Chapter 5
An Overview of the Autoregressive Latent Trajectory (ALT) Model

Kenneth A. Bollen and Catherine Zimmer

Abstract Autoregressive cross-lagged models and latent growth curve models are frequently applied to longitudinal or panel data. Though often presented as distinct and sometimes competing methods, the Autoregressive Latent Trajectory (ALT) model (Bollen and Curran, 2004) combines the primary features of each into a single model. This chapter: (1) presents the ALT model, (2) describes the situations when this model is appropriate, (3) provides an empirical example of the ALT model, and (4) gives the reader the input and output from an ALT model run on the empirical example. It concludes with a discussion of the limitations and extensions of the ALT model. Our focus is on repeated measures of continuous variables.

5.1 Introduction

There are two intuitive ways to approach the modeling of longitudinal data. The first relies on the idea and common observation that one of the best determinants of the current value of a variable is its value in the preceding period. So a student's reading performance in 2008 is well-determined by her reading performance in 2007, and this is true for all students in the population. This perspective can be formalized into what is known as an autoregressive model where the current value of a variable is determined by its past value. A second intuitively appealing method is to treat each subject as having a separate trajectory of change over time. Some cases might have a generally upward trend, others a downward trend, and still others might be

Kenneth A. Bollen
Odum Institute for Research in Social Science and Department of Sociology, University of North Carolina, Chapel Hill, USA
e-mail: bollen@unc.edu

Catherine Zimmer
Odum Institute for Research in Social Science, University of North Carolina, Chapel Hill, USA
e-mail: cathy_zimmer@unc.edu

K. van Montfort et al., *Longitudinal Research with Latent Variables*,
DOI 10.1007/978-3-642-11760-2_5,

relatively stable with regard to the outcome of interest. Here individual variability in change is permitted and each case can have different parameter values where these values describe the nature of the trajectory. This second approach we refer to as a latent (growth)[1] *curve* or *latent trajectory* model.

The autoregressive and latent curve models have long but largely independent histories. In the social and behavioral sciences autoregressive models were and are of substantial interest to economists who commonly use autoregressive time-series models to study economic indicators and lagged endogenous variables in panel data. The autoregressive models spread throughout the social and behavioral sciences beyond just economic applications. Anderson (1960), Humphreys (1960), Heise (1969), Wiley and Wiley (1970), Jöreskog (1970), and Werts, Jöreskog, and Linn (1971) provide just a few examples of publications that examined autoregressive models of a single outcome. Campbell (1963), Bohrnstedt (1969), Duncan (1969), Heise (1969), and Jöreskog (1979) are some of the earlier social science examples of authors who looked at autoregressive and cross-lagged models for two or more outcome variables in panel data. Kessler and Greenberg (1981) provided a book length treatment of these autoregressive and cross-lagged models. These have been and continue to be popular modeling approaches for longitudinal data.

The growth curve models of biostatistics have a long history (Bollen and Curran, 2006, pp. 9-14). The merger of the growth curve models with the factor analysis of longitudinal data resulted in the contemporary latent curve models and the resulting latent curve models date back to the 1950s (Bollen, 2007). Rao (1958) and Tucker (1958) were key works linking growth curve and exploratory factor analysis models. Meredith and Tisak (1984) was a seminal paper connecting confirmatory factor analysis to growth curve models leading to the latent curve model tradition that is influential today. In contrast to the autoregressive models, the repeated measures are reflective of an underlying pattern of change or trajectory. The trajectory is described by a set of parameters (e.g., random intercept and random slope) and these parameters can differ by individuals permitting a rich variety of trends across the cases in a sample.

Popularity of the autoregressive models preceded that of the growth curve models in the social and behavioral sciences. Early proponents of the growth curve model in these disciplines argued that the autoregressive and growth curve models were in direct competition (e.g., Bast and Reitsma, 1997; Kenny and Campbell, 1989; Rogosa and Willett, 1985) and some advocates argued that growth curve models were inherently superior to the autoregressive models (e.g., Rogosa, Brandt, and Zimowski, 1982, p. 744).

More recently the autoregressive and latent curve model have been combined into what is called the Autoregressive Latent Trajectory (ALT) model (Bollen and Curran, 2004; Curran and Bollen, 2001). The ALT model incorporates features of both the autoregressive and the latent curve model in a single framework. It is developed in recognition of the usefulness and appeal of each model and it permits modeling

[1] "Growth" suggests that the outcome variable is always increasing in magnitude and is misleading in those cases where the outcome decreases or is stable. For this reason, we sometimes omit this and refer to the models as latent curve or latent trajectory models.

data that has features of both models. Furthermore, it permits tests that provide information on whether the data more closely conform to the autoregressive or to the latent curve model. So if one or the other models is best, the ALT model will help to reveal that, whereas if both processes are operating both can be accommodated by the ALT model.

It also is important to distinguish the ALT model from a more established one that is a latent curve model with an autoregressive disturbance. For example, Chi and Reinsel (1989), Browne and du Toit (1991), Diggle, Liang and Zeger (1994), and Goldstein, Healy and Rasbash (1994) discuss modifications of the standard growth curve model to permit an autoregressive disturbance. In these types of models the autoregression of the disturbance is a type of nuisance association that is relegated to the disturbance and it is given little substantive explanation. In the ALT model the autoregressive relation is between the repeated measures, not the disturbances.[2] Furthermore, the lagged effect of the earlier value on the current value should be substantively meaningful when using the ALT model.

The purposes of this chapter are: (1) to present the ALT model, (2) to describe the situations when this model is appropriate, (3) to provide an empirical example of the ALT model, and (4) to give the reader the input and output from an ALT model run on the empirical example. Much of the technical presentation of the ALT model is based on Bollen and Curran (2004; 2006). Applications of the ALT are in many fields, such as psychology to study developmental psychopathology (Curran and Willoughby, 2003), daily anxiety and panic expectancy (Rodebaugh, Curran, and Chambliss, 2002), job performance over time (Zyphur, Chaturvedi and Arvey, 2008), and changes in eating behavior among first-year undergraduate women (Boyd, 2007). Addiction researchers have found the ALT model useful for studying how adolescent and peer substance use changes over time and affects each other (Simons-Morton and Chen, 2006). Wan, Zhang and Unruh (2006) used the ALT model to investigate outcome improvement in residents of nursing homes.

The next several sections present single variable and two variable ALT models, a general equation for all models, the implied moment matrices, and a section on the estimation and testing of these models. After these we present an empirical example. A conclusion summarizes the ALT model and its use.

