Multivariate and Shared Parameter Mixed-Effects Models for Intensive Longitudinal Data

Donald Hedeker
Department of Public Health Sciences
University of Chicago

Juned Siddique
Department of Preventive Medicine
Northwestern University

Xingruo Zhang
Department of Public Health Sciences
University of Chicago

Bonnie Spring
Department of Preventive Medicine
Northwestern University

[1] AUTHORS' NOTE: The authors thank the Volume Editors Donald E.K. Martin, Arni S.R. Srinivasa Rao, C.R. Rao, and an anonymous reviewer for helpful and constructive comments on an earlier version of this chapter. Preparation of this chapter was supported by National Institute of Diabetes and Digestive and Kidney Diseases, Grant Award Number: R01 DK125414.

Chapter 1

Multivariate and Shared Parameter Mixed-Effects Models for Intensive Longitudinal Data

ABSTRACT
Intensive longitudinal data are increasingly common in many research areas. Such data are often collected using Ecological Momentary Assessment (EMA), Experience Sampling Methods (ESM), internet studies, wearable devices, and diary methods. In such studies, typically there are a large number of repeated observations per subject, and often many variables are measured. In this chapter, we first describe a multivariate mixed-effects model that simultaneously considers several dependent variables as joint outcomes. This model includes several random subject effects for each outcome, and considers these random effects to be correlated. In this way, one can assess the correlation of the outcomes conditionally adjusting for model covariates. Furthermore, the multivariate model allows one to test if covariates have the same or different effects on the outcomes. We then describe how the covariance parameters of the random effects can be reformulated as regression effects, leading to a shared parameter modeling of the joint outcomes. This then allows one to consider interactions with the regression versions of the covariance parameters. For example, one can examine whether the association of the outcomes varies by other model covariates (e.g., subject sex or age). Furthermore, these association parameters can interact with each other (e.g., does the association of two outcomes vary as a function of another outcome). This flexibility of the shared parameter approach is highlighted and offers data analysts the possibility of considering novel research questions for intensive longitudinal data. To illustrate these approaches, we use a study investigating weight loss, in which subjects provided daily weight measurements over a treatment intervention period of 3 months and a follow-up period of 9 months. Since subjects have varying numbers of weight measurements in both the treatment and follow-up periods, the use of the mixed model, which does not assume complete data across time, is attractive. In supplemental materials, we present syntax for both frequentist and Bayesian approaches to estimate the parameters of such models using standard statistical software packages.

KEYWORDS
random effects, joint models, irregular measurement

1.1 INTRODUCTION

Modern data collection procedures, such as ecological momentary assessments (EMA) (Stone & Shiffman 1994, Smyth & Stone 2003, Shiffman et al. 2008), experience sampling (de Vries 1992, Scollon et al. 2003, Feldman Barrett & Barrett 2001), and diary methods (Bolger et al. 2003), have been developed to record the momentary events and experiences of subjects in daily life. These procedures yield relatively large numbers of subjects and observations per subject, and data from these designs are sometimes referred to as intensive longitudinal data (Walls & Schafer 2006). Often, these studies include many variables measured over time and there is interest in modeling several variables jointly. In this chapter, we describe a multivariate mixed-effects model that simultaneously considers several dependent variables as joint outcomes. This model includes several random subject effects for each outcome, and considers these random effects to be correlated. We also consider a shared parameter model for the multivariate outcomes that allows further examination of the correlations of the outcomes, by moving the random effect covariance parameters to be regression coefficients in the multivariate model. This allows one to examine interactions with the random effects, including interactions of the random effects with each other.

1.2 MULTIVARIATE MIXED-EFFECTS MODELS

Here, we present a description of a bivariate model, as the methods extend in a logical way for more than two outcomes. For this, Thiébaut et al. (2002) described a practical way of using mixed model software for bivariate outcomes, which we will follow. Specifically, define the dependent variable vector of the two outcomes (considered jointly), $\boldsymbol{y}_i^{(1)}$ (with $n_i^{(1)}$ observations) and $\boldsymbol{y}_i^{(2)}$ (with $n_i^{(2)}$ observations), for subject i ($i = 1, 2, \ldots, N$ subjects) as:

$$\boldsymbol{y}_i = \begin{bmatrix} \boldsymbol{y}_i^{(1)} \\ \boldsymbol{y}_i^{(2)} \end{bmatrix} . \tag{1.1}$$

Here, if the two outcomes are measured at the same timepoints, then $n_i^{(1)} = n_i^{(2)} = n_i$, however this is not required. Thus, subjects can provide more/less numbers of observations on these two outcomes. Notice that the dependent variable vector $\boldsymbol{y}_i$, which stacks the two outcome vectors $\boldsymbol{y}_i^{(1)}$ and $\boldsymbol{y}_i^{(2)}$ on top of each other, is of size $(n_i^{(1)} + n_i^{(2)}) \times 1$.

Similarly, define the covariate matrix for the two outcomes of subject i as:

$$\boldsymbol{X}_i = \begin{bmatrix} \boldsymbol{X}_i^{(1)} & 0 \\ 0 & \boldsymbol{X}_i^{(2)} \end{bmatrix} , \tag{1.2}$$

where $\boldsymbol{X}_i^{(1)}$ is a $n_i^{(1)} \times p_i^{(1)}$ matrix of covariates for the first outcome, and $\boldsymbol{X}_i^{(2)}$ is a $n_i^{(2)} \times p_i^{(2)}$ matrix of covariates for the second outcome. The first column of these two matrices would generally be ones for the intercepts, and the remaining columns would contain the $p^{(1)} - 1$ and $p^{(2)} - 1$ covariates, respectively. These covariates could be time-invariant or time-varying, and include dummy-codes, contrast variables, polynomials, and interactions, for example. In some cases these covariate matrices for the two outcomes could be the same, however this is not required.

Models for longitudinal data generally include multiple random subject effects (*e.g.*, random intercepts and slopes). For this, define the random effect design matrix for the two outcomes of subject i as:

$$\boldsymbol{Z}_i = \begin{bmatrix} \boldsymbol{Z}_i^{(1)} & 0 \\ 0 & \boldsymbol{Z}_i^{(2)} \end{bmatrix}, \tag{1.3}$$

where $\boldsymbol{Z}_i^{(1)}$ is a $n_i^{(1)} \times r_i^{(1)}$ matrix of random effect variables for the first outcome, and $\boldsymbol{Z}_i^{(2)}$ is a $n_i^{(2)} \times r_i^{(2)}$ matrix of random effect variables for the second outcome. As an example, for a random intercept and time model, the first columns of $\boldsymbol{Z}_i^{(1)}$ and $\boldsymbol{Z}_i^{(2)}$ would consist of ones, and the second columns would be the values of time at the $n_i^{(1)}$ and $n_i^{(2)}$ timepoints, respectively. If the two dependent variables were measured at the same timepoints, then generally $\boldsymbol{Z}_i^{(1)} = \boldsymbol{Z}_i^{(2)}$, however, again, this is not required. Also, if time is included as a random effect in these matrices, it would usually also be a variable in the covariate matrices $\boldsymbol{X}_i^{(1)}$ and $\boldsymbol{X}_i^{(2)}$. In this way, the random time effects represent subject deviations from the population trends.

Putting this together, we now have the following bivariate mixed model for subject i:

$$\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{\upsilon}_i + \boldsymbol{\epsilon}_i , \tag{1.4}$$

where $\boldsymbol{\beta}$ are the $(p^{(1)} + p^{(2)}) \times 1$ vector of regression coefficients, $\boldsymbol{\upsilon}_i$ are the $(r^{(1)} + r^{(2)}) \times 1$ vector of random subject effects, and $\boldsymbol{\epsilon}_i$ is the $(n_i^{(1)} + n_i^{(2)}) \times 1$ error vector. The population distribution of the random effects is assumed to be a normal distribution with zero mean and variance-covariance matrix $\boldsymbol{\Sigma}_\upsilon$. The errors $\boldsymbol{\epsilon}_i$ are also assumed to be normally distributed in the population with zero mean and variance covariance matrix $\boldsymbol{\Sigma}_\epsilon$, and independent of the random effects. Here, $\boldsymbol{\Sigma}_\upsilon$ represents the between-subject (BS) (co)variance, and $\boldsymbol{\Sigma}_\epsilon$ is the within-subject (WS) (co)variance.

Suppose that each dependent variable had two random effects, an intercept and a time trend, then the BS (co)variance matrix would be:

$$\boldsymbol{\Sigma}_\upsilon = \begin{bmatrix} \sigma^2_{\upsilon_0^{(1)}} & \sigma_{\upsilon_0^{(1)}\upsilon_1^{(1)}} & \sigma_{\upsilon_0^{(1)}\upsilon_0^{(2)}} & \sigma_{\upsilon_0^{(1)}\upsilon_1^{(2)}} \\ \sigma_{\upsilon_0^{(1)}\upsilon_1^{(1)}} & \sigma^2_{\upsilon_1^{(1)}} & \sigma_{\upsilon_1^{(1)}\upsilon_0^{(2)}} & \sigma_{\upsilon_1^{(1)}\upsilon_1^{(2)}} \\ \sigma_{\upsilon_0^{(1)}\upsilon_0^{(2)}} & \sigma_{\upsilon_1^{(1)}\upsilon_0^{(2)}} & \sigma^2_{\upsilon_0^{(2)}} & \sigma_{\upsilon_0^{(2)}\upsilon_1^{(2)}} \\ \sigma_{\upsilon_0^{(1)}\upsilon_1^{(2)}} & \sigma_{\upsilon_1^{(1)}\upsilon_1^{(2)}} & \sigma_{\upsilon_0^{(2)}\upsilon_1^{(2)}} & \sigma^2_{\upsilon_1^{(2)}} \end{bmatrix}, \tag{1.5}$$

where $\upsilon_0^{(1)}$ and $\upsilon_1^{(1)}$ denote the (co)variance parameters for the intercepts and time trends of the first outcome, and $\upsilon_0^{(2)}$ and $\upsilon_1^{(2)}$ denote the (co)variance parameters for the intercepts and time trends of the second outcome. Often the covariances of the two time trends will be of interest. For example, $\sigma_{\upsilon_0^{(1)}\upsilon_0^{(2)}}$ would represent the association of the intercepts of the two variables, and $\sigma_{\upsilon_1^{(1)}\upsilon_1^{(2)}}$ would represent the association of the time trends of the two variables. These covariance parameters highlight an advantage of the bivariate model. Namely, it allows one to assess whether there is a significant association of the time trends of the two dependent variables.

