X. Panel Modeling: Random Coefficients and Covariance Structures

J.J. Hox
Utrecht University

x.1 Multilevel Models for Change Over Time

Multilevel or mixed models are becoming standard modeling tools for longitudinal or repeated measures data. Compared to the classic Manova approach, they have several advantages. Firstly, they deal efficiently with panel dropout; because there is no assumption that each subject must be measured on the same number of occasions, subjects with incomplete data are simply retained in the data set. The assumption is that incomplete data are missing at random (Little, 1995), which is weaker than the assumption of missing completely at random, which is made by applying listwise deletion in Manova. Secondly, it is possible to include time-varying covariates in the model. Thirdly, using polynomial functions or piecewise regression the change over time can be modeled very flexibly. Finally, by allowing regression coefficients for the change model to vary across subjects, different subjects can have their own trajectory of change, which can in turn be modeled by time invariant subject characteristics.

It is useful to distinguish between repeated measures that are collected at fixed or at varying occasions. If the measurements are taken at fixed occasions, all individuals provide measurements for the same set of occasions, usually regularly spaced, such as once every year. When occasions are varying, a different number of measures is collected at different points in time for different individuals. Such data occur, for instance, in growth studies, where physical or psychological characteristics are studied for a set of individuals at different moments in their development. The data collection could be at fixed moments in the year, but the individuals would have different ages at that moment. For a multilevel analysis of the resulting data, the difference between fixed and varying occasions is not very important. For fixed occasion designs, especially when the occasions are regularly spaced and there are no missing data, repeated measures Manova is a viable alternative for multilevel analysis. Another possibility in such designs is latent curve analysis, also known as latent growth curve analysis. This is a structural equation model (cf. Willett & Sayer, 1994) that models a repeated measures polynomial analysis of variance. This chapter focuses on fixed occasion data. In addition to the familiar multilevel model equations, it uses path diagrams to clarify the models, but the analysis concentrates on the multilevel regression approach. In fact, multilevel models and structural equation models for change over time are just different representations of the same underlying model. Bollen and Curran (2006) provide a thorough discussion of longitudinal models from a structural equation perspective, and Hedeker and Gibbons (2006) provide a comparable discussion from the multilevel regression perspective. Duncan, Duncan & Strycker (2006) provide an introduction to longitudinal modeling from both perspectives.

The multilevel regression model for longitudinal data is a straightforward application of the standard multilevel regression mode, with measurement occasions within subjects replacing the subjects within groups structure. At the lowest, the repeated measures level, we have:

$$Y_{ti} = \pi_{0i} + \pi_{1i}T_{ti} + \pi_{2i}X_{ti} + e_{ti}\ . \qquad \text{(x.1)}$$

In repeated measures applications, the coefficients at the lowest level are often indicated by the Greek letter π. This has the advantage that the subject level coefficients, which are in repeated measures modeling at the second level, can be represented by the usual Greek letter β, and so on. In equation (x.1), Y_{ti} is the response variable of individual i measured at time point t, T is the time variable that indicates the time point, and X_{ti} is a time varying covariate. Subject characteristics, such as gender, are time invariant covariates, which enter the equation at the second level:

$$\pi_{0i} = \beta_{00} + \beta_{01}Z_i + u_{0i} \text{ ,(x.2)}$$
$$\pi_{1i} = \beta_{10} + \beta_{11}Z_i + u_{1i} \text{ ,(x.3)}$$
$$\pi_{2i} = \beta_{20} + \beta_{21}Z_i + u_{2i} \text{ .(x.4)}$$

By substitution, we get the single equation model:

$$Y_{ti} = \beta_{00} + \beta_{10}T_{ti} + \beta_{20}X_{ti} + \beta_{01}Z_i + \beta_{11}Z_iT_{ti} + \beta_{21}Z_iX_{ti} + u_{1i}T_{ti} + u_{2i}X_{ti} + u_{0i} + e_{ti}\,. \qquad \text{(x.5)}$$

In multilevel models for subjects within groups, there is an assumed dependency between the subjects who are in the same group. Most often, there is no need to assume a specific structure for this dependency. Subjects within the same group are assumed exchangeable, and the intraclass correlation refers to the average correlation between two randomly chosen subjects from the same group. In multilevel models for occasions within subjects, measurement occasions are not freely exchangeable, because they are ordered in time. In such models, it often does make sense to assume a structure for the relationships between measurements across time. For example, an intuitively attractive assumption is that correlations between measures taken at different measurement occasions are higher when these occasions are close to each other in time.