[2] Hamaker (2005) has an interesting paper where she shows that an ALT model that has an equal autoregressive coefficient and is not written with the first wave outcome as predetermined is mathematically equivalent to an alternative growth curve model with autoregressive disturbances. These two forms of the model would have different substantive meanings in that the ALT model hypothesizes that the lagged repeated measure has an impact on the current repeated measure whereas the autoregressive disturbance model assumes that the prior disturbance influences the current disturbance. In the autoregressive disturbance model there is no direct effect of the repeated measures on other repeated measures and only a direct effect between disturbances. Also the equivalency does not hold if the autoregressive parameter differs across waves or if the first wave of the outcome is treated as a predetermined variable as recommended in Bollen and Curran (2004).

5.2 Autoregressive Latent Trajectory (ALT) Model

5.2.1 Single Variable Unconditional ALT Model

In this subsection we present the single variable, unconditional ALT model. By single variable we mean that there is only one outcome observed over time. By unconditional, we refer to the fact that the model has no explanatory variables or covariates that determine the random intercepts, random slopes, or the repeated measures other than the lagged value of the repeated measures. Suppose that y_{it} is the repeated measure of y for the ith observation at the tth time point. The ALT model is

$$y_{it} = \alpha_i + \Lambda_t \beta_i + \rho_{t,t-1} y_{i,t-1} + \varepsilon_{it} \tag{5.1}$$

where the i indexes the individual in the sample and the t indexes the time with $t = 2,3,...,T$. The α_i is the random intercept, β_i is the random slope, and Λ_t is the time trend variable that describes the pattern of growth so that for a linear growth model it would $0,1,2,\dots$. The autoregressive parameter is $\rho_{t,t-1}$,[3] $y_{i,t-1}$ is the lagged value of y, and ε_{it} is the disturbance of the equation. We assume $E(\varepsilon_{it}) = 0$, $COV(\varepsilon_{it}, y_{i,t-1}) = 0$, $COV(\varepsilon_{it}, \beta_i) = 0$, and $COV(\varepsilon_{it}, \alpha_i) = 0$. We also assume $E(\varepsilon_{it}, \varepsilon_{jt}) = 0$ for all t and $i \neq j$, $E(\varepsilon_{it}, \varepsilon_{it}) = \sigma^2_{\varepsilon_t}$ for each t and i, and $COV(\varepsilon_{it}, \varepsilon_{i,t+k}) = 0$ for $k \neq 0$ though in some cases this latter restriction could be removed.

If we assume that $VAR(\beta_i)$, $VAR(\alpha_i)$, and $E(\beta_i)$ are all zero, then we get

$$y_{it} = \alpha + \rho_{t,t-1} y_{i,t-1} + \varepsilon_{it} \tag{5.2}$$

which is an autoregressive model with an intercept that does not change over time. If the true model corresponds to an autoregressive model, then we would expect the variances of the random intercepts and random slopes, and the mean of the slope to be zero in the ALT model.

Alternatively, suppose that $\rho_{t,t-1}$ in the ALT model is zero for all t. Now the resulting model is

$$y_{it} = \alpha_i + \Lambda_t \beta_i + \varepsilon_{it} \tag{5.3}$$

which corresponds to a latent curve model with random intercept α_i and random slope β_i.

These preceding constraints give us information on whether the autoregressive or latent curve model are sufficient to describe data or whether the full ALT model is required. The basic task is to estimate the ALT model. If the variances of the random intercepts and random slopes and the mean of the slope are essentially zero, then the

---

[3] In general we assume that $|\rho_{t,t-1}| < 1$ to insure that y_{it} does not grow infinitely as t goes to infinity. In the time series literature, this is a stationarity condition (e.g., Box and Jenkins, 1976). In nonstationary data, the autoregressive parameter can equal or exceed one but in our experience such nonstationary series are rare in panel data. This condition is not critical for our developments here.

autoregressive model is appropriate as long as $\rho_{t,t-1}$ is not zero. Alternatively, if the random intercepts and random slopes have nonzero variances and $\rho_{t,t-1}$ is always zero, then the latent curve model is preferred. If neither of these conditions are true, then the full ALT model should be considered.

One complication that we have not mentioned has to do with the first wave of data. Although Bollen and Curran (2004) show how to model all repeated measures as endogenous variables, they suggest that there are some useful simplifications that result when the first wave of the outcome is treated as a predetermined variable as is shown in Figure 5.1. One advantage follows in that we cannot estimate equation (5.1) for the first wave of data since by definition we do not have the lagged value of the first wave of the outcome variable. Treating this first wave as predetermined bypasses this problem. The equation for the first wave outcome variable then becomes

$$y_{i1} = \nu_1 + \varepsilon_{i1} \tag{5.4}$$

where ν_1 is the mean of y_{i1}.

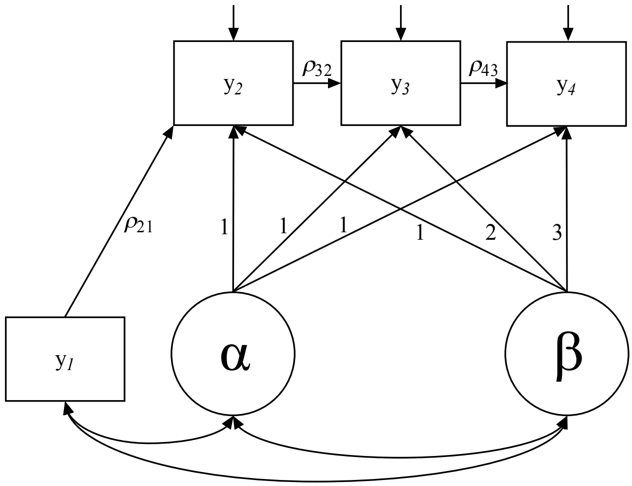

Fig. 5.1 Autoregressive Latent Trajectory (ALT) model with single variable over four waves and y_1 predetermined.

The other two equations to make the single variable ALT model complete are

$$\alpha_i = \mu_\alpha + \zeta_{\alpha i} \tag{5.5}$$
$$\beta_i = \mu_\beta + \zeta_{\beta i} \tag{5.6}$$

where μ_α and μ_β are the means of the random intercepts and random slopes, respectively, and $\zeta_{\alpha i}$ and $\zeta_{\beta i}$ are the random deviations around the respective means. The predetermined y_{i1}, α_i, and β_i are allowed to correlate.