The form of the WS (co)variance matrix $\boldsymbol{\Sigma}_\epsilon$ depends on what is being assumed about the correlation of the errors of the two dependent variables. The simplest assumption would be that these are independent errors, in which case it would be:

$$\boldsymbol{\Sigma}_\epsilon = \begin{bmatrix} \sigma^2_{\epsilon^{(1)}} \boldsymbol{I}_{n_i^{(1)}} & 0 \\ 0 & \sigma^2_{\epsilon^{(2)}} \boldsymbol{I}_{n_i^{(2)}} \end{bmatrix}, \tag{1.6}$$

where $\boldsymbol{I}_{n_i^{(1)}}$ is an identity matrix of size $n_i^{(1)} \times n_i^{(1)}$, and $\boldsymbol{I}_{n_i^{(2)}}$ is an identity matrix of size $n_i^{(2)} \times n_i^{(2)}$. In this case, there are simply two error (co)variance parameters, one variance for each dependent variable, $\sigma^2_{\epsilon^{(1)}}$ and $\sigma^2_{\epsilon^{(2)}}$. This would assume that there is no residual association both between and within the repeated observations of the two variables. This is an assumption of conditional independence, meaning that the random effects are accounting for all of the correlation both within and between the repeated observations of the two variables. This could be relaxed by allowing for autocorrelated errors for the within-variable residual correlation, and by allowing for cross-variable association parameters for the between-variable residual correlation. It is often the case in modeling a single longitudinal outcome that the assumption of conditional independence for the residual correlation is made, though not always (Chi & Reinsel 1989, Hedeker 1989). In terms of the between-variable association, if the two dependent variables are measured at the same timepoints, then this association would likely be important to include. In other words, there would likely be an association of the errors of the two variables at the same timepoints. Alternatively, if the two dependent variables are not measured at the same timepoints, then the timing of the two variables does not coincide and so the errors would generally be treated as independent of each other.

1.3 EXAMPLE

Data for the analyses reported here come from a longitudinal weight loss management study (Pfammatter et al. 2019), consisting of an active treatment period of approximately 3 months and a follow-up period of approximately 9 months.

The outcome that we will focus on is the self-reported daily weight (lbs) of the subject. Here, we restrict our analyses to those subjects that provided at least 2 daily weight reports during both the active and the follow-up periods. Our interest is in modeling the weights from both periods in a bivariate model to examine whether there is an association in the subject random effects from these two periods, and to examine for possible group, time, and interaction effects. The primary weight assessments were made at baseline, 3, 6, and 12 months. Here, we will focus instead on the daily self reported weight data that subjects' provided beyond these primary weight assessments. In all, there were 196 subjects with 12,404 and 8,299 daily observations from the active and follow-up periods, respectively. Across all of the 20,703 observations, the average weight equaled 203.45lbs (range = 136.47 to 323.20, quartiles = 178.57, 197.76, 226.42).

The timing variable `Date` represented the number of days past the study start for a given subject that the measurement was made (range = 0 to 389). We converted this to the variable `Mon` (i.e., month) by dividing `Date` by 30. Thus, our timing variable will represent weight change per 30 days. Subjects varied in terms of both the number of weight reports and the timing of their weight reports in both periods. During the 3-month active period, the average number of observations per subject was approximately 63 (range = 2 to 85, quartiles = 56.5, 71, and 77), while in the 9-month follow-up period the average was approximately 42 (range = 2 to 268, quartiles = 4, 11, and 54). For each subject, we calculated the average value of the `Mon` variable for both periods. The average of these subject averages in the active period was approximately 1.34 (range = 0.08 to 2.47, quartiles = 1.28, 1.37, and 1.42), while in the follow-up period the average was 5.18 (range = 2.85 to 11.2, quartiles = 3.10, 4.77, and 7.06). Note that in a more traditional longitudinal study in which the timing of the measurements is fixed (e.g., baseline, 3, 6, and 12 months), all subjects would have the same average for the timing variable if all subjects were measured at all timepoints, and nearly the same average of the timing variable if there are some missing data. Here, however, there was a fair degree of variation in these averages across subjects, indicating that subjects varied to some degree in terms of when their measurements were made in both periods. Figure 1.1 provides spaghetti plots of the data in the active and follow-up periods, respectively. Here, a random sample of 10 subjects (i.e., N=10) was selected so that the plots are not overly crowded.

****

Insert Figure 1.1 about here

****

As can be seen, there are more data in the active period, and a great deal of subject heterogeneity in the weight measurements across time in both the active period and the follow-up period. The duration of the follow-up period also differs significantly by subject. In terms of time trends, it is not clear if there are any significant changes in weight over time.

We created centered versions of our timing variable `Mon` for each of these two periods (active and follow-up). In this way, the model intercepts will represent the grand mean across time. Furthermore, since the average values of our timing variable `Mon` varied across subjects, we decomposed the timing effect in terms of its between-subjects (BS) and within-subjects (WS) parts (Neuhaus & Kalbfleisch 1998, van de Pol & Wright 2009), and did this separately for each of the two periods (active and follow-up). Namely, for a time-varying variable X_{ij} (for subject i at timepoint j), we can express the variable as $X_{ij} = \bar{X}_i + (X_{ij} - \bar{X}_i)$. The first part ($\bar{X}_i$) represents the BS component of the variable, and the second part ($X_{ij} - \bar{X}_i$) is the WS component of the variable. In terms of our timing variable `Mon`, the BS effect would express whether subjects that are measured, on average, earlier or later in the period have lower/higher average weights in the period. The WS effect would indicate for a given subject whether there is weight change across time during the period. In the equations below, the variable denoted as `Mon_bs` is the BS version, and `Mon_ws` is the WS version. Again, this was done separately for the two periods.

1.3.1 Bivariate Mixed Model

Denoting the active period variables and parameters by the superscript $^{(A)}$ and the follow-up period variables and parameters by the superscript $^{(F)}$, our bivariate mixed model (for subject i at timepoint j) is given by:

$$\texttt{wt}_{ij}^{(A)} = \beta_0^{(A)} + \beta_1^{(A)}\texttt{Mon\_ws}_{ij}^{(A)} + \beta_2^{(A)}\texttt{Mon\_bs}_i^{(A)} + \upsilon_{0i}^{(A)} + \upsilon_{1i}^{(A)}\texttt{Mon\_ws}_{ij}^{(A)} + \epsilon_{ij}^{(A)} , \tag{1.7}$$

$$\texttt{wt}_{ij}^{(F)} = \beta_0^{(F)} + \beta_1^{(F)}\texttt{Mon\_ws}_{ij}^{(F)} + \beta_2^{(F)}\texttt{Mon\_bs}_i^{(F)} + \upsilon_{0i}^{(F)} + \upsilon_{1i}^{(F)}\texttt{Mon\_ws}_{ij}^{(F)} + \epsilon_{ij}^{(F)} . \tag{1.8}$$

Here, with our centering of the timing variable, the intercepts ($\beta_0^{(A)}$ and $\beta_0^{(F)}$) represent the average weights for the two periods (for subjects with average values of the BS timing variable `Mon_bs`). Our linear time effects ($\beta_1^{(A)}$ and $\beta_1^{(F)}$) represent the average weight changes (per month) for the two periods. The BS effects ($\beta_2^{(A)}$ and $\beta_2^{(F)}$) express whether the average timing of the measurements for subjects relates to their average weight during each of the two periods. The random effects allow heterogeneity across subjects in both of the time trend parameters (i.e., intercepts and slopes) for both periods. Of particular interest is whether there are cross-period associations of these random effects. In other words, is a subject's weight trajectory during the active period related to their weight trajectory during the follow-up period. The covariance parameters of the random-effects (co)variance matrix $\boldsymbol{\Sigma}_v$ will help us to address this.

It should be noted that we are using a simple linear trend for the effect of time. One might argue that this is overly simplistic and opt for more extended time effects like polynomials (e.g., linear, quadratic, cubic, etc.), splines, or non-linear time effects. However, we are interested here in expressing whether subjects' gained, lost, or essentially remained the same in terms of weight across time. Thus, the objective is not to model weight change across time in a more precise manner, but to have model coefficients that summarize weight change in a basic and understandable way.

Table 1.1 list the results of this bivariate mixed model.

****

Insert Table 1.1 about here

****

During the active period, the grand mean is estimated as 207.23 pounds, and is highly significant. The test is that this parameter equals 0, which would be impossible, so the test for the grand mean is not especially interesting. The estimated linear trend is -2.85 ($p < .0001$), which indicates that the average weight loss per month was 2.85 pounds during the active period. The BS timing effect is also significant ($\hat{\beta} = 3.43, p = .012$), indicating that subjects with average measurements later in the active period had greater average weight during this period. The effect is estimated to be 3.43 pounds per month. Turning to the regression effects during the follow-up period, we see that the grand mean is estimated to be 202.88 pounds, and highly significant (again, this test is not especially interesting). The linear trend during the follow-up period is estimated to be -0.22, which suggests weight loss of approximately one-fifth of a pound per month during the follow-up period, however, this is not statistically significant at the 0.05 level. Thus, we can conclude that there is significant weight loss during the active period, but not during the follow-up period.

Turning to the random effect variances, there is considerable subject heterogeneity in the trend parameters of both the active and follow-up periods. This agrees with the visual impression from the spaghetti plots in Figure 1.1. Thus, subjects vary in their overall weight in both periods, and subjects also vary in their trends over time in both periods. Table 1.1 expresses the covariances of the random effects as correlations. In terms of these correlations of the random effects, the correlation of the intercepts for the two periods is very high ($\hat{r} = 0.982$). With the centering of our time variable, the intercepts represent grand means, and so this indicates that subjects' grand mean from the active period is highly correlated with subjects' grand mean from the follow-up period. Additionally, there is a significant correlation between the active period linear trend and the follow-up grand mean ($\hat{r} = 0.187$). Thus, if a subject lost less/more weight during the active period, then they had overall higher/lower weight during the follow-up period. The correlation of the linear time trends from the two periods is also positive ($\hat{r} = 0.115$), however, not statistically significant.

