in the multiple-groups framework (e.g., Model M1.5 can be specified where only the mean of the intercept differs between groups), with the additional flexibility that a wide variety of intermediate models can be used to test group differences. The key issue is more a philosophical one—Do the intercept and slope components of change hold specific interpretive value in isolation, or are they only meaningful in bulk as multiple components of a trajectory?

Ordering of Model Comparisons

There are two possible ways to structure the comparisons among the four multiple-group models, M1 to M4. One approach is to fit and compare the models in *ascending* order. That is, begin with Model M1, the most constrained model, and proceed to Models M2, M3, and finally M4, the model with the fewest constraints. In this approach, the first model in the sequence has the smallest number of parameters and the greatest misfit, in terms of the size of the $-2LL$ and χ^2, and the last model in the sequence has the greatest number of parameters and the smallest misfit. Advantages of this sequencing include the following: (1) the simplest models (in terms of the number of estimated parameters) are fit first, (2) the second model in the sequence has direct parallels with the time-invariant covariate model and thus provides a common starting point for describing between-person differences in within-person change, and (3) the sequence generally follows the expansion of ANOVA (assuming equal variances followed by unequal variances). The alternative approach is to fit and compare models in *descending* order, beginning with the most relaxed model, the means, covariances, and residual variances model, and then gradually imposing constraints in sequence through the models—to M3, M2, and then M1. This approach follows the sequencing most often used when studying factorial invariance (Meredith, 1993) over groups or time. No matter the sequence, the same conclusions should be drawn.

MOVING FORWARD

The multiple-group approach to studying between-person differences in within-person change is powerful, even more so than we have covered. Thus far, we have limited our presentation to linear growth models for describing within-person change and between-person differences in change. In later chapters, we move to more complex nonlinear models and highlight the potential for studying group differences in the shape of change over time—an examination that is not available when working with models that include a time-invariant covariate as a predictor of the intercept and slope. For example, in some situations we may expect one group to follow a linear growth trajectory (e.g., control group), whereas another group would follow exponential growth (e.g., intervention group). As soon as we push beyond hypotheses about differences in the magnitude of within-person change trajectories, and consider the possibility that groups of individuals may follow different within-person change trajectories, the usefulness of the multiple-groups framework is magnified.

We have presented the multiple-groups framework as an alternative to the time-invariant covariates approach; however, the two frameworks can be integrated. As outlined in the previous chapter, time-invariant covariates are used to explain between-person differences in the intercept and slope. Adding time-invariant covariates into the multiple-groups framework allows us to explain variability in the intercept and slope within each group. We can then test whether the relations between the time-invariant covariates and the intercept and slope (regression parameters) differ across groups. Models where the regression parameters are invariant across groups allow for main effects of time-invariant covariates. Models where the regression parameters differ across groups examine how the time-invariant covariate and the grouping variable interact to affect individuals' change trajectories.

The multiple-groups framework is the foundation for understanding latent class growth models (Nagin, 1999) and growth mixture models (Muthén & Shedden, 1999). These models have received considerable attention in the last decade as an additional way of examining between-person differences in within-person change. Yet, they are not so different from the models just covered. The key distinction between these models and the multiple-group model is whether or not the grouping variable is known *a priori* (Ram & Grimm, 2009). In the multiple-groups framework, the grouping variable is known ahead of time. It is in the datafile, and it can be used to identify the cases that belong to each group. Group-specific models are used to understand how the groups differ in their growth trajectories. In contrast, for both the latent class growth model and the growth mixture model, the grouping variable is not known *a priori*. Rather, the grouping variable is inferred from differences manifested in the growth trajectories. In brief, these models attempt to recover the grouping variable by pulling apart the sample in ways that maximize between-group or between-class differences and minimize within-group or within-class differences. In the next chapter, we follow the approach taken here and illustrate how unknown or latent classes can differ in various aspects of the change trajectory—latent variable means, latent variable covariances, and residual variances (e.g., Ram & Grimm, 2009).