Given the relations between the ALT model, the latent curve model, and the autoregressive model just described helps in interpreting the ALT model in equation (5.1). Consider first the latent curve model without autoregressive effects as in equation (5.3). In a latent curve model each case can have a distinct trajectory of the outcome variable. The trajectories are captured by having a random intercept and random slope that can vary by case. Once you control for the random intercepts and random slopes there is no influence of prior values of y on current values of y, that is, there is no autoregressive impact net of the trajectory parameters.

Alternatively, in the pure autoregressive model as in equation (5.2), the current y_{it} is driven by the past $y_{i,t-1}$ plus a random disturbance. Each case in the sample has the same autoregressive coefficient, $\rho_{t,t-1}$. Once the prior value of y is controlled, there are no individual trajectories for the cases in the sample.

From one perspective the ALT model is a latent curve model with random intercepts and random slopes where each individual can have a distinct trajectory. But now once we control for the random intercepts and random slopes there remains an autoregressive relationship between the ys. Taking a different perspective, the ALT model is an autoregressive model where the lagged value of a repeated measure partially determines the current value, but even taking account of the autoregressive relation each case can have a distinct trajectory. To understand the change in y we need to know the prior value of y and the individual trajectory of change for that individual. In other words both an autoregressive and growth curve model characterize the process. Neither a LCM or an autoregressive one alone is sufficient to describe the change.

5.2.2 Single Variable Conditional ALT Model

So far we have limited our description to an unconditional model where the random intercepts (α_i), random slopes (β_i), and the first wave of the repeated measures (y_{i1}) do not include covariates that determine them and they are only represented as a function of their means and deviations from their respective means (see eqs. (5.4) to (5.6)). A natural extension allows for covariates to predict α_i, β_i, and y_{i1}. To illustrate consider the incorporation of two time invariant exogenous predictors, z_{i1} and z_{i2} (though it is easy to generalize this model to any number of covariates). We modify equations (5.4) to (5.6) by adding these covariates resulting in

$$\alpha_i = \mu_\alpha + \gamma_{\alpha 1} z_{i1} + \gamma_{\alpha 2} z_{i2} + \zeta_{\alpha i} \tag{5.7}$$

$$\beta_i = \mu_\beta + \gamma_{\beta 1} z_{i1} + \gamma_{\beta 2} z_{i2} + \zeta_{\beta i} \tag{5.8}$$

$$y_{i1} = \nu_1 + \gamma_{y1} z_{i1} + \gamma_{y2} z_{i2} + \varepsilon_{i1} \tag{5.9}$$

where μ_α, μ_β, and ν_1 now represent regression intercepts rather than unconditional means. The γs represent the fixed regressions of the random intercepts (α_i), random slopes (β_i), and the predetermined y_{i1} on the two covariates. Figure 5.2 is a path diagram of the conditional ALT model for four waves of data and with two covariates. We assume that the disturbances (i.e., $\zeta_{\alpha i}$, $\zeta_{\beta i}$, ε_{i1}) have zero means and are uncorrelated with the exogenous variables (zs). Further, we permit $\zeta_{\alpha i}$, $\zeta_{\beta i}$, ε_{i1} to correlate with each other, but none of these is correlated with later values of ε_{it} where $t = 2, 3, \ldots$. Finally, we assume the exogenous variables are measured without error.

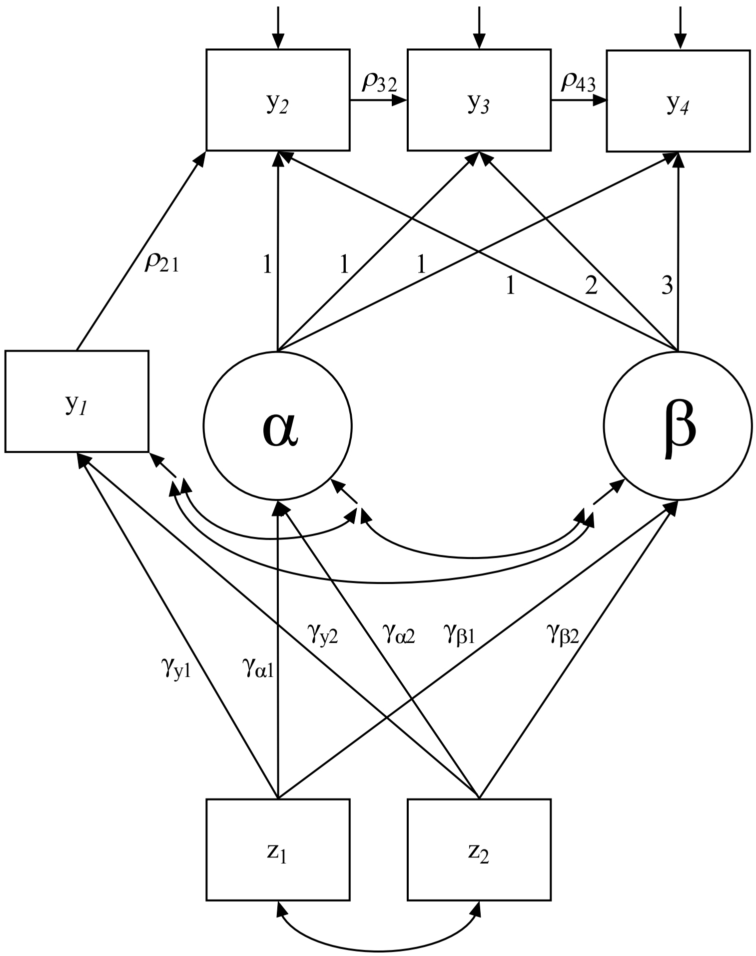

Fig. 5.2 Conditional Autoregressive Latent Trajectory (ALT) model with single variable over four waves and two covariates.