1.3.2 Bivariate Shared Parameter Mixed Model

We now consider a shared parameter model in which the cross-period covariances are reformulated as regression effects. Specifically, the model for the weight outcomes in the two periods is:

$$\texttt{wt}_{ij}^{(A)} = \beta_0^{(A)} + \beta_1^{(A)}\texttt{Mon\_ws}_{ij}^{(A)} + \beta_2^{(A)}\texttt{Mon\_bs}_i^{(A)} + \upsilon_{0i}^{(A)} + \upsilon_{1i}^{(A)}\texttt{Mon\_ws}_{ij}^{(A)} + \epsilon_{ij}^{(A)} , \tag{1.9}$$

$$\begin{aligned}\texttt{wt}_{ij}^{(F)} = {} & \beta_0^{(F)} + \beta_1^{(F)}\texttt{Mon\_ws}_{ij}^{(F)} + \beta_2^{(F)}\texttt{Mon\_bs}_i^{(F)} + \beta_3^{(F)}\upsilon_{0i}^{(A)} + \beta_4^{(F)}\upsilon_{1i}^{(A)} \\ & + \beta_5^{(F)}(\upsilon_{0i}^{(A)} \times \texttt{Mon\_ws}_{ij}^{(F)}) + \beta_6^{(F)}(\upsilon_{1i}^{(A)} \times \texttt{Mon\_ws}_{ij}^{(F)}) \\ & + \upsilon_{0i}^{(F)} + \upsilon_{1i}^{(F)}\texttt{Mon\_ws}_{ij}^{(F)} + \epsilon_{ij}^{(F)} . \end{aligned} \tag{1.10}$$

Notice that the active period random effects ($\upsilon_{0i}^{(A)}$ and $\upsilon_{1i}^{(A)}$) and their interactions with time ($\texttt{Mon\_ws}_{ij}^{(F)}$) are included as regressors in the model for the follow-up weights ($\texttt{wt}_{ij}^{(F)}$). Since these random effects are included in both models, the term "shared-parameter" is given to this bivariate model. Here, we can assess whether a subject's active period time trend (both intercept and slope) influence their weight in the follow-up period, both as main effects and as interactions with time. As a result, the cross-period covariance parameters are set to zero in the (co)variance matrix of the random effects:

$$\boldsymbol{\Sigma}_\upsilon = \begin{bmatrix} \sigma^2_{\upsilon_0^{(A)}} & \sigma_{\upsilon_0^{(A)}\upsilon_1^{(A)}} & 0 & 0 \\ \sigma_{\upsilon_0^{(A)}\upsilon_1^{(A)}} & \sigma^2_{\upsilon_1^{(A)}} & 0 & 0 \\ 0 & 0 & \sigma^2_{\upsilon_0^{(F)}} & \sigma_{\upsilon_0^{(F)}\upsilon_1^{(F)}} \\ 0 & 0 & \sigma_{\upsilon_0^{(F)}\upsilon_1^{(F)}} & \sigma^2_{\upsilon_1^{(F)}} \end{bmatrix} . \tag{1.11}$$

The cross-period associations are now captured by the regression coefficients $\beta_3^{(F)}$, $\beta_4^{(F)}$, $\beta_5^{(F)}$, and $\beta_6^{(F)}$. Note that the total number of parameters is the same as in the previous bivariate mixed model.

The results for this bivariate shared parameter mixed model are given in Table 1.2.

****

Insert Table 1.2 about here

****

Though not listed in the tables, the model deviance ($-2\log L$ value) is exactly the same as the bivariate mixed model, namely $-2\log L = 91085$. Thus, the bivariate shared parameter mixed model is simply a reparameterized version of the bivariate mixed model. Additionally, comparing the estimates in Tables 1.1 and 1.2, one sees that many of the parameter estimates are identical or near-identical, though not all. Interestingly, the random effect intercept variance for the follow-up weight model is dramatically reduced (from 1206.74 to 22.95). To understand this, it is helpful to write the model for $\mathtt{wt}_{ij}^{(F)}$ in its multilevel representation (Goldstein 2011):

level-1

$$\mathtt{wt}_{ij}^{(F)} = b_{0i}^{(F)} + b_{1i}^{(F)}\mathtt{Mon\_ws}_{ij}^{(F)} + \epsilon_{ij}^{(F)} , \tag{1.12}$$

level-2

$$b_{0i}^{(F)} = \beta_0^{(F)} + \beta_2^{(F)}\mathtt{Mon\_bs}_i^{(F)} + \left[\beta_3^{(F)}\upsilon_{0i}^{(A)} + \beta_4^{(F)}\upsilon_{1i}^{(A)}\right] + \upsilon_{0i}^{(F)} , \tag{1.13}$$

$$b_{1i}^{(F)} = \beta_1^{(F)} + \left[\beta_5^{(F)}\upsilon_{0i}^{(A)} + \beta_6^{(F)}\upsilon_{1i}^{(A)}\right] + \upsilon_{1i}^{(F)} . \tag{1.14}$$

Here, level-1 is for the time-varying outcome and any time-varying predictors, and level-2 is for subject-varying predictors. In the level-1 equation, the subject intercept is $b_{0i}^{(F)}$ and the slope (effect of $\mathtt{Mon\_ws}_{ij}^{(F)}$) is $b_{1i}^{(F)}$. These are then explained in terms of subject-varying variables in the level-2 equations, where the terms in brackets are the shared random effects from the active phase. Note that these would not be in the level-2 equations for the bivariate mixed model, but are now included in the bivariate shared parameter mixed model. Thus, the multilevel representation reveals that a subject's intercept not only has $\mathtt{Mon\_bs}_i^{(F)}$ as a predictor in the bivariate mixed model, but also includes the active phase random effects $\upsilon_{0i}^{(A)}$ and $\upsilon_{1i}^{(A)}$ as predictors in the bivariate shared parameter mixed model. Similarly, for the equation of a subject's slope, though the latter does not include $\mathtt{Mon\_bs}_i^{(F)}$ (since we are not considering a $\mathtt{Mon\_bs}_i^{(F)}$ by $\mathtt{Mon\_ws}_{ij}^{(F)}$ interaction). In addition to the predictors, both level-2 equations include the follow-up phase random effects $\upsilon_0^{(F)}$ and $\upsilon_1^{(F)}$, which can be thought of as level-2 residuals (Goldstein 2011). Thus, to the extent that these additional predictors are important, the residual level-2 variances are reduced. This is precisely why the intercept variance for the follow-up weight model is so much lower in the shared parameter model. The active phase random effects are explaining much of the heterogeneity across subjects in their follow-up weight intercept (i.e., grand mean). As can be seen in Table 1.2, the active phase random effects are highly statistically significant in the follow-up weight model ($\hat{\beta}_3^{(F)} = 0.996, \hat{\beta}_4^{(F)} = 1.598$, both with $p = 0.0001$), whereas their interactions with time are not statistically significant ($\hat{\beta}_5^{(F)} = 0.0006, p = 0.87; \hat{\beta}_6^{(F)} = 0.061, p = 0.24$). The conclusions are the same as what was found in the bivariate mixed model, namely (1) subjects' grand mean from the active period is positively associated with subjects' grand mean in the follow-up period, and (2) subjects' time trend during the active period is also positively associated with subjects' grand mean in the follow-up period.

Another interesting difference is that the correlation of the two random effects in the follow-up period is now highly significant ($\hat{r} = 0.597, p = 0.0001$), whereas this same correlation was not significant in the bivariate mixed model ($\hat{r} = 0.116, p = 0.17$). Again, the multilevel representation of the model for $\texttt{wt}_{ij}^{(F)}$ helps to make sense of this. In the bivariate mixed model, the random intercept was only conditional on $\texttt{Mon\_bs}_i^{(F)}$, and the random slope was not conditional on any subject covariates. As noted, in the bivariate shared parameter mixed model, both are conditional on the active phase random effects. Thus, the highly significant positive correlation in the latter model is more akin to a partial correlation that has partialled out these additional effects. The positive nature of this association means that subjects with higher/lower grand means have more positive/negative time trends, after controlling for the active phase random effects.

Table 1.3 provides estimates obtained using the Bayesian software program STAN.

****

Insert Table 1.3 about here

****

Sample code for this is provided in the Appendix B. The results are based on 1500 post-warm-up iterations of four chains and a maximum tree depth of 20. Since non-informative priors were used, the Bayesian estimates are nearly identical to the maximum likelihood estimates. Notice that instead of p-values, the 95% credible intervals are provided in the table. These indicate that, given the data, there is a 95% probability that the true estimate lies within the interval.

1.3.3 Bivariate Shared Parameter Mixed Model with Interactions

An advantage of the shared parameter model is that we can now consider interactions of the shared random effects with other variables. Thus, for example, we can examine the degree to which a variable might moderate the association between the active and follow-up period random effects. In the current case, the variable we will consider is `Cond`, which denotes the condition that a subject was randomized to (0 = app only, 1 = app + coaching). The augmented model, including Cond and Cond interactions, is given below. Because $\texttt{Cond}_i$ does not change across the two periods, there is no need for a superscript on this subject-level variable.

$$\begin{aligned}\texttt{wt}_{ij}^{(A)} &= \beta_0^{(A)} + \beta_1^{(A)}\texttt{Mon\_ws}_{ij}^{(A)} + \beta_2^{(A)}\texttt{Mon\_bs}_i^{(A)} \\ &+ \beta_3^{(A)}\texttt{Cond}_i + \beta_4^{(A)}(\texttt{Cond}_i \times \texttt{Mon\_ws}_{ij}^{(A)}) + \beta_5^{(A)}(\texttt{Cond}_i \times \texttt{Mon\_bs}_i^{(A)}) \\ &+ \upsilon_{0i}^{(A)} + \upsilon_{1i}^{(A)}\texttt{Mon\_ws}_{ij}^{(A)} + \epsilon_{ij}^{(A)} \ , \end{aligned} \tag{1.15}$$

$$\begin{aligned}\texttt{wt}_{ij}^{(F)} &= \beta_0^{(F)} + \beta_1^{(F)}\texttt{Mon\_ws}_{ij}^{(F)} + \beta_2^{(F)}\texttt{Mon\_bs}_i^{(A)} \\ &+ \beta_3^{(F)}\texttt{Cond}_i + \beta_4^{(F)}(\texttt{Cond}_i \times \texttt{Mon\_ws}_{ij}^{(F)}) + \beta_5^{(F)}(\texttt{Cond}_i \times \texttt{Mon\_bs}_i^{(F)}) \\ &+ \beta_6^{(F)}\upsilon_{0i}^{(A)} + \beta_7^{(F)}\upsilon_{1i}^{(A)} + \beta_8^{(F)}(\upsilon_{0i}^{(A)} \times \texttt{Mon\_ws}_{ij}^{(F)}) + \beta_9^{(F)}(\upsilon_{1i}^{(A)} \times \texttt{Mon\_ws}_{ij}^{(F)}) \\ &+ \beta_{10}^{(F)}(\texttt{Cond}_i \times \upsilon_{0i}^{(A)}) + \beta_{11}^{(F)}(\texttt{Cond}_i \times \upsilon_{1i}^{(A)}) \\ &+ \beta_{12}^{(F)}(\texttt{Cond}_i \times \upsilon_{0i}^{(A)} \times \texttt{Mon\_ws}_{ij}^{(F)}) + \beta_{13}^{(F)}(\texttt{Cond}_i \times \upsilon_{1i}^{(A)} \times \texttt{Mon\_ws}_{ij}^{(F)}) \\ &+ \upsilon_{0i}^{(F)} + \upsilon_{1i}^{(F)}\texttt{Mon\_ws}_{ij}^{(F)} + \epsilon_{ij}^{(F)} \ . \end{aligned} \tag{1.16}$$

Here, we are including $\texttt{Cond}_i$, $\texttt{Cond}_i$ by $\texttt{Mon\_ws}_{ij}$, and $\texttt{Cond}_i$ by $\texttt{Mon\_bs}_i$ in both models (i.e., six additional parameters). Also, in the follow-up model, we are interacting $\texttt{Cond}_i$ with the four shared parameter terms we introduced in the bivariate shared parameter mixed model: (1) active random intercept $\upsilon_{0i}^{(A)}$, (2) active random slope $\upsilon_{1i}^{(A)}$, active random intercept by time $\upsilon_{0i}^{(A)} \times \texttt{Mon\_ws}_{ij}^{(F)}$, and (4) active random slope by time $\upsilon_{1i}^{(A)} \times \texttt{Mon\_ws}_{ij}^{(F)}$. Thus, there is a total of 10 additional parameters in this model, relative to the previous one.