Longitudinal designs often concern the analysis of structured change, such as growth or decline over time. The appropriate model for such research problems is a latent curve model, where change in the outcome variable is modeled as a function of time. As outlined above, the use of polynomials and varying regression coefficients makes this a very flexible analysis tool. As will be explained in more detail below, allowing coefficients for time to vary across subjects implies specific dependency structures. Since the focus is on modeling individual trajectories, the possibility of specifying specific structures across time is usually not explored.

Panel designs are longitudinal designs where the emphasis is on changes that do not follow a pattern of growth or decline across time. An example is using a panel to monitor satisfaction with the government. Satisfaction with the government is not expected to increase or decrease continuously. However, it is expected to fluctuate, and we may be able to predict these fluctuations with time varying covariates that capture events that occur at different occasions. Here, time is not a relevant predictor variable. Still, it is logical to assume that there is a dependency structure over time, which cannot be ignored. In this case, exploring specific structures across time is very important.

Although modeling the dependencies over time can be done implicitly, by allowing random coefficients for the time-variable, or explicitly, by specifying a specific structure, these approaches can also be combined. The remainder of this chapter discusses latent curve modeling, explicit modeling of dependency over time, and combining these approaches. The discussion takes up the issue when specific approaches are useful.

x.2 Multilevel Models for Structured Change Over Time

For structured change over time, we will use an example data set constructed by Patrick Curran This data set, hereafter called the *Curran data*, was compiled from a large longitudinal data set. Supporting documentation and the original data files are available on Internet (Curran, 1997); the following description is excerpted from Curran (1997).

The Curran data are a sample of 405 children who were within the first two years of entry to elementary school. The data consist of four repeated measures of both the child's antisocial behavior and the child's reading recognition skills. In addition, at the first measurement occasion, measures were collected of emotional support and cognitive stimulation provided by the mother. These data are a sub-sample from the National Longitudinal Survey of Youth (NLSY), based on three key criteria. First, children must have been between the ages of 6 and 8 years at the first wave of measurement. Second, children must have complete data on all measures used at the first measurement occasion. Third, only one child was considered from each mother. All N=405 children and mothers were interviewed at measurement occasion 1; on the three following occasions the sample sizes were 374, 297 and 294. Only 221 cases were interviewed at all four occasions.
The time varying variables are Antisocial Behavior (anti1–4) and Reading Recognition (read1–4). The time invariant variables are Emotional Support to the child (homeemo) and Cognitive Stimulation (homecog), Mother's Age (momage) and Child Age (kidage) in years at Time 1, and the child’s gender (kidgen).

In this example, reading recognition is the outcome variable. A simple start model for the effect over time is to include measurement occasion as a predictor variable (coded 0,1,2,3), and allow only the intercept to vary across subjects. It turns out that the relationship between occasion and reading is nonlinear, and a quadratic term is added to the model. The results are given in the first column of Table x.1 (REML estimates)

Table x.1 shows in the fixed part the regression coefficients, and in the random part the residual variance at the lowest level and the (co)variances at the second level, with standard errors in parentheses. All parameter estimates are significant by the Wald test, the variances are also significant using the more accurate likelihood ratio test.

----- insert table x.1 about here -----

The regression coefficients in the first column of Table x.1 indicate a mean reading recognition score of 2.54 at the first measurement occasion. The linear effect indicates that the reading score goes up at each occasion, and the negative quadratic effect indicates that this effect levels off at later occasions.