5.2.3 Bivariate Unconditional ALT Model

In the conditional univariate ALT model, we considered the influences of one or more time invariant covariates. However, there are many instances in which there might be interest in the relationship between two repeated measures, each of which is functionally related to the passage of time. We can extend the single repeated ALT model to include two or more repeated measures say, y_{it} and x_{it}. We write the bivariate ALT model for $t = 2, 3, ..., T$ as

$$y_{it} = \alpha_{yi} + \Lambda_{yt2}\beta_{yi} + \rho_{y_t y_{t-1}} y_{i,t-1} + \rho_{y_t x_{t-1}} x_{i,t-1} + \varepsilon_{yit} \tag{5.10}$$

$$x_{it} = \alpha_{xi} + \Lambda_{xt2}\beta_{xi} + \rho_{x_t y_{t-1}} y_{i,t-1} + \rho_{x_t x_{t-1}} x_{i,t-1} + \varepsilon_{xit} \tag{5.11}$$

We maintain similar assumptions about the disturbances (ε's) as before (means of zero, not autocorrelated, uncorrelated with the right-hand side variables and random coefficients). We permit some ε_{yit} to correlate with ε_{xit} as long as model identification is maintained. For this model we treat the y_{i1} and x_{i1} variables as predetermined and the random intercepts and random slopes as exogenous. Their equations are

$$y_{i1} = \nu_{y1} + \varepsilon_{yi1} \tag{5.12}$$

$$x_{i1} = \nu_{x1} + \varepsilon_{xi1} \tag{5.13}$$

$$\alpha_{yi} = \mu_{y\alpha} + \zeta_{y\alpha i} \tag{5.14}$$

$$\beta_{yi} = \mu_{y\beta} + \zeta_{y\beta i} \tag{5.15}$$

$$\alpha_{xi} = \mu_{x\alpha} + \zeta_{x\alpha i} \tag{5.16}$$

$$\beta_{xi} = \mu_{x\beta} + \zeta_{x\beta i} \tag{5.17}$$

All disturbances in these equations have means of zero. Generally, we permit ε_{yi1}, ε_{xi1}, $\zeta_{y\alpha i}$, $\zeta_{y\beta i}$, $\zeta_{x\alpha i}$, and $\zeta_{x\beta i}$ to correlate with each other, but these are assumed not to correlate with ε_{yit} and ε_{xit} for $t = 2, 3, ..., T$. Figure 5.3 is the path diagram of a bivariate unconditonal ALT model for four waves of data.

Each of the equations (5.10) and (5.11) are similar to the unconditional single variable ALT model except for the extra cross-lag term either $\rho_{y_t x_{t-1}} x_{i,t-1}$ in equation (5.10) or $\rho_{x_t y_{t-1}} y_{i,t-1}$ in equation (5.11). This is a noteworthy difference in that it permits the repeated measure from one series to directly impact the repeated measure of another. The bivariate ALT model not only allows the lagged dependent variable to enter the equation along with the random intercepts and random slopes, but it also permits a second repeated measure to have an impact once we control for the lagged and latent curve effects on the repeated measure. The flexibility of this model is considerable in that depending on the result of estimation the model could be an autoregressive model (when $VAR(\beta_i)$, $VAR(\alpha_i)$, $E(\beta_i)$, $\rho_{y_t x_{t-1}}$, and $\rho_{x_t y_{t-1}}$ all equal zero), a cross-lag model (when $VAR(\beta_i)$, $VAR(\alpha_i)$, and $E(\beta_i)$ all equal zero) or a latent curve model ($\rho_{y_t y_{t-1}}$, $\rho_{y_t x_{t-1}}$, $\rho_{x_t x_{t-1}}$ and $\rho_{x_t y_{t-1}}$ all zero).

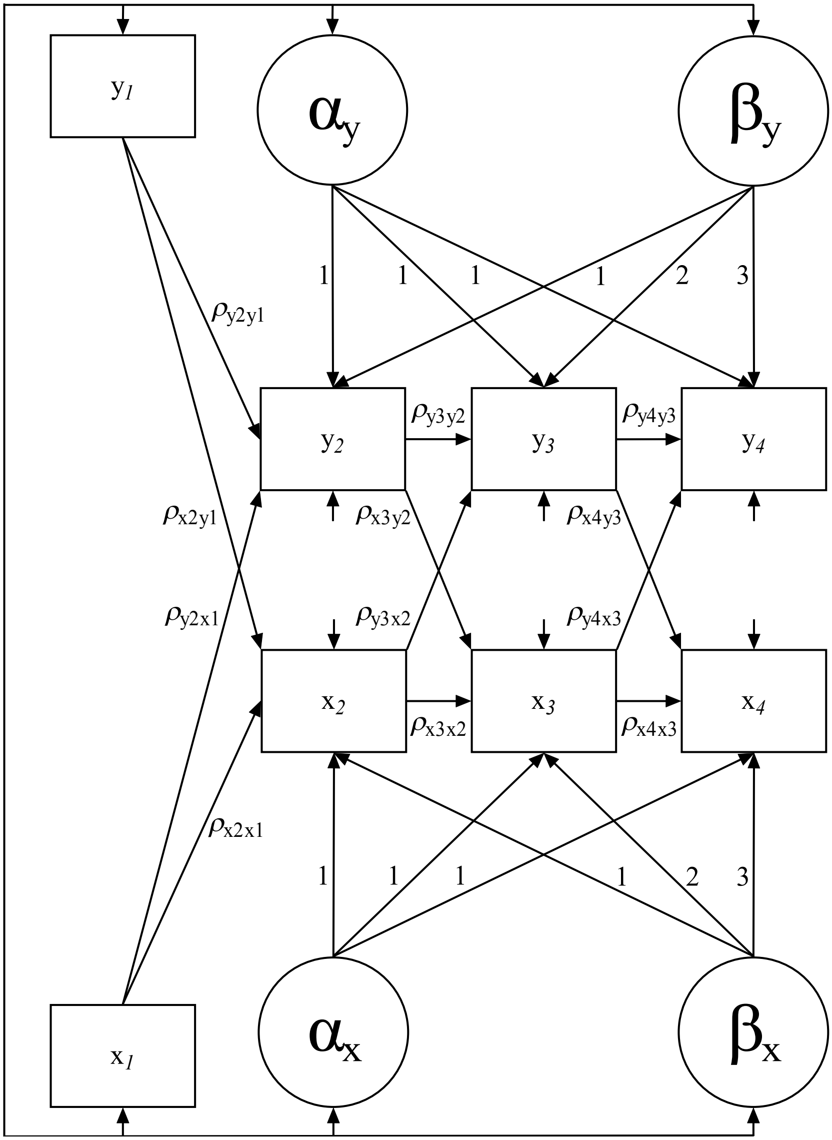

Fig. 5.3 Autoregressive Latent Trajectory (ALT) model for two variables over four waves and lagged effects between observed variables.

Furthermore additional or different lagged values of the repeated measures could enter these equations as dictated by the substantive knowledge driving the research. For instance, if the current value of a repeated measure was affected not only by say, y_{t-1}, but also by y_{t-2}, then both y_{t-1} and y_{t-2} should be included as predictors of y_t in the equations. In addition, the type of model devised for each repeated measure need not be the same. So if a latent curve model without autoregressive terms fits the x series best and an ALT model is needed for the y series, there is no reason not to include different structures for each repeated measure.