Table 1.4 provides the results of this extended bivariate shared parameter mixed model, which has a model deviance equal to 91065.

****

Insert Table 1.4 about here

****

The inclusion of these 10 additional parameters provides improved model fit as indicated by a likelihood ratio test ($\chi^2_{10} = 91085 - 91065 = 20, p < 0.03$). Turning to the individual tests of these 10 parameters, only the Cond by Mon_ws during the active phase ($\hat{\beta}_4^{(A)} = -1.38, p = 0.0006$) and the Cond by act1 by Mon_ws follow-up phase ($\hat{\beta} = 0.236, p = 0.02$) interactions are seen to be statistically significant at the 0.05 level. The first indicates a greater weight loss during the active phase of 1.38 pounds per month for the group randomized to receive App + Coaching, relative to the group that received App only. The estimate for the App only group is weight loss of 2.13 pounds per month (p = 0.0001), and so the estimated weight loss for the App + Coaching group is 1.38 + 2.13 = 3.51 pounds per month. Turning to the second significant result (`Cond` by act1 by `Mon_ws` follow-up phase), let us first consider the act1 by `Mon_ws` follow-up phase result. This interaction represents the association of the active period time trend with the follow-up period time trend for the group that received App only. Note that this parameter is estimated to be approximately zero and non-significant ($\hat{\beta} = -0.038$, $p = 0.58$). Thus, there is no significant relationship between the time trends of the two periods for the App only group. However, the three way interaction (`Cond` by act1 by `Mon_ws` follow-up phase) is significant ($\hat{\beta} = 0.236$, $p = 0.02$). This indicates that this association of the time trends is significantly more positive for the App + Coaching group, relative to the App only group. Adding these two together (as a linear combination of parameters), we get an estimate of 0.236 - 0.038 = 0.198, with $p = 0.01$. For the App + Coaching group, but not for the App only group, this indicates that there is a positive association between the weight loss/gain during the active phase with the weight loss/gain during the follow-up phase.

To get a visual sense of the difference in this association of time trends by condition, we generated the subject random effect (empirical Bayes) estimates of these time trends based on the bivariate shared parameter model without the interactions. Specifically, these are the random effects $\upsilon_{1i}^{(A)}$ and $\upsilon_{1i}^{(F)}$ from Equations (1.9) and (1.10), respectively. A scatterplot of these random effect estimates for the App only and the App + Coaching conditions are provided in Figure 1.2.

****

Insert Figure 1.2 about here

****

In this plot, the active period estimates are on the y-axis and the follow-up period estimates are on the x-axis. The estimates from the two conditions (App only and App + Coaching) are represented by different symbols, and regression lines are included in the plot for each condition (with different line patterns). As can be seen, there is virtually no relationship in the App only condition (correlation: $\hat{r}$ = 0.035), but a positive association in the App + Coaching condition (correlation: $\hat{r}$ = 0.236).

1.3.4 Bivariate Shared Parameter Mixed Model with Random Effect Interactions

Beyond interacting the shared parameters with observed variables, as was illustrated in the previous model with $\texttt{Cond}_i$, one can also interact the random effects themselves in the shared parameter model. Here, we will consider this by interacting the active period and follow-up period random effects, including interactions with time. Here, we will return to the model that does not include $\texttt{Cond}_i$, but instead build random effect interaction terms in the model for $\texttt{wt}_{ij}^{(F)}$. Below is this version of the bivariate shared parameter mixed model.

$$\texttt{wt}_{ij}^{(A)} = \beta_0^{(A)} + \beta_1^{(A)}\texttt{Mon\_ws}_{ij}^{(A)} + \beta_2^{(A)}\texttt{Mon\_bs}_i^{(A)} + \upsilon_{0i}^{(A)} + \upsilon_{1i}^{(A)}\texttt{Mon\_ws}_{ij}^{(A)} + \epsilon_{ij}^{(A)} , \tag{1.17}$$

$$\begin{aligned}\texttt{wt}_{ij}^{(F)} &= \beta_0^{(F)} + \beta_1^{(F)}\texttt{Mon\_ws}_{ij}^{(F)} + \beta_2^{(F)}\texttt{Mon\_bs}_i^{(F)} + \beta_3^{(F)}\upsilon_{0i}^{(A)} + \beta_4^{(F)}\upsilon_{1i}^{(A)} \\ &\quad + \beta_5^{(F)}(\upsilon_{0i}^{(A)} \times \texttt{Mon\_ws}_{ij}^{(F)}) + \beta_6^{(F)}(\upsilon_{1i}^{(A)} \times \texttt{Mon\_ws}_{ij}^{(F)}) \\ &\quad + \beta_7^{(F)}\left(\upsilon_{0i}^{(A)} \times \upsilon_{0i}^{(F)}\right) + \beta_8^{(F)}\left(\upsilon_{1i}^{(A)} \times \upsilon_{0i}^{(F)}\right) \\ &\quad + \beta_9^{(F)}\left(\upsilon_{0i}^{(A)} \times \upsilon_{1i}^{(F)} \times \texttt{Mon\_ws}_{ij}^{(F)}\right) + \beta_{10}^{(F)}\left(\upsilon_{1i}^{(A)} \times \upsilon_{1i}^{(F)} \times \texttt{Mon\_ws}_{ij}^{(F)}\right) \\ &\quad + \upsilon_{0i}^{(F)} + \upsilon_1^{(F)}\texttt{Mon\_ws}_{ij}^{(F)} + \epsilon_{ij}^{(F)} .\end{aligned} \tag{1.18}$$

Notice that the model for $\texttt{wt}_{ij}^{(F)}$ now includes the random-effect product terms with regression coefficients $\beta_7^{(F)}, \beta_8^{(F)}, \beta_9^{(F)}$, and $\beta_{10}^{(F)}$.

Again, it is instructive to consider the multilevel version of the model for $\texttt{wt}_{ij}^{(F)}$, given below. The random-effect interaction terms are included in brackets to distinguish what they add to the model.

level − 1

$$\texttt{wt}_{ij}^{(F)} = b_{0i}^{(F)} + b_{1i}^{(F)}\texttt{Mon\_ws}_{ij}^{(F)} + \epsilon_{ij}^{(F)} , \tag{1.19}$$

level − 2

$$\begin{aligned}b_{0i}^{(F)} &= \beta_0^{(F)} + \beta_2^{(F)}\texttt{Mon\_bs}_i^{(F)} + \beta_3^{(F)}\upsilon_{0i}^{(A)} + \beta_4^{(F)}\upsilon_{1i}^{(A)} , \\ &\quad + \left[\beta_7^{(F)}\left(\upsilon_{0i}^{(A)} \times \upsilon_{0i}^{(F)}\right) + \beta_8^{(F)}\left(\upsilon_{1i}^{(A)} \times \upsilon_{0i}^{(F)}\right)\right] + \upsilon_{0i}^{(F)} ,\end{aligned} \tag{1.20}$$

$$\begin{aligned}b_{1i}^{(F)} &= \beta_1^{(F)} + \beta_5^{(F)}\upsilon_{0i}^{(A)} + \beta_6^{(F)}\upsilon_{1i}^{(A)} \\ &\quad + \left[\beta_9^{(F)}\left(\upsilon_{0i}^{(A)} \times \upsilon_{1i}^{(F)}\right) + \beta_{10}^{(F)}\left(\upsilon_{1i}^{(A)} \times \upsilon_{1i}^{(F)}\right)\right] + \upsilon_{1i}^{(F)} .\end{aligned} \tag{1.21}$$

Consider first, the equation for the follow-up period grand mean $b_{0i}^{(F)}$, which is a function of the population grand mean $\beta_0^{(F)}$, the subject's average measurement time $\beta_2^{(F)}\texttt{Mon\_bs}_i^{(F)}$, their follow-up grand mean deviation $\upsilon_{0i}^{(F)}$, their active period random effects $\beta_3^{(F)} \upsilon_{0i}^{(A)} + \beta_4^{(F)} \upsilon_{1i}^{(A)}$, and the interaction of their follow-up grand mean deviation with their active period random effects

$$\left[\beta_7^{(F)} \left(\upsilon_{0i}^{(A)} \times \upsilon_{0i}^{(F)}\right) + \beta_8^{(F)} \left(\upsilon_{1i}^{(A)} \times \upsilon_{0i}^{(F)}\right)\right] .$$

If we further focus on contribution from the parameters involving the follow-up grand mean deviation $\upsilon_{0i}^{(F)}$, we have:

$$\left(1 \ + \ \beta_7^{(F)} \ \upsilon_{0i}^{(A)} \ + \ \beta_8^{(F)} \ \upsilon_{1i}^{(A)}\right) \ \upsilon_{0i}^{(F)} \ . \tag{1.22}$$

The follow-up grand mean deviations $\upsilon_{0i}^{(F)}$ represent the heterogeneity across subjects in the grand means. If these are more towards zero, there is less heterogeneity, and as these deviate from zero, there is increased heterogeneity. The interaction terms $\beta_7^{(F)}$ and $\beta_8^{(F)}$, multiplied by the active period random effects $\upsilon_{0i}^{(A)}$ and $\upsilon_{1i}^{(A)}$, respectively, thus moderate this heterogeneity (to the extent that they are non-zero). If these interaction effects are positive, heterogeneity in the grand means is increased, whereas if these interactions are negative, then heterogeneity in the grand means is decreased. Similarly, for the time trend $b_{1i}^{(F)}$ and its subject-specific deviation $\upsilon_{1i}^{(F)}$. The interaction terms $\beta_9^{(F)}$ and $\beta_{10}^{(F)}$ indicate the degree to which heterogeneity in the follow-up period time trends is affected by the active period random effects.

The results for this model are presented in Table 1.5.