The second column in Table x.1 presents the results from the model where the regression coefficient for occasion varies across subjects. The variance of this coefficient is clearly significant. A third model with a varying coefficient for the quadratic component indicated no variance for that coefficient. To model the variance of the occasion coefficient interactions of other predictor variables with occasion must be added to the model. Cognitive stimulation and the child’s age predict reading, but there are no significant interactions. To simplify the exposition, the other variables are not included in the model here. The model with varying coefficients for the linear effect of occasion is presented as a SEM diagram in Figure x.1. Note that for complete correspondence with the multilevel regression approach, the variances of the four errors must be constrained to be the same.

----- insert figure x.1 about here -----

The two models that underlie Table x.1 have different consequences for the pattern of covariances between reading measures over time. The combined model for the varying coefficient for occasion is:

$$Y_{ti} = \beta_{00} + \beta_{10}T_{ti} + u_{1i}T_{ti} + u_{0i} + e_{ti}\,. \qquad \text{(x.6)}$$

In this model, the variance at a specific measurement occasion is given by (Goldstein, 2002; Raudenbush, 2002):

$$Var\left(Y_{ti}\right) = \sigma_{u0}^2 + 2T_{ti}\sigma_{u0u1} + T_{ti}^2\sigma_{u1}^2 + \sigma_e^2\,. \qquad \text{(x.7)}$$

The covariance between two measurement occasions is given by (Goldstein, 2002; Raudenbush, 2002):

$$Cov\left(T_{ti,t'i}\right) = \sigma_{u0}^2 + \left(T_{ti} + T_{t'i}\right)\sigma_{u0u1} + T_{ti}T_{t'i}\sigma_{u1}^2\,. \qquad \text{(x.8)}$$

Together, equations x.7 and x.8 specify a very restricted pattern for the variances and covariances across time. The pattern for the fixed occasion model is even more specific. By removing terms that refer to T_{ti} we obtain:

$$Var\left(Y_{ti}\right) = \sigma_{u0}^2 + \sigma_e^2 \qquad \text{(x.9)}$$

and

$$Cov\left(T_{ti,t'i}\right) = \sigma_{u0}^2\,. \qquad \text{(x.10)}$$

The model with only a random intercept assumes that all variances are the same, and all covariances are the same. In the Manova context, this assumption is known as compound symmetry, and considered highly restrictive.

----- insert table x.2 about here -----

Table x.2 presents the observed means and variances and the means and variances implied by the Occasion Fixed and Occasion Random model. It is clear that the random coefficient model is predicting the observed variances fairly well. It should be noted that some discrepancy is to be expected, because the observed means and variances are based on the non-missing cases at each measurement occasion, and the model predictions are predictions for the entire sample, assuming missing at random for the missing data.

x.3 Multilevel Models for Unstructured Change Over Time

As noted in the introduction, there are situations where it makes no sense to assume perpetual growth or decline, while it is still interesting to model change and predictors of change. The term ‘unstructured change’ is used to indicate that there is no long-term trend to model.

The example data are simulated to reflect a diary study, in which changes are expected but no overall trend. In this hypothetical study, a sample of 60 workers who work in a stressful work environment are asked to fill in a diary for 2 weeks (only working days). The study uses a State-Trait Anxiety Inventory. Trait anxiety (TraitAnx), which is assumed to be a relatively stable individual characteristic, is measured only on the first day. State anxiety (StateAnx), which is assumed to be a transitory mood state, is measured each day. Both scales are commonly normed to T-scores which have a mean of 50 and a standard deviation of 10 in the test's norm group. In addition, the study collects daily data on perceived job demands (7-point scale) and perceived social support (7-point scale).

For such data, no large differences are expected for the average anxiety on different days. This is borne out by a repeated measures Manova which finds no differences across the 10 days. The linear trend over time tested by Manova is also not significant. Figure x.2 presents a SEM diagram for repeated measures Manova. To test equality of means in a SEM context, the model in Figure x.2 is compared to a model where the means are constrained to be equal.