5.2.4 Bivariate Conditional ALT Model

As we described for the univariate conditional ALT model, we can incorporate one or more exogenous predictors in the bivariate ALT model as well. This is again accomplished by the extension of the equations for the random trajectories. Specifically, we modify equations (5.14) through (5.17) to include time invariant covariates z_{i1} and z_{i2} such that

$$\alpha_{yi} = \mu_{y\alpha} + \gamma_{\alpha y1} z_{i1} + \gamma_{\alpha y2} z_{i2} + \zeta_{y\alpha i} \tag{5.18}$$
$$\beta_{yi} = \mu_{y\beta} + \gamma_{\beta y1} z_{i1} + \gamma_{\beta y2} z_{i2} + \zeta_{y\beta i} \tag{5.19}$$

and

$$\alpha_{xi} = \mu_{x\alpha} + \gamma_{\alpha x1} z_{i1} + \gamma_{\alpha x2} z_{i2} + \zeta_{x\alpha i} \tag{5.20}$$
$$\beta_{xi} = \mu_{x\beta} + \gamma_{\beta x1} z_{i1} + \gamma_{\beta x2} z_{i2} + \zeta_{x\beta i} \tag{5.21}$$

As before, the set of gammas represent the fixed regressions of the random trajectory components on the two correlated exogenous variables. It is possible to have the random intercepts or random slopes as explanatory variables in equations (5.18) to (5.21). For instance, the random intercept from the y series (α_{yi}) might affect the random slope of the x series leading to $\beta_{xi} = \mu_{x\beta} + \gamma_{\beta_x \alpha_y} \alpha_{yi} + \gamma_{\beta x1} z_{i1} + \gamma_{\beta x2} z_{i2} + \zeta_{x\beta i}$ or the slope of one series could alter the slope of the other, for example, $\beta_{yi} = \mu_{y\beta} + \gamma_{\beta_y \beta_x} \beta_x + \gamma_{\beta y1} z_{i1} + \gamma_{\beta y2} z_{i2} + \zeta_{y\beta i}$.

In the bivariate unconditional ALT model, we let the initial repeated measures correlate with the random intercepts and random slopes. In the conditional bivariate ALT model, we must regress x_{i1} and y_{i1} on the set of exogenous measures. Thus, the equations for the initial measures for x_{i1} and y_{i1} are

$$y_{i1} = \nu_{y1} + \gamma_{y1} z_{i1} + \gamma_{y2} z_{i2} + \varepsilon_{yi1} \tag{5.22}$$
$$x_{i1} = \nu_{x1} + \gamma_{x1} z_{i1} + \gamma_{x2} z_{i2} + \varepsilon_{xi1} \tag{5.23}$$

The same assumptions described for the univariate conditional ALT model hold here as well.

5.3 General Equation for All Models

Up to this point we have presented unconditional and conditional ALT models for a single and two repeated measures using a scalar notation. These and variants of these models are expressable in a general matrix notation that is convenient for presenting the estimation and assessment of fit of these models. The matrix model is (Bollen and Curran, 2004):

$$\boldsymbol{\eta}_i = \boldsymbol{\mu} + \mathbf{B}\boldsymbol{\eta}_i + \boldsymbol{\zeta}_i \tag{5.24}$$

$$\mathbf{o}_i = \mathbf{P}\boldsymbol{\eta}_i \tag{5.25}$$

where the first equation provides the structural relations between variables, $\boldsymbol{\eta}_i$ is a vector that contains both the repeated measures and the random intercepts and random slopes, $\boldsymbol{\mu}$ is a vector of means or intercepts, $\mathbf{B}$ is a coefficient matrix that gives the coefficients for the relationships of $\boldsymbol{\eta}_i$s on each other, and $\boldsymbol{\zeta}_i$ is the disturbance vector for the variables in $\boldsymbol{\eta}_i$. We assume that $E(\boldsymbol{\zeta}_i) = \mathbf{0}$. The nature of the covariances of $\boldsymbol{\zeta}_i$ with $\boldsymbol{\eta}_i$ will vary depending on the model, but for identification purposes at least some of these covariances will be zero or known values. The second equation functions to pick out the observed variables, $\mathbf{o}_i$, from the latent variables of equation 5.24.

In more detail,

$$\boldsymbol{\eta}_i = \begin{bmatrix} \mathbf{y}_i \\ \mathbf{x}_i \\ \mathbf{z}_i \\ \boldsymbol{\alpha}_i \\ \boldsymbol{\beta}_i \end{bmatrix} \tag{5.26}$$

where $\mathbf{y}_i$ and $\mathbf{x}_i$ are two variables repeatedly measured for T time periods, $\mathbf{z}_i$ is a q x 1 vector of exogenous determinants of the latent trajectory parameters or of the repeated measures, α_i is the 2 x 1 vector of α_{yi} and α_{xi}, the random intercepts for the two sets of repeated measures, and β_i is the 2 x 1 vector of β_{yi} and β_{xi} the random slopes for the two repeated measures. The $\boldsymbol{\mu}$ vector is

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_y \\ \boldsymbol{\mu}_x \\ \boldsymbol{\mu}_z \\ \boldsymbol{\mu}_\alpha \\ \boldsymbol{\mu}_\beta \end{bmatrix} \tag{5.27}$$

where $\boldsymbol{\mu}_y$ and $\boldsymbol{\mu}_x$ are vectors of means/intercepts for the $\mathbf{y}_i$ and $\mathbf{x}_i$ observed repeated measures, $\boldsymbol{\mu}_z$ is the vector of means for the exogenous covariates in the model, $\boldsymbol{\mu}_\alpha$ is a vector of means/intercepts for the random intercepts, α_{yi} and α_{xi}, and $\boldsymbol{\mu}_\beta$ is a vector of the means/intercepts of β_{yi} and β_{xi}.