****

Insert Table 1.5 about here

****

Comparing this model to a model without the four interaction terms yields a likelihood-ratio test statistic of $\chi_4^2 = 53$, $p = .001$. Thus, there is clear evidence of interaction between the active period and follow-up period random effects. The estimates of these four interaction terms are listed under "Random effect interactions" in Table 1.5; all four are observed to be statistically significant. At first glance, one might think that these effects are not large given that they are all close to zero. However, one must consider the scaling of the random effects. For example, if we standardize the random effects (to have zero mean and unit standard deviation), in terms of the effects involving the follow-up grand mean deviation $\upsilon_{0i}^{(F)}$, we have (using θ to represent the standardized random effects):

$$\left(1 \ + \ \hat{\beta}_7^{(F)} \ \hat{\sigma}_{\upsilon_0^{(A)}} \ \theta_{0i}^{(A)} \ + \ \beta_8^{(F)} \ \hat{\sigma}_{\upsilon_1^{(A)}} \ \theta_{1i}^{(A)}\right) \ \hat{\sigma}_{\upsilon_0^{(F)}} \ \theta_{0i}^{(F)} \ . \tag{1.23}$$

For the interaction involving the grand means of the two periods, multiplying the regression estimates by the estimated random effect standard deviations yields

$0.009817 \times \sqrt{1151.58} \times \sqrt{23.0094} = 1.598$. Similar calculations yield estimates of -0.633 (active time trend by follow-up grand mean), 0.329 (active grand mean by follow-up time trend), and -0.267 (active time trend by follow-up time trend) for the other interactions. Thus, the general pattern is that higher/lower values of the active period grand mean increases/decreases heterogeneity of both follow-up random effects, and higher/lower values of the active period time trends reduces/increases heterogeneity of both follow-up random effects.

A visual representation of these interaction effects is provided in Figure 1.3.

****
Insert Figure 1.3 about here
****

The top two figures display the effects of the active period grand mean and time trend (random effects), respectively, on the estimated follow-up average weight for three different values of the standardized active period random effects (-1, 0, 1). The slope of the lines in these figures represents the heterogeneity in the follow-up average weight. As can be seen, in the top left figure, the slope is more pronounced for higher values of the active period grand mean, relative to lower values. Thus, as the active period grand mean is higher, heterogeneity in the follow-up average weight is increased. Conversely, the effect of the active period time trend on the follow-up average weight, which is displayed in the top right figure, is the opposite. Here, as the active period time trend is higher, heterogeneity in the follow-up average weight is decreased. This suggests that some subjects who lost more weight in the active period (i.e., negative active period time trends) were susceptible to relapsing and gaining the weight back.

The bottom two figures in Figure 1.3 display the effects of the active period grand mean and time trend (random effects), respectively, on the estimated follow-up linear trend for three different values of the standardized active period random effects (-1, 0, 1). Here, negative/positive linear trends indicate weight loss/gain in the follow-up period. First, focusing on the figure in lower left-hand side, one sees that higher values of the active period grand mean led to greater negative time trends in the follow-up period if the follow-up random effect was negative. Conversely, higher values of the active period grand mean led to greater positive time trends in the follow-up period if the follow-up random effect was positive. Thus, there was increased heterogeneity in the follow-up time trends (i.e., a wider range of time trends) as the active period grand mean increased. The figure in the lower right-side displays the effect of the active period time trend on the follow-up period time trend. Here, lower values (i.e., more negative) of the active period time trend are associated with greater heterogeneity in the follow-up period time trends. Again, this suggests that some subjects who lost more weight in the active period (i.e., negative active period time trends) were susceptible to relapsing and gaining the weight back. Another interesting finding is the convergence of the three lines for positive follow-up time trends (say, trends of 0.5 to 1.0). Since trends in this range indicate weight gain in the follow-up period, this convergence indicates that follow-up weight gain was essentially independent of the value of the active period time trend.

1.4 DISCUSSION

This chapter has illustrated how multivariate mixed models can be used to model the longitudinal data from several outcomes. This builds on the work of Thiébaut et al. (2002), who described a practical way of using mixed model software for bivariate outcomes. Their paper focused on mixed models for two outcomes measured concurrently. Here, we have considered two outcomes measured at different time periods, and have shown how elements of the random effect variance covariance matrix can be shifted to be shared parameters in the mixed models. The advantage of this is that these random effect associations can then interact with other variables, either observed variables or random effects themselves. In this way, one can examine whether such variables moderate the associations of the outcome variables, in terms of their random effects.

Appendix A indicates how to structure the dataset for both the bivariate mixed model and the bivariate shared parameter mixed model, and provides sample syntax for all of the models considered in this chapter. In particular, the software program SAS PROC NLMIXED was used for the analyses presented in this chapter, and sample NLMIXED scripts are provided in Appendix A for all of the models. NLMIXED is useful because it allows for programming statements, which also makes it somewhat more complicated to use than standard mixed model software. Also provided in Appendix B is sample syntax using STAN for Bayesian analysis of the bivariate mixed model. Hopefully, the scripts provided in the appendices will help researchers use the models presented in this chapter for their own research.

In this chapter, we have focused on continuous normally-distributed outcomes. However, the same general approach can be applied to other outcome types and even mixtures of outcome types (i.e., a binary outcome with a continuous normal outcome). Software like SAS PROC NLMIXED, with its programming capabilities, makes this possible, and Blozis et al. (2020) provide NLMIXED code for a model that has both a continuous log-normal component and a logistic component. Alternatively, Bayesian software like JAGS, STAN, and WinBUGS can be used as well, and offer great flexibility, but they need additional specifications for prior distributions of the parameters. As an example, Siddique et al. (2023) describe a bivariate model of the frequency and duration of accelerometer-measured physical activity data, including JAGS code, using a mixed-effects Poisson hurdle sub-model for the number of bouts per day and a mixed-effects location scale gamma regression sub-model to characterize the duration of the bouts and their variance.

BIBLIOGRAPHY

Blozis, S. A., McTernan, M., Harring, J. R. & Zheng, Q. (2020), ‘Two-part mixed-effects location scale models’, *Behavioral Research Methods* **52(5)**, 1836–1847.

Bolger, N., Davis, A. & Rafaeli, E. (2003), ‘Diary methods: capturing life as it is lived’, *Annual Review of Psychology* **54**, 579–616.

Chi, E. M. & Reinsel, G. C. (1989), ‘Models for longitudinal data with random effects and ar(1) errors’, *Journal of the American Statistical Association* **84**(406), 452–459.

de Vries, M. (1992), *The experience of psychopathology: Investigating mental disorders in their natural settings*, Cambridge University Press, New York.

Feldman Barrett, L. & Barrett, D. (2001), ‘An introduction to computerized experience sampling in psychology’, *Social Science Computer Review* **19**, 175–185.

Goldstein, H. (2011), *Multilevel Statistical Models, 4th edition*, Wiley, Chichester, UK.

Hedeker, D. (1989), *Random Regression Models with Autocorrelated Errors*, Ph.D. thesis, University of Chicago, Department of Psychology.

Neuhaus, J. M. & Kalbfleisch, J. D. (1998), ‘Between- and within- cluster covariate effects in the analysis of clustered data’, *Biometrics* **54**, 638–645.

Pfammatter, A. F., Nahum-Shani, I., DeZelar, M., Scanlan, L., McFadden, H. G., Siddique, J., Hedeker, D. & Spring, B. (2019), ‘SMART: Study protocol for a sequential multiple assignment randomized controlled trial to optimize weight loss management’, *Contemporary Clinical Trials* **82**, 36–45.

Scollon, C. N., Kim-Prieto, C. & Diener, E. (2003), ‘Experience sampling: promises and pitfalls, strengths and weeknesses’, *Journal of Happiness Studies* **4**, 5–34.

Shiffman, S., Stone, A. A. & Hufford, M. R. (2008), ‘Ecological momentary assessment’, *Annual Review of Clinical Psychology* **4**, 1–32.

Siddique, J., Daniels, M. J., Inan, G., Battalio, S., Spring, B. & Hedeker, D. (2023), ‘Joint modeling the frequency and duration of accelerometer-measured physical activity from a lifestyle intervention trial’, *Statistics in Medicine* **42**.

Smyth, J. M. & Stone, A. A. (2003), ‘Ecological momentary assessment research in behavioral medicine’, *Journal of Happiness Studies* **4**, 35–52.

Stone, A. & Shiffman, S. (1994), ‘Ecological Momentary Assessment (EMA) in behavioral medicine’, *Annals of Behavioral Medicine* **16**, 199–202.

Thiébaut, R., Jacqmin-Gadda, H., Chêne, G., Leport, C. & Commenges, D. (2002), ‘Bivariate linear mixed models using SAS proc MIXED’, *Computer Methods and Programs in Biomedicine* **69(3)**, 249–256.

van de Pol, M. & Wright, J. (2009), ‘A simple method for distinguishing within- versus between-subject effects using mixed models’, *Animal Behaviour* **77**, 753–758.

Walls, T. A. & Schafer, J. L. (2006), *Models for intensive longitudinal data*, Oxford University Press, New York.

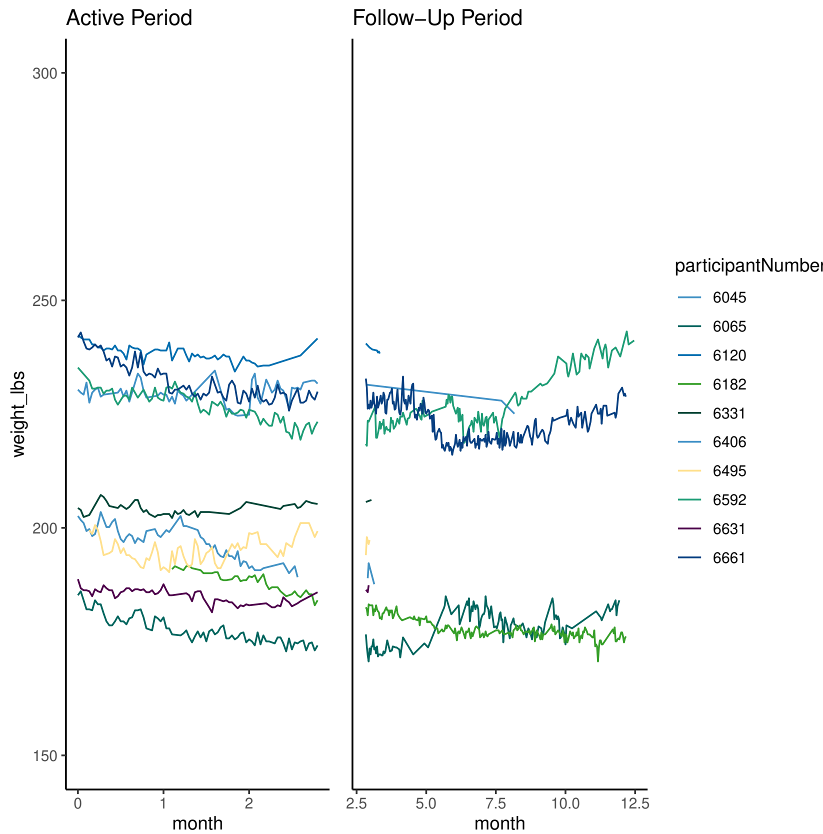

FIGURE 1.1 Spaghetti Plot of Weight - Active Period and Follow-Up Period.