----- insert figure x.2 about here -----

Note that the path diagram explicitly shows that Manova estimates the variances and covariances for all measures over time. Manova is an unstructured model, there is no specific structure assumed for this covariance matrix. To model correlated errors in multilevel regression, we use a multivariate response model with a lowest level for the repeated measures, and a full set of dummy variables indicating the different occasions. Thus, we have 10 dummy variables, one for each day. The intercept term is removed from the model, and the variance of the lowest level residuals is constrained to zero. The dummy variables are all allowed to have random slopes at the second level. The equation for a model without further explanatory variables becomes:

$$Y_{ti} = \beta_1 D_1 \ldots \beta_{10} D_{10} + u_1 D_1 \ldots u_1 D_1 . \qquad \text{X.11}$$

Having ten random slopes at level two provides us with a 106×10 covariance matrix for the ten consecutive days. The regression slopes β_1 to β_{10} are simply the estimated means at the ten occasions. Equation 5.11 defines the multilevel model that is equivalent to the Manova approach. Maas and Snijders (2002) discuss this model at length, and show how the familiar *F*-ratio's can be calculated from the multilevel software output.

The model in equation 5.11 is fully saturated; it estimates all means and all (co)variances. Both the fixed part and the random part can be simplified. We can replace the fixed part by a regression equation that includes predictors such as the state anxiety and the time varying predictors job demands and social support. This models the state anxiety in a more interesting way, as the result of the combination of trait anxiety and different pressures at work. In addition we have a more parsimonious model, since we replace ten estimated means with four estimates for the intercept and the three regression coefficients.

The covariance matrix for the ten occasions has no restrictions. If we impose the restriction that all variances are equal, and that all covariances are equal, we have again the compound symmetry model. This shows that the model with occasion as fixed is one way to impose the compound symmetry structure on the random part of the model. Consequently, these models are nested, and we can use the overall chi-square test based on the deviance of the two models to test if the assumption of compound symmetry is tenable.

Models that assume a saturated model for the error structure are very complex. If there are k time points, the number of elements in the covariance matrix for the occasions is $k(k+1)/2$. So, with ten occasions, we have 55 (co)variance parameters to be estimated. If the assumption of compound symmetry is tenable, this model is preferable, because the random part contains only two parameters to be estimated. However, the compound symmetry model is very restrictive, because it assumes that there is one single value for all correlations between time points. This assumption is not very realistic, because the error term contains all omitted sources of variation, which may be correlated over time. Different assumptions about the autocorrelation over time lead to different structures of the covariance matrix across the occasions. For instance, it is reasonable to assume that occasions that are close together in time have a higher correlation than occasions that are far apart. Accordingly, the elements in the covariance matrix Σ should become smaller, the further away they are from the diagonal. Such a correlation structure is called a simplex. A more restricted version of the simplex is to assume that the autocorrelation between the occasions follow the first order autoregressive model

$$e_t = \rho e_{t-1} + \varepsilon \qquad \text{(x.12)}$$

where e_t is the error term at occasion t, ρ is the autocorrelation, and ε is a residual error with variance σ_ε^2. The error structure in equation (5.15) is a first order autoregressive process. This leads to a covariance matrix of the form:

$$\Sigma\,(Y) = \frac{\sigma_\varepsilon^2}{\left(1-\rho^2\right)} \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{k-1} \\ \rho & 1 & \rho & \cdots & \rho^{k-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{k-1} & \rho^{k-2} & \rho^{k-3} & \cdots & 1 \end{pmatrix} \qquad \text{(x.13)}$$

The first term $\sigma_\varepsilon^2/(1-\rho^2)$ is a constant, and the autocorrelation coefficient ρ is between –1 and +1, but typically positive. It is possible to have second order autoregressive processes, and other models for the error structure over time. The autoregressive model that produces the simplex in equation x.13 estimates one variance plus an autocorrelation. It is just as parsimonious as the compound symmetry model, but does not assume constant variances and covariances. The path diagram in Figure x.3 illustrates this model.