For the ALT model, the $\mathbf{B}$ matrix is

$$\mathbf{B} = \begin{bmatrix} \mathbf{B_{yy}} & \mathbf{B_{yx}} & \mathbf{B_{yz}} & \mathbf{B_{y\alpha}} & \mathbf{B_{y\beta}} \\ \mathbf{B_{xy}} & \mathbf{B_{xx}} & \mathbf{B_{xz}} & \mathbf{B_{x\alpha}} & \mathbf{B_{x\beta}} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{B_{\alpha z}} & \mathbf{B_{\alpha\alpha}} & \mathbf{B_{\alpha\beta}} \\ \mathbf{0} & \mathbf{0} & \mathbf{B_{\beta z}} & \mathbf{B_{\beta\alpha}} & \mathbf{B_{\beta\beta}} \end{bmatrix} \tag{5.28}$$

where the double subscript notation in the partition matrix indicates that the submatrix contains those coefficients related to effects among the subscripted variables. For instance, $\mathbf{B_{yy}}$ contains the effects of the repeated y variables on each other, and $\mathbf{B_{\beta z}}$ contains the impact of the exogenous $\mathbf{z}_i$ on the random slopes, β_{yi} and β_{xi}, for the ys and xs. The $\mathbf{z}_i$ consists of exogenous variables.

The disturbance vector for equation 5.24 is

$$\boldsymbol{\zeta}_{\mathbf{i}} = \begin{bmatrix} \boldsymbol{\varepsilon}_{yi} \\ \boldsymbol{\varepsilon}_{xi} \\ \boldsymbol{\varepsilon}_{zi} \\ \boldsymbol{\zeta}_{\alpha i} \\ \boldsymbol{\zeta}_{\beta i} \end{bmatrix} \tag{5.29}$$

with covariance matrix $\boldsymbol{\Sigma}_{\zeta\zeta}$. Since $\mathbf{z}_i$ is exogenous, the variance of $\boldsymbol{\varepsilon}_{zi}$ is equivalent to the variance of $\mathbf{z}_i$.

The $\mathbf{P}$ matrix is

$$\mathbf{P} = \begin{bmatrix} \mathbf{I}_T & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_T & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_q & \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{5.30}$$

where $\mathbf{I}_T$ is a T x T identity matrix with dimensions that depend on the number of repeated measures and $\mathbf{I}_q$ is a q x q identity matrix with q exogenous variables. The matrix picks out the observed variables in a given model where $\mathbf{o}_i$ is

$$\mathbf{o}_i = \begin{bmatrix} \mathbf{y}_i \\ \mathbf{x}_i \\ \mathbf{z}_i \end{bmatrix} \tag{5.31}$$

Bollen and Curran (2004) demonstrate how this matrix expression enables a researcher to incorporate all of the models discussed as well as others. For instance, the standard autoregressive model for a single repeated measure has

$$\boldsymbol{\eta}_i = [\mathbf{y}_i] \tag{5.32}$$

$$\boldsymbol{\mu} = [\boldsymbol{\mu}_{\mathbf{y}}] \tag{5.33}$$

$$\mathbf{B} = [\mathbf{B_{yy}}] \tag{5.34}$$

$$\boldsymbol{\zeta}_{\mathbf{i}} = [\boldsymbol{\varepsilon}_i] \tag{5.35}$$

$$\mathbf{o}_i = \boldsymbol{\eta}_i \tag{5.36}$$

with

$$\mathbf{B_{yy}} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ \rho_{21} & 0 & 0 & \cdots & 0 \\ 0 & \rho_{32} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \rho_{T,T-1} & 0 \end{bmatrix} \tag{5.37}$$

to capture a first order autoregressive relation.

The unconditional latent curve model has

$$\boldsymbol{\eta}_i = \begin{bmatrix} \mathbf{y}_i \\ \alpha_i \\ \beta_i \end{bmatrix} \tag{5.38}$$

$$\boldsymbol{\mu} = \begin{bmatrix} \mathbf{0} \\ \mu_\alpha \\ \mu_\beta \end{bmatrix} \tag{5.39}$$

where the $\mathbf{0}$ vector in $\boldsymbol{\mu}$ represents the zero fixed intercepts for the repeated measures in a latent trajectory model. The $\mathbf{B}$ matrix is

$$\mathbf{B} = \begin{bmatrix} \mathbf{0} & \mathbf{B_{y\alpha}} & \mathbf{B_{y\beta}} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{5.40}$$

$$\mathbf{B_{y\alpha}} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \qquad \mathbf{B_{y\beta}} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ T-1 \end{bmatrix} \tag{5.41}$$

The $\boldsymbol{\zeta}_\mathbf{i}$ and $\mathbf{P}$ matrices are

$$\boldsymbol{\zeta}_\mathbf{i} = \begin{bmatrix} \boldsymbol{\varepsilon}_i \\ \zeta_{\alpha i} \\ \zeta_{\beta i} \end{bmatrix} \tag{5.42}$$

$$\mathbf{P} = \begin{bmatrix} \mathbf{I} \; \mathbf{0} \; \mathbf{0} \end{bmatrix} \tag{5.43}$$

As a last example the unconditional univariate ALT model has

$$\mathbf{B} = \begin{bmatrix} \mathbf{B_{yy}} & \mathbf{B_{y\alpha}} & \mathbf{B_{y\beta}} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{5.44}$$

where $\mathbf{B_{yy}}$ is the same as equation 5.37 and

$$\mathbf{B}_{\mathbf{y}\alpha} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \qquad \mathbf{B}_{\mathbf{y}\beta} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ T-1 \end{bmatrix} \tag{5.45}$$

for a model where y_{1i} is predetermined. Furthermore

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_y \\ \mu_\alpha \\ \mu_\beta \end{bmatrix} \tag{5.46}$$

with

$$\boldsymbol{\mu}_y = \begin{bmatrix} \mu_{y1} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{5.47}$$

and

$$\boldsymbol{\zeta}_{\mathbf{i}} = \begin{bmatrix} \boldsymbol{\varepsilon}_i \\ \zeta_{\alpha i} \\ \zeta_{\beta i} \end{bmatrix} \tag{5.48}$$

The variances of ε_{1i}, $\zeta_{\alpha i}$, and $\zeta_{\beta i}$ are equal to the variances of the predetermined variables, y_{1i}, α_i, and β_i, respectively.