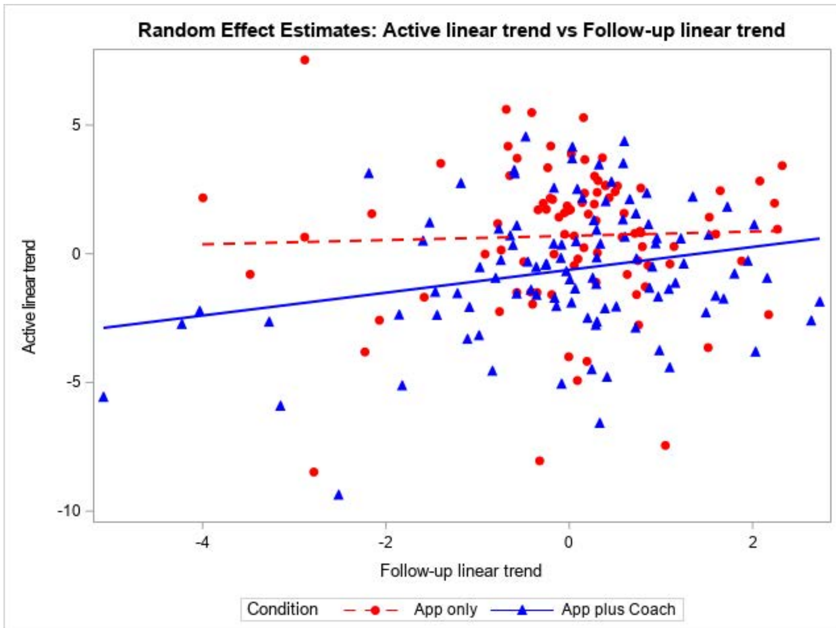

FIGURE 1.2 Plot of estimated linear trends in active and follow-up by condition.

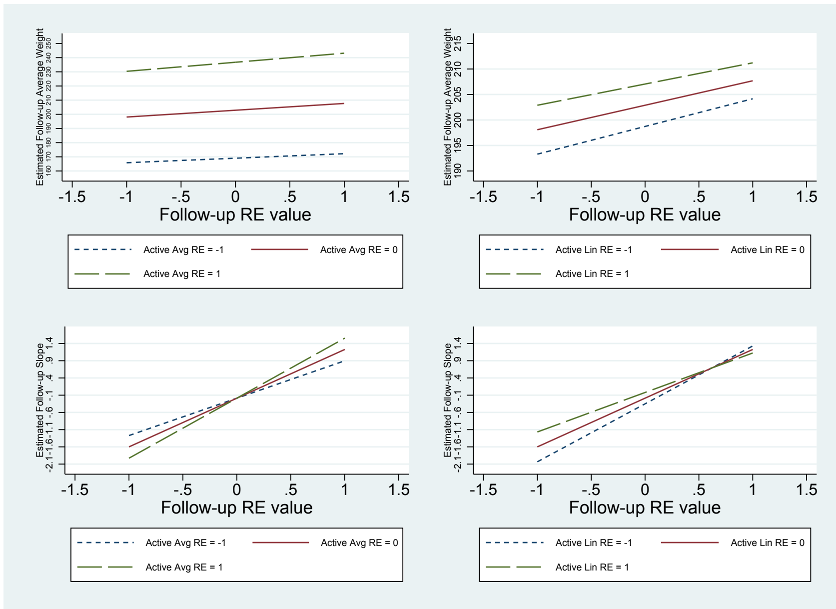

FIGURE 1.3 Estimated weights and time trends based on interactions of random effects.

TABLE 1.1 Bivariate mixed-effects model of weight, $N = 196, \sum n_i = 20,703$, maximum likelihood estimates, standard errors, z-values, and p-values.

Variable	Estimate	Std Error	z-value	p-value
Regression coefficients				
active				
`Intercept`	207.23	2.4265	85.40	0.0001
`Mon_ws`	-2.8521	0.2016	-14.15	0.0001
`Mon_bs`	3.4319	1.3491	2.54	0.0118
follow-up				
`Intercept`	202.88	2.4822	81.74	0.0001
`Mon_ws`	-0.2161	0.1310	-1.65	0.1005
`Mon_bs`	-0.1155	0.1561	-0.74	0.4603
Random effect variances				
active				
`Intercept`	1153.96	118.61	9.73	0.0001
`Mon_ws`	7.7154	0.8196	9.41	0.0001
follow-up				
`Intercept`	1206.74	124.02	9.73	0.0001
`Mon_ws`	2.1962	0.2672	8.22	0.0001
*Correlations of random effects**				
act01	0.06116	0.07261	0.84	0.4007
fol01	0.1164	0.08402	1.39	0.1675
act0_fol0	0.9822	0.00250	392.97	0.0001
act1_fol0	0.1873	0.07053	2.66	0.0086
act0_fol1	0.02056	0.08541	0.24	0.8100
act1_fol1	0.1147	0.09538	1.20	0.2308
Error variances				
active	2.9875	0.03855	77.49	0.0001
follow-up	5.8648	0.09284	63.17	0.0001

* 0 = `Intercept`, 1 = `Mon_ws`, act = active, fol = follow-up

TABLE 1.2 Bivariate shared parameter mixed-effects model of weight, $N = 196, \sum n_i = 20,703$, maximum likelihood estimates, standard errors, z-values, and p-values.

Variable	Estimate	Std Error	z-value	p-value
Regression coefficients				
active				
`Intercept`	207.23	2.4265	85.40	0.0001
`Mon_ws`	-2.8521	0.2016	-14.15	0.0001
`Mon_bs`	3.4319	1.3491	2.54	0.0118
follow-up				
`Intercept`	202.88	2.4822	81.74	0.0001
`Mon_ws`	-0.2161	0.1310	-1.65	0.1005
`Mon_bs`	-0.1155	0.1561	-0.74	0.4603
shared parameter effects in follow-up model*				
act0	0.9964	0.01050	94.88	0.0001
act1	1.5976	0.1329	12.02	0.0001
act0 by `Mon_ws`	0.000593	0.003739	0.16	0.8742
act1 by `Mon_ws`	0.06074	0.05162	1.18	0.2408
Random effect variances				
active				
`Intercept`	1153.70	116.57	9.90	0.0001
`Mon_ws`	7.7154	0.8196	9.41	0.0001
follow-up				
`Intercept`	22.9483	2.4264	9.46	0.0001
`Mon_ws`	2.1669	0.2652	8.17	0.0001
*Correlations of random effects**				
act01	0.06113	0.07263	0.84	0.4010
fol01	0.5966	0.05766	10.35	0.0001
Error variances				
active	2.9875	0.03855	77.49	0.0001
follow-up	5.8648	0.09284	63.17	0.0001

* 0 = `Intercept`, 1 = `Mon_ws`, act = active, fol = follow-up

TABLE 1.3 Bivariate shared parameter mixed-effects model of weight, $N = 196, \sum n_i = 20,703$, Bayesian estimates and 95% Credible Intervals via Stan.

Variable	Estimate	95% Credible Interval
Regression coefficients		
active		
`Intercept`	203.79	(197.65 to 209.80)
`Mon_ws`	-2.8457	(-3.2605 to -2.3888)
`Mon_bs`	3.4130	(0.5985 to 6.1209)
follow-up		
`Intercept`	204.71	(199.50 to 209.57)
`Mon_ws`	-0.2230	(-0.4803 to 0.0290)
`Mon_bs`	-0.1382	(-0.4485 to 0.1673)
shared parameter effects in follow-up model*		
act0	0.9960	(0.9746 to 1.0171)
act1	1.5942	(1.3290 to 1.8514)
act0 by `Mon_ws`	0.0005127	(-0.007364 to 0.007821)
act1 by `Mon_ws`	0.06130	(-0.03594 to 0.16005)
Random effect variances		
active		
`Intercept`	1162.03	(952.96 to 1411.69)
`Mon_ws`	7.7660	(6.2709 to 9.6012)
follow-up		
`Intercept`	23.5487	(19.0222 to 29.1019)
`Mon_ws`	2.2228	(1.7404 to 2.8091)
*Correlations of random effects**		
act01	0.06072	(-0.07919 to 0.2000)
fol01	0.5916	(0.4697 to 0.6957)
Error variances		
active	2.9884	(2.9113 to 3.0653)
follow-up	5.8672	(5.6862 to 6.0518)

* 0 = `Intercept`, 1 = `Mon_ws`, act = active, fol = follow-up

TABLE 1.4 Bivariate shared parameter mixed-effects model of weight with condition interactions, $N = 196, \sum n_i = 20,703$, maximum likelihood estimates, standard errors, z-values, and p-values.

Variable	Estimate	Std Error	z-value	p-value
Regression coefficients				
active				
Intercept	207.70	3.5216	58.98	0.0001
Mon_ws	-2.1271	0.2876	-7.40	0.0001
Mon_bs	3.1991	1.6732	1.91	0.0574
Cond	-0.8806	4.8580	-0.18	0.8563
Cond by Mon_ws	-1.3800	0.3948	-3.50	0.0006
Cond by Mon_bs	-0.4806	2.8702	-0.17	0.8672
follow-up				
Intercept	204.61	3.6266	56.42	0.0001
Mon_ws	-0.2183	0.1876	-1.16	0.2461
Mon_bs	-0.03737	0.2204	-0.17	0.8656
Cond	-3.2964	4.9674	-0.66	0.5077
Cond by Mon_ws	0.02693	0.2606	0.10	0.9178
Cond by Mon_bs	-0.1750	0.3145	-0.56	0.5786
shared parameter effects in follow-up model*				
act0	1.0043	0.01476	68.03	0.0001
act1	1.5691	0.1874	8.37	0.0001
act0 by Mon_ws	0.003533	0.005079	0.70	0.4875
act1 by Mon_ws	-0.03790	0.06795	-0.56	0.5776
Cond by act0	-0.01624	0.02097	-0.77	0.4394
Cond by act1	0.04407	0.2712	0.16	0.8711
Cond by act0 by Mon_ws	-0.00528	0.007301	-0.72	0.4701
Cond by act1 by Mon_ws	0.2358	0.09992	2.36	0.0193
Random effect variances				
active				
Intercept	1152.97	116.49	9.90	0.0001
Mon_ws	7.3632	0.7817	9.42	0.0001
follow-up				
Intercept	22.7933	2.4124	9.45	0.0001
Mon_ws	2.0566	0.2530	8.13	0.0001
*Correlations of random effects**				
act01	0.06340	0.07267	0.87	0.3840
fol01	0.6067	0.05643	10.75	0.0001
Error variances				
active	2.9872	0.03855	77.49	0.0001
follow-up	5.8645	0.09283	63.17	0.0001

* 0 = Intercept, 1 = Mon_ws, act = active, fol = follow-up

TABLE 1.5 Bivariate shared parameter mixed-effects model of weight with random effect interactions, $N = 196, \sum n_i = 20,703$, maximum likelihood estimates, standard errors, z-values, and p-values.