----- insert figure x.3 about here -----

Presenting the model as a SEM path diagram immediately suggests other structures for the dependency over time. Some often used structures are discussed by Hox (2002) and in more detail by Hedeker and Gibbons (2006). When multilevel regression is used, some structures, such as compound symmetry or the saturated model are easy to specify. Other structures are more difficult or impossible, unless the software producers have built in an option for specific structures. Many programs have these, for instance, both specialized programs like HLM and SuperMix and general packages like SPSS and SAS have a number of structures for dependency across time built in.

x.4 Choosing Between Structures and Combined Models

x.4.1 Choosing Between Structures

As explained in section x.2, allowing time varying covariates, including indicators for the measurement occasions, to vary across subjects, implies certain covariance structures over time. In addition, some software allows direct specification of specific covariance structures, for example an autoregressive model. As a consequence, an observed set of relationships over time can often be modeled about equally well by two different approaches. Any of such models is nested within the fully saturated model, which means that a likelihood ratio test or the equivalent chi-square deviance difference test can be used to asses their fit. However, a model allowing random slopes and a model directly specifying a covariance structure are not nested, and can only be compared using absolute fit indices such as Akaike's AIC of Schwarz's BIC. The AIC can be calculated from the deviance d and the number of estimated parameters q:

$$AIC = d + 2q, \tag{x.14}$$

and the BIC can be calculated as:

$$BIC = d + q\mathrm{Ln}(N) . \tag{x.15}$$

When the deviance goes down, indicating a better fit, both the AIC and the BIC also tend to go down. However, the AIC and the BIC include a penalty function based on the number of estimated parameters q. When the number of estimated parameters goes up, the AIC and BIC tend to go up too. For most sample sizes, the BIC places a larger penalty on complex models, which leads to a preference for smaller models compared to AIC. A problem with BIC in multilevel analysis is what the relevant sample size is: the number of groups or the total number of individuals? Most software that reports BIC uses the latter. However, in the context of longitudinal data, the number of subjects (i.e. the number of second level units) appears also reasonable. When a SEM package is used to specify the model, N is always the number of subjects. When a multilevel regression package is used, either level can supply the N, and the manual should be consulted to find out what the model does, or BIC must be calculated manually. Hedeker and Gibbons (2006), referring to Raftery (1995), advise to use the number of subjects, a practice we will follow here.

As an example, we model the state-anxiety data in several ways. The fixed part has three predictors: Trait Anxiety at the subject level and the two time-varying predictors Job Demands and Social Support. The first model in Table x.3 models the random part using a saturated model. Models 2–3 use the compound symmetry and the lag-1 autocorrelation model for the covariance structure. Model 4 derives the structure for the covariances from a random intercept plus a random slope for Job Demands. Model 5 will be described in section x.4.2. The table reports for each model the number of parameters estimated, and the overall fit statistics deviance, AIC, BIC based on the number of subjects, and BIC based on the total number of measurements (subjects×occasions). Full Maximum Likelihood estimation is used, so both regression coefficients and (co)variances enter the likelihood function.

All models are nested within the saturated model, so they can be tested against that model using the deviance difference test. The difference between the deviances of the models is a chi-square variate with degrees of freedom equal to the difference in number of astimated parameters. The column p in Table x.3 presents the p-value from the test of the model against the saturated model. Model 2 and 3 differ significantly from the saturated model, which means that they do not replicate the covariances well. Model four is significantly different from the saturated model at the 5% alpha level, but not at the 1% level. It does a better job at replicating the covariances than model two and three.

It is clear that if one explores these data from a covariance structure perspective, the likely choice is for a model with all predictors fixed and a lag-1 autocorrelation. From a random slopes perspective, the likely choice is for a model with a random intercept plus a random slope for the variable Job Demands. The fit indices point towards the random slope model.

x.4.2 Combined Models

The choice for a particular approach is not an either/or choice, the two approaches can be combined in a single model. The last row in Table x.3 presents the results for a model where the intercept and the slope for Job Demands have variation, and the remaining structure over time is modeled using a lag-1 autocorrelation. All fit indices prefer this model to the simpler Intercept + Slope model. Not all combinations of regression model and covariance structure are possible. For instance, a saturated model for the covariances over time leaves no place for intercept or slope variance. Highly restricted models such as compound symmetry or lag-1 autocorrelation, which both estimate only two parameters for the (co)variances, leave much room for varying slopes.