5.4 Implied Moment Matrices

Structural equation models (SEMs) typically involve expressing the means and covariance matrix of the observed variables as a function of the parameters ($\boldsymbol{\theta}$) in a model. These expressions of the implied mean vector ($\boldsymbol{\mu}(\boldsymbol{\theta})$) and the implied covariance matrix ($\boldsymbol{\Sigma}(\boldsymbol{\theta})$) also are referred to as the implied moment matrices and they are useful in estimation and the assessment of model fit. Bollen and Curran (2004) show that the implied mean vector is

$$\boldsymbol{\mu}(\boldsymbol{\theta}) = E(\mathbf{o}_i) = \mathbf{P}(\mathbf{I}-\mathbf{B})^{-1}\boldsymbol{\mu} \tag{5.49}$$

and the implied covariance matrix of observed variables is

$$\begin{aligned} \boldsymbol{\Sigma}(\boldsymbol{\theta}) &= [E(\mathbf{o}_i\mathbf{o}_i') - E(\mathbf{o}_i)E(\mathbf{o}_i')] \\ &= \mathbf{P}(\mathbf{I}-\mathbf{B})^{-1}\boldsymbol{\Sigma}_{\zeta\zeta}(\mathbf{I}-\mathbf{B})^{-1\prime}\mathbf{P}' \end{aligned} \tag{5.50}$$

The exact value of these implied moments depends on the value of the matrices for the particular type of ALT model, but once the matrices that correspond to the model of interest are substituted into these expressions, the implied moments are revealed.

One valuable aspect of the implied moment matrices is in determining the identification of the model parameters. A parameter is identified if it is possible to find a unique value for it. In SEMs we have

$$\boldsymbol{\mu} = \boldsymbol{\mu}(\boldsymbol{\theta}) \tag{5.51}$$

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta}) \tag{5.52}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the mean vector and covariance matrix of the observed variables and we have already defined their corresponding implied moments. Demonstrating that each $\boldsymbol{\theta}$ is solvable as a unique value of a function of one or more elements of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ demonstrates that the parameters are identified. In general, we require four waves of data for the ALT model to be identified if the autoregressive parameter is equal over time and five waves without the equality restriction on the autoregression coefficient. If there are only three waves of data, then y_{i1} must be made endogenous and the coefficients for the paths from α_i and β_i to y_{i1} require nonlinear constraints for estimation. Bollen and Curran (2004) discuss this special case in more detail.

5.5 Estimation and Testing

SEMs are estimable with a wide variety of estimators. The most appropriate estimator depends on whether the endogenous observed variables are continuous or categorical and the distribution of these variables. In the most straightforward case of continuous endogenous variables, the Full Information Maximum Likelihood (FIML) estimator is available in all SEM software:

$$F_{ml} = \ln|\boldsymbol{\Sigma}(\boldsymbol{\theta})| + tr[\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{S}] - \ln|\mathbf{S}| - p + (\bar{\mathbf{z}} - \boldsymbol{\mu}(\boldsymbol{\theta}))'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})(\bar{\mathbf{z}} - \boldsymbol{\mu}(\boldsymbol{\theta})) \tag{5.53}$$

where $\boldsymbol{\theta}$ is a vector that contains all of the parameters (i.e., coefficients, variances, and covariances of exogenous variables and errors) in the model that we wish to estimate, $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ is the covariance matrix of the observed variables that is implied by the model structure, $\boldsymbol{\mu}(\boldsymbol{\theta})$ is the mean vector of the observed variables implied by the model, $\mathbf{S}$ is the sample covariance matrix of the observed variables, $\bar{\mathbf{z}}$ is the vector of sample means of the observed variables, and p is the number of observed variables in the model. The implied covariance matrix $[\boldsymbol{\Sigma}(\boldsymbol{\theta})]$ and the implied mean vector $[\boldsymbol{\mu}(\boldsymbol{\theta})]$ are in 5.49 and 5.50, respectively.

The classical derivation of F_{ml} begins with the assumption that the observed variables come from multivariate normal distributions (see, e.g., Bollen, 1989a, pp. 131-135). The FIML estimator of the parameters, $\widehat{\boldsymbol{\theta}}$, has several desirable properties: the estimator is consistent, asymptotically unbiased, asymptotically normally distributed, asymptotically efficient, and its covariance matrix equals the inverse of the information matrix (Lawley and Maxwell 1971). Fortunately, the FIML has desirable properties under less restrictive conditions. Browne (1984) proves that the preceding properties hold as long as the observed variables come from distributions with no excess multivariate kurtosis. There also are robustness studies that provide conditions where many of these properties hold even with excess multivariate kurtosis (see e.g., Satorra, 1990). Even when the robustness conditions fail there are corrections to the significance tests and bootstrapping procedures that permit significance tests (e.g., Satorra and Bentler, 1988; Bollen and Stine, 1990; 1993). Thus, with continuous outcome variables, estimation is possible even with excess multivariate kurtosis. Categorical dependent observed variables require procedures that take account of their categorical nature, but this is beyond the scope of our chapter. See Bollen and Curran (2006, Ch. 8) for discussion.

A first step in assessing model fit is a test of $H_0 : \mathbf{\Sigma} = \mathbf{\Sigma}(\boldsymbol{\theta})$ and $\boldsymbol{\mu} = \boldsymbol{\mu}(\boldsymbol{\theta})$ where $\mathbf{\Sigma}$ is the population covariance matrix of the observed variables, $\mathbf{\Sigma}(\boldsymbol{\theta})$ is the covariance matrix implied by the model that is a function of the parameters of the model, $\boldsymbol{\mu}$ is the population mean vector of the observed variables, and $\boldsymbol{\mu}(\boldsymbol{\theta})$ is the implied mean vector that is a function of the model parameters. These implied moment matrices were described above. If the model is true, then H_0 should be true. If the model structure is incorrect, then we should reject H_0. The test statistic of $T_{ml} = F_{ml}(N-1)$ is asymptotically distributed as a χ^2 with degrees of freedom $df = (p(p+1)/2+p)-t$ where p is the number of observed variables and t is the number of estimated parameters. A significant chi-square test statistic is evidence against $H_0 : \mathbf{\Sigma} = \mathbf{\Sigma}(\boldsymbol{\theta})$ and $\boldsymbol{\mu} = \boldsymbol{\mu}(\boldsymbol{\theta})$ while a nonsignificant test statistic is consistent with the null hypothesis and hence the validity of the model. It is possible to compare two or more nested models where the parameters of one model are a restrictive form of the parameters of another. For instance, if we had an ALT model with no restrictions on the autoregressive parameter and a second identical to the first except that the autoregressive parameters were constrained to be equal, then the equal autoregressive ALT model would be nested in the ALT model where the autoregressive parameters were freely estimated. The difference in the chi-square test statistics for these individual ALT models would itself be asymptotically distributed as a chi-square variate with degrees of freedom equal to the difference in the degrees of freedom of the two models. The null hypothesis in this comparison of nested models is that the model with the greatest number of restrictions fits as well as the less restrictive model. A significant chi-square would be evidence in favor of the less restrictive model whereas a nonsignificant chi-square is evidence favoring the more restrictive model.