Variable	Estimate	Std Error	z-value	p-value
Regression coefficients				
active				
`Intercept`	207.23	2.4241	85.49	0.0001
`Mon_ws`	-2.8939	0.1985	-14.58	0.0001
`Mon_bs`	1.8822	1.1639	1.62	0.1075
follow-up				
`Intercept`	202.89	2.4728	82.05	<.0001
`Mon_ws`	-0.1865	0.1292	-1.44	0.1505
`Mon_bs`	-0.01468	0.1339	-0.11	0.9128
shared parameter effects in follow-up model*				
act0	0.9981	0.01016	98.26	0.0001
act1	1.5225	0.1410	10.80	0.0001
act0 by `Mon_ws`	1.507E-6	0.003589	0.00	0.9997
act1 by `Mon_ws`	0.06042	0.05066	1.19	0.2345
*Random effect interactions**				
act0 by fol0	0.009817	0.001544	6.36	0.0001
act1 by fol0	-0.04830	0.01903	-2.54	0.0119
act0 by fol1	0.006851	0.002112	3.24	0.0014
act1 by fol1	-0.06931	0.02283	-3.04	0.0027
Random effect variances				
active				
`Intercept`	1151.58	116.35	9.90	0.0001
`Mon_ws`	7.4626	0.7791	9.58	0.0001
follow-up				
`Intercept`	23.0094	2.8518	8.07	0.0001
`Mon_ws`	1.9982	0.2761	7.24	0.0001
*Correlations of random effects**				
act01	0.02721	0.07335	0.37	0.7111
fol01	0.5996	0.05444	11.01	<.0001
Error variances				
active	2.9860	0.03852	77.51	0.0001
follow-up	5.8624	0.09281	63.17	0.0001

* 0 = `Intercept`, 1 = `Mon_ws`, act = active, fol = follow-up

Appendix A

SAS PROC NLMIXED code

First, let us get a sense of the dataset structure that is needed for the bivariate mixed model. Suppose that subject 10 is measured three times during the active period and twice in the follow-up period, and that his/her weights (and time values) were 200 (0), 195 (1), and 190 (2) during the active phase; and 189 (0) and 191 (3) during the follow-up phase. Using `act` to denote the active period, and `fol` to denote the follow-up period, this subject would then have the between-subject average values of `act_Mon_bs` = 1 and `fol_Mon_bs` = 1.5. The `act_Mon_ws` and `fol_Mon_ws` values would then be relative to these means, namely -1,0,1 for the active period and -1.5,1.5 for the follow-up period. Then, this subject would contribute five records in the dataset with variables and values as the following.

id	wt_lbs	act_int	act_Mon_ws	act_Mon_bs	fol_int	fol_Mon_ws	fol_Mon_bs
10	200	1	-1	1	0	0	0
10	195	1	0	1	0	0	0
10	190	1	1	1	0	0	0
10	189	0	0	0	1	-1.5	1.5
10	191	0	0	0	1	1.5	1.5

Here, `id` is a subject identifier, `wt_lbs` denotes the outcome (with the active period weights stacked on top of the follow-up period weights); `act_int`, `act_Mon_ws`, and `act_Mon_bs` are the active period regressors; and `fol_int`, `fol_Mon_ws`, and `fol_Mon_bs` are the follow-up period regressors. Note that the follow-up variables have values of 0 for the three active period rows, and the active period variables have 0 values for the two follow-up period rows.

Once the dataset is organized in this way, sample syntax necessary to run the bivariate mixed-effects model described in this chapter is given below. In this syntax, upper case is used for SAS specific syntax and lower or mixed case is used for user defined entities.

```
/* Bivariate mixed model via PROC MIXED */
PROC MIXED COVTEST METHOD=ML;
CLASS id;
MODEL  wt_lbs = act_int act_Mon_ws act_Mon_bs
                    fol_int fol_Mon_ws fol_Mon_bs / S NOINT;
RANDOM act_int act_Mon_ws act fol_int fol_Mon_ws
       / SUBJECT = id TYPE=UN G GCORR;
REPEATED / LOCAL=EXP(fol_int);
```

```
RUN;
```

Here it is important to specify the NOINT option on the MODEL statement, since we are estimating separate intercepts for the two periods by including act_int and fol_int as regressors. Also, the specification on the REPEATED statement will yield the two error variance estimates, one for the active period and the second for the difference between the follow-up and active periods (on the log scale). Other options are COVTEST to provide standard error estimates and tests of the (co)variance parameters, METHOD=ML to select maximum likelihood estimation, S (on the MODEL statement) to print out the estimated regression coefficients, TYPE=UN (on the RANDOM statement) to allow the random effects to be correlated, G (on the RANDOM statement) to print out the estimated variance-covariance matrix of the random effects, and GCORR (on the RANDOM statement) to print out the estimated variance-covariance matrix of the random effects as a correlation matrix.

One can also use SAS PROC NLMIXED to estimate the parameters of the bivariate mixed model. Strictly speaking, this is not necessary since PROC MIXED can do it, however, NLMIXED is necessary for the shared-parameter models included in this chapter. So, it is helpful to use NLMIXED for the bivariate mixed model, and then to modify the code for the shared-parameter models. Below is NLMIXED code for the bivariate mixed model. NLMIXED has many estimation options, which sometimes need to be modified to achieve convergence. Here, we have selected the options `TECH=TRUREG OPTCHECK HESCAL=1`. Also, all model parameters need to be named and given starting values on the `PARMS` statement. Here, we use the estimates from the previous PROC MIXED run as the starting values.

```
/* Bivariate mixed model via PROC NLMIXED */
/* these results are listed in Table 1    */
PROC NLMIXED TECH=TRUREG OPTCHECK HESCAL=1 ;
PARMS b0_act=202.64 b_act_Mon_ws=-2.85 b_act_Mon_bs=3.43
      b0_fol=203.15 b_fol_Mon_ws=-0.22 b_fol_Mon_bs=-0.12
      v0_act=1159 c01_act=-5.02 v1_act=8.23 act0_fol0=1148.3 act1_fol0=16.19
      v0_fol=1232.2 act0_fol1=-3.3 act1_fol1=1.08 c01_fol=-8.66 v1_fol=2.51
      vare_act=3 vare_fol=3.5;
/* initialize the likelihood functions to zero */
ll_act=0; ll_fol=0;   pi = ARCOS(-1);
IF (fol_int EQ 0) THEN /* active period model */
DO;
   mu_act = (b0_act + u0_act) + (b_act_Mon_ws + u1_act)*act_Mon_ws
          + b_act_Mon_bs*act_Mon_bs;
   ll_act = LOG(1 / (SQRT(2*pi*vare_act)))
          + (-(wt_lbs-mu_act)**2)/ (2*vare_act);
END;
IF (fol_int EQ 1) THEN  /* follow-up period model */
DO;
   mu_fol = (b0_fol + u0_fol) + (b_fol_Mon_ws + u1_fol)*fol_Mon_ws
          + b_fol_Mon_bs*fol_Mon_bs;
   ll_fol = log(1 / (SQRT(2*pi*vare_fol)))
          + (-(wt_lbs-mu_fol)**2)/ (2*vare_fol);
END;
ll = ll_act+ll_fol;
MODEL wt_lbs ~ GENERAL(ll);
RANDOM u0_act u1_act u0_fol u1_fol ~ normal([0,0,0,0],
        [v0_act,c01_act,v1_act,act0_fol0,act1_fol0,
         v0_fol,act0_fol1,act1_fol1,c01_fol,v1_fol]) SUBJECT=id;

/* express covariances as correlations  */
ESTIMATE 'corr_act01' c01_act / (SQRT(v0_act) * SQRT(v1_act));
ESTIMATE 'corr_fol01' c01_fol / (SQRT(v0_fol) * SQRT(v1_fol));
ESTIMATE 'corr_act0_fol0' act0_fol0 / (SQRT(v0_act) * SQRT(v0_fol));
ESTIMATE 'corr_act1_fol0' act1_fol0 / (SQRT(v1_act) * SQRT(v0_fol));
ESTIMATE 'corr_act0_fol1' act0_fol1 / (SQRT(v0_act) * SQRT(v1_fol));
ESTIMATE 'corr_act1_fol1' act1_fol1 / (SQRT(v1_act) * SQRT(v1_fol));
RUN;
```

Below is NLMIXED syntax for the shared parameter bivariate mixed model, corresponding to Equations (1.9) and (1.10) and with results in Table 1.2. For the starting values, we used estimates from the bivariate mixed model.

```
/* Shared Parameter bivariate mixed model via PROC NLMIXED */
/* these results are listed in Table 2    */
PROC NLMIXED TRUREG OPTCHECK HESCAL=1;
PARMS b0_act=202.64 b_act_Mon_ws=-2.85 b_act_Mon_bs=3.43
      b0_fol=203.15 b_fol_Mon_ws=-0.22 b_fol_Mon_bs=-0.12
      b_act0_fol0=0 b_act1_fol0=0 b_act0_fol1=0 b_act1_fol1=0
      v0_act=1159 c01_act=-5.02 v1_act=8.23
      v0_fol=1232.2 c01_fol=-8.66 v1_fol=2.51
      vare_act=3 vare_fol=3.5;
/* initialize the likelihood functions to zero */
ll_act=0; ll_fol=0;   pi = arcos(-1);
IF (fol_int EQ 0) THEN /* active period model */
DO;
   mu_act = (b0_act + u0_act) + (b_act_Mon_ws + u1_act)*act_Mon_ws
          + b_act_Mon_bs*act_Mon_bs;
   ll_act = LOG(1 / (SQRT(2*pi*vare_act)))
          + (-(wt_lbs-mu_act)**2)/ (2*vare_act);
END;
IF (fol_int EQ 1) THEN  /* follow-up period model */
DO;
   mu_fol = (b0_fol + u0_fol + b_act0_fol0*u0_act  + b_act1_fol0*u1_act)
          + (b_fol_Mon_ws + u1_fol + b_act0_fol1*u0_act
          + b_act1_fol1*u1_act)*fol_Mon_ws
          + b_fol_Mon_bs*fol_Mon_bs;
   ll_fol = LOG(1 / (SQRT(2*pi*vare_fol)))
          + (-(wt_lbs-mu_fol)**2)/ (2*vare_fol);
END;
ll = ll_act+ll_fol;
MODEL wt_lbs ~ GENERAL(ll);
RANDOM u0_act u1_act u0_fol u1_fol ~ NORMAL([0,0,0,0],
        [v0_act,c01_act,v1_act,0,0,
         v0_fol,0,0,c01_fol,v1_fol]) SUBJECT=id;

/* express covariances as correlations  */
ESTIMATE 'corr_act01' c01_act / (SQRT(v0_act) * SQRT(v1_act));
ESTIMATE 'corr_fol01' c01_fol / (SQRT(v0_fol) * SQRT(v1_fol));
RUN;
```