The last row in Table x.3 presents the fit information of a model that combines the random intercept plus slope model with a lag-1 autocorrelation structure. It fits very well, the *p*-value is 1.00, meaning that the model does not significantly differ from the saturated model. That is impressive, especially since the saturated model contains 59 parameters, and the combined model only nine. Since model four and five are also nested (model four is model five minus the autocorrelation part) they can also be compared using the formal test. The chi-square is 59.6, with 2 degrees of freedom, and the difference between the two models is clearly significant. Thus, model five a significant improvement on model four.

To assess the impact of various choices for the covariance structure on the fixed estimates, Table x.4 presents the parameter estimates for the fixed part, and for the most important parameters in the random part. The values within brackets are the standard errors. These are given for the saturated model, the autocorrelation model, the random intercept plus slope model, and the combined random intercept and slope plus

----- insert table x.4 about here -----

Although all corresponding estimates are similar, they are clearly not identical. The combined model could be improved by removing the evidently nonsignificant covariance between the intercept and the slope. If that is done, the deviance difference test for the variance of the intercept is significant ($\chi^2 = 10.08$, df=1, p=0.002). The associated changes in parameter estimates are very small, so they are not reported here. The varying coefficient for Job Demands can in principle be explained by adding a cross-level interaction of Job Demands with Trait Anxiety to the model. However, the coefficient for this interaction is not significant, and the variance of the Job Demands slopes remains unexplained.

The autocorrelation in the combined model is a conditional autocorrelation (Singer & Willet, 2003), conditional on the predictors in the model and the random effects of the intercept and of Job Demands. In the autocorrelation model, without the random intercept and slopes, the autocorrelation is much higher. Since there are no random effects in this model, the structure of the covariances over time must completely be explained by the autocorrelation function. In the combined model, part of the covariance structure is explained by the random effects in the model, which leaves a lower autocorrelation.

The interpretation of the model results for the combined model is the following. The time-varying predictors Job Demands and Social Support are significant. On days that Job Demands are high, a higher state anxiety is reported. On days that social support is high, a lower state anxiety is reported. Subjects who score high on trait anxiety report in general a higher state anxiety as well. The slope variation for Job Demands shows that some subjects are more sensitive to changes in Job Demands than others. There is a medium size correlation between the residuals for state anxiety from one day to the next, which means that state anxiety has a certain amount of short-term stability over time.

x.5 Conclusions

This chapter makes a distinction between longitudinal data where indicators for time are predictors in the model, to model growth or decline over time, and models where time is not a predictor, and time-varying covariates are used to model change over time. When time or other time-varying covariates have varying regression slopes, the dependency structure in the covariances over time is modeled implicitly, as a consequence of the estimated parameters in the random part. If there are no varying coefficients, the data must be modeled otherwise.

x.6 References

Bollen, K.A. & Curran, P.J. (2006). *Latent curve models*. Hoboken, NJ: Wiley.

Curran, P.J. (1997). *Supporting Documentation for Comparing Three Modern Approaches to Longitudinal Data Analysis: An Examination of a Single Developmental Sample***.** http://www.unc.edu/~curran/srcd.htmlhttp://www.unc.edu/~curran/srcd.html.

Duncan, T.E., Duncan, S.C. & Strycker, L.A. (2006). *An introduction to latent variable growth curve modeling*. Mahwah, NJ: Erlbaum.

Hedeker, D. & Gibbons, R.D. (2006). *Longitudinal data analysis*. Hoboken, NJ: Wiley.

Little, R.J.A. (1995). Modeling the drop-out mechanism in repeated measures studies. *Journal of the American Statistical Association, 90*, 1112-1121.

Maas, C.J.M., & Snijders, T.A.B. (2003). The multilevel approach to repeated measures with missing data. *Quality & Quantity*, 37, 71-89.

Raudenbush, S.W. (2002). Alternative covariance structures for polynomial models of individual growth and change. Pp. 25-57 in D.S. Moskowitz & S.L. Hersberger (eds). *Modeling intraindividual variability with repeated measures data*. Mahwah, NJ: Erlbaum.

Singer, J.D. & Willett, J.B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrences*. New York: Oxford.

Figure captions

Figure x.1. Path diagram corresponding to growth model with four occasions

Figure x.2. SEM diagram corresponding to Manova on 10 consecutive anxiety measures

Figure x.3. Path diagram for autoregressive model