Below is NLMIXED syntax for the shared parameter bivariate mixed model that includes interactions with the Cond variable. This corresponds to Equations (1.15) and (1.16) and with results in Table 1.4. For the starting values, we used estimates from the previous shared parameter bivariate mixed model, and set all of the interaction effects with starting values of zero.

```
/* Shared parameter bivariate mixed model via PROC NLMIXED */
/* with Cond and Cond interactions  */
/* these results are listed in Table 3    */
PROC NLMIXED TRUREG OPTCHECK HESCAL=1;
PARMS b0_act=207.23 b_act_Mon_ws=-2.85 b_act_Mon_bs=3.43
      b_act_cond=0 b_act_cond_Mon_ws=0 b_act_cond_Mon_bs=0
      b0_fol=202.88 b_fol_Mon_ws=-0.22 b_fol_Mon_bs=-0.12
      b_fol_cond=0 b_fol_cond_Mon_ws=0 b_fol_cond_Mon_bs=0
      b_act0_fol0=0.99 b_act1_fol0=1.60 b_act0_fol1=0 b_act1_fol1=0.061
      b_cond_act0_fol0=0 b_cond_act1_fol0=0 b_cond_act0_fol1=0 b_cond_act1_fol1=0
      v0_act=1154 c01_act=-5.77 v1_act=7.72 v0_fol=22.95  c01_fol=4.21 v1_fol=2.17
      vare_act=3 vare_fol=5.86;
ll_act=0; ll_fol=0;   pi = arcos(-1);
IF (fol_int EQ 0) THEN /* active period model */
DO;
   mu_act = (b0_act + u0_act) + (b_act_Mon_ws + u1_act)*act_Mon_ws
          + b_act_Mon_bs*act_Mon_bs
          + b_act_cond*cond + b_act_cond_Mon_ws*cond*act_Mon_ws
          + b_act_cond_Mon_bs*cond*act_Mon_bs;
   ll_act = LOG(1 / (SQRT(2*pi*vare_act)))
          + (-(wt_lbs-mu_act)**2)/ (2*vare_act);
END;
IF (fol_int EQ 1) THEN   /* follow-up period model */
DO;
   mu_fol = (b0_fol + u0_fol + b_act0_fol0*u0_act  + b_act1_fol0*u1_act)
          + (b_fol_Mon_ws + u1_fol + b_act0_fol1*u0_act
          + b_act1_fol1*u1_act)*fol_Mon_ws
          + b_fol_Mon_bs*fol_Mon_bs
          + b_fol_cond*cond + b_fol_cond_Mon_ws*cond*fol_Mon_ws
          + b_fol_cond_Mon_bs*cond*fol_Mon_bs
          +  b_cond_act0_fol0*u0_act*cond  + b_cond_act1_fol0*u1_act*cond
          + (b_cond_act0_fol1*u0_act*cond
          + b_cond_act1_fol1*u1_act*cond)*fol_Mon_ws;
   ll_fol = LOG(1 / (SQRT(2*pi*vare_fol)))
          + (-(wt_lbs-mu_fol)**2)/ (2*vare_fol);
END;
ll = ll_act+ll_fol;
MODEL wt_lbs ~ GENERAL(ll);
RANDOM u0_act u1_act u0_fol u1_fol ~ NORMAL([0,0,0,0],
        [v0_act,c01_act,v1_act,0,0,
         v0_fol,0,0,c01_fol,v1_fol]) SUBJECT=id;

/* express covariances as correlations  */
ESTIMATE 'corr_act01' c01_act / (SQRT(v0_act) * SQRT(v1_act));
ESTIMATE 'corr_fol01' c01_fol / (SQRT(v0_fol) * SQRT(v1_fol));
RUN;
```

Below is NLMIXED syntax for the shared parameter bivariate mixed model that includes interactions with the random effects. This corresponds to Equations (1.17) and (1.18) and with results in Table 1.5. For the starting values, we used estimates from the shared parameter bivariate mixed model (without interactions), and set all of the interaction effects with starting values of zero.

```
/* Shared parameter bivariate mixed model via PROC NLMIXED */
/* with random effect interactions */
/* these results are listed in Table 4    */
PROC NLMIXED TRUREG OPTCHECK HESCAL=1;
PARMS b0_act=207.23 b_act_Mon_ws=-2.85 b_act_Mon_bs=3.43
      b0_fol=202.88 b_fol_Mon_ws=-0.22 b_fol_Mon_bs=-0.12
      b_act0_fol0=1 b_act1_fol0=1.6 b_act0_fol1=0 b_act1_fol1=0.06
      b_INT_act0_fol0=0 b_INT_act1_fol0=0 b_INT_act0_fol1=0 b_INT_act1_fol1=0
      v0_act=1154 c01_act=3.10 v1_act=7.72 v0_fol=22.95  c01_fol=3.95 v1_fol=2.17
      vare_act=3 vare_fol=5.86;
ll_act=0; ll_fol=0;   pi = arcos(-1);
IF (fol_int EQ 0) THEN /* active period model */
DO;
   mu_act = (b0_act + u0_act) + (b_act_Mon_ws + u1_act)*act_Mon_ws
          + b_act_Mon_bs*act_Mon_bs;
   ll_act = LOG(1 / (SQRT(2*pi*vare_act)))
          + (-(wt_lbs-mu_act)**2)/ (2*vare_act);
END;
IF (fol_int EQ 1) THEN   /* follow-up period model */
DO;
   mu_fol = (b0_fol + u0_fol + b_act0_fol0*u0_act  + b_act1_fol0*u1_act
          + b_INT_act0_fol0*u0_act*u0_fol + b_INT_act1_fol0*u1_act*u0_fol)
          + (b_fol_Mon_ws + u1_fol + b_act0_fol1*u0_act  + b_act1_fol1*u1_act
          + b_INT_act0_fol1*u0_act*u1_fol
          + b_INT_act1_fol1*u1_act*u1_fol)*fol_Mon_ws
          + b_fol_Mon_bs*fol_Mon_bs;
   ll_fol = LOG(1 / (SQRT(2*pi*vare_fol)))
          + (-(wt_lbs-mu_fol)**2)/ (2*vare_fol);
END;
ll = ll_act+ll_fol;
MODEL wt_lbs ~ GENERAL(ll);
RANDOM u0_act u1_act u0_fol u1_fol ~ NORMAL([0,0,0,0],
        [v0_act,c01_act,v1_act,0,0,
         v0_fol,0,0,c01_fol,v1_fol]) SUBJECT=id;

/* express covariances as correlations  */
ESTIMATE 'corr_act01' c01_act / (SQRT(v0_act) * SQRT(v1_act));
ESTIMATE 'corr_fol01' c01_fol / (SQRT(v0_fol) * SQRT(v1_fol));
RUN;
```

Appendix B

STAN code

```
//Shared parameter bivariate mixed model via Stan
//Use the same starting values as the algorithm's maximum-likelihood counterpart using PROC NLMIXED

//The input data is a vector 'y' of length 'N'.
data {
  int<lower=1> N; //number of data points
  int<lower=1> nsubj; //number of subjects
  int<lower=1, upper=nsubj> subject[N]; //subject numbers
  real act_Mon_ws[N];
  real act_Mon_bs[N];
  real fol_int[N];
  real fol_Mon_ws[N];
  real fol_Mon_bs[N];
  real y[N]; //outcome
}


parameters {
    //fixed parameters for active period
    real b0_act;
    real b_act_Mon_ws;
    real b_act_Mon_bs;
    //fixed parameters for follow-up period
    real b0_fol;
    real b_fol_Mon_ws;
    real b_fol_Mon_bs;
    //active on follow-up
    real b_act0_fol0;
    real b_act1_fol0;
    real b_act0_fol1;
    real b_act1_fol1;
    //random effects
    //active_period
    vector[2] u_act[nsubj];
    //follow-up period
    vector[2] u_fol[nsubj];
    //random effects variances & covariances
    cov_matrix[2] cov_act;
    cov_matrix[2] cov_fol;
    //error variance
    real<lower = 0> vare_act;
    real<lower = 0> vare_fol;
```

```
}

model {
  //priors
  cov_act ~ inv_wishart(3, diag_matrix(rep_vector(1, 2)));
  cov_fol ~ inv_wishart(3, diag_matrix(rep_vector(1, 2)));
  //random effects
  for (j in 1:nsubj) {
    u_act[j] ~ multi_normal(rep_vector(0, 2), cov_act);
    u_fol[j] ~ multi_normal(rep_vector(0, 2), cov_fol);
  }
  //likelihood
  for (i in 1 : N) {
    if(fol_int[i] == 0){
      y[i] ~ normal(b0_act + u_act[subject[i]][1] + (b_act_Mon_ws + u_act[subject[i]][2]) * act_Mon_ws[i] +
      b_act_Mon_bs * act_Mon_bs[i],
      sqrt(vare_act));
    }else{
       y[i] ~ normal(b0_fol + u_fol[subject[i]][1] + b_act0_fol0 * u_act[subject[i]][1] +
       b_act1_fol0 * u_act[subject[i]][2] +
       (b_fol_Mon_ws + u_fol[subject[i]][2] + b_act0_fol1 * u_act[subject[i]][1] +
       b_act1_fol1 * u_act[subject[i]][2]) * fol_Mon_ws[i] +
       b_fol_Mon_bs * fol_Mon_bs[i],
      sqrt(vare_fol));
    }
  }
}

generated quantities{
  //correlations between random effects
  real corr_act;
  real corr_fol;

  corr_act = cov_act[1,2]/sqrt(cov_act[1,1])/sqrt(cov_act[2,2]);
  corr_fol = cov_fol[1,2]/sqrt(cov_fol[1,1])/sqrt(cov_fol[2,2]);
}
```