In practice, the chi-square test statistics are not the sole means of assessing model fit. Even if we use test statistics that correct for excess multivariate kurtosis, the power of the chi-square test statistic generally is large when the sample size is large. Structural misspecifications that might otherwise be judged as minor might result in a statistically significant chi-square or chi-square difference test. For this reason, researchers frequently use additional fit statistics to supplement the chi-square test statistic. There are numerous fit statistics available (Bollen and Long, 1993), but here we present several that we use in our example section: the Incremental Fit Index (IFI, Bollen 1989b), 1 minus the Root Mean Square Error of Approximation

($1-RMSEA$, Steiger and Lind, 1980), and the Bayesian Information Criterion (*BIC*, Schwartz, 1978; Raftery, 1995):

$$IFI = \frac{T_b - T_h}{T_b - df_h} \tag{5.54}$$

$$(1-RMSEA) = 1 - \sqrt{\frac{T_h - df_h}{(N-1)df_h}} \tag{5.55}$$

$$BIC = T_h - df(\ln(N)) \tag{5.56}$$

where T_b and T_h are the likelihood ratio test statistics for a baseline and the hypothesized models, df_b and df_h are the df for the baseline and hypothesized models, N is the sample size and t is the number of free parameters in the model. The hypothesized model is simply the model that the researcher is testing and the baseline model is a highly restrictive model to which the fit of the hypothesized model is being compared. Typically the baseline model freely estimates the variances and means of the observed variables but forces their covariances to zero. A value of 1 is an ideal fit for the IFI and $(1-RMSEA)$. For the *BIC*, a negative value is evidence that favors the hypothesized model over the saturated model whereas a positive value favors the saturated model.[4] Although judgement is required in evaluating these fit indices, values less than .90 are typically considered to signify an inadequate fit to the data for the IFI and $(1-RMSEA)$.

5.6 Examples

5.6.1 Data

The data for these examples are repeated measures of Rosenberg's self-esteem scale from the National Longitudinal Study of Youth (NLSY). The data are organized by age of respondent rather than by wave of the survey. Using age to measure time creates missing data so we need to use the direct maximum likelihood estimator to take account of the missing values. There are 5622 respondents between the ages of 15 and 30 put in two year groupings, ages 15-16 to ages 29-30 with each assessed a minimum of once and a maximum of 6 times. The observed mean levels of self-esteem by age group are 3.058 for 15 and 16, 3.090 for 17 and 18, 3.113 for 19 and 20, 3.120 for 21 and 22, 3.125 for 23 and 24, 3.146 for 25 and 26, 3.141 for 27 and 28, and 3.127 for 29 and 30. The average mothers' education level is 11.544 years.

[4] This interpretation holds when calculating *BIC* as in equation (5.56), but this interpretation will not be true if different formulas are used.

5.6.2 Models

We present several models. First, we estimate the unconditional autoregressive (AR) model. Then we estimate the unconditional latent curve model (LCM). Third, we present the results from the unconditional ALT model. Fourth, and finally, we add the respondents' mothers' years of education in 1994 as an exogenous predictor to produce a conditional ALT (cALT) model. All estimation was conducted using Mplus 5.2. The programs that produced the results and the results themselves are available in Chapter 5 at the book website `http://www.econ.upf.edu/~satorra/longitudinallatent/readme.html`. Table 5.1 shows the fit statistics corresponding to the five models we estimated. The first model, AR with equal intercepts, has a statistically significant chi-square, low values of the IFI and a positive value for the BIC which suggests that the saturated model fits better than the hypothesized one. The only fit index that suggests a good fit is the $(1 - RMSEA)$. We also estimated the AR model with unconstrained intercepts. The fit was very close to that of the autoregressive model we report in Table 5.1 ($T_{ML}(18) = 262.39$, $p = 0.006; IFI = 0.82; 1 - RMSEA = 0.95; BIC = 106.96$).

The second model, the Latent Curve Model (LCM), has a fit that is much better than the AR one in that the IFI and $(1 - RMSEA)$ are high and the BIC is a large negative value. Combining features of both models in the ALT model we find for the first time a nonsignificant chi-square, an IFI and $(1 - RMSEA)$ that are near their ideal values, and a large negative BIC. However, closer examination of the parameter estimates and their standard errors reveals that the mean, variance, and the covariances of the slope are all not significantly different from zero. This suggests that the slope factor is not needed in this model. Furthermore, the autoregressive coefficients appear near equal when their standard errors are taken into account. This led us to respecify the ALT model without the slope term and with the autoregressive parameters set equal. The fit statistics suggest that this model fits very well. This model suggests that there are stable individual differences in self-esteem and that there is an impact of past self-esteem feelings on current ones.

Table 5.1 Overall fit of Autoregressive, Latent Curve, and Autoregressive Latent Trajectory models for self-esteem, ages 15-30 (N = 5622)

Overall Fit	(1) Autoregressive Model	(2) Latent Curve Model	(3) Unconditional ALT Model	(4) No Slope Unconditional ALT Model	(5) No Slope Conditional ALT Model
T_{ML}	280.69	69.93	22.76	39.03	49.36
df	24	28	18	28	34
p-value	<0.001	<0.001	0.200	0.080	0.043
IFI	0.81	0.97	1.00	0.99	0.99
$1 - RMSEA$	0.96	0.98	0.99	0.99	0.99
BIC	73.46	-171.83	-132.66	-202.73	-244.21

Which of these four models is best? The question is complicated by the fact that not all of these models are nested. However, some are. If in the unconditional ALT model (see column (3) in Table 5.1) we set the $VAR(\beta_i)$, $VAR(\alpha_i)$, $E(\beta_i)$, $COV(\beta_i, \alpha_i)$, $COV(\beta_i, y_1)$, and $COV(\alpha_i, y_1)$ to zero, then we are led to equation (5.2) which is the AR model with equal intercepts reported in column (1) of Table 5.1. A nested chi-square difference test leads to a highly significant difference ($T_{ML}(6) = 280.69 - 22.76 = 257.93,\ p < 0.001$) lending support to the ALT over the AR model. The "No Slope Unconditional ALT Model" of column (4) is nested in the "Unconditional ALT Model" of column (3) and the chi-square difference test is not significant ($T_{ML}(10) = 39.032 - 22.763 = 16.69$, $p = 0.092$) lending support to the ALT model without a slope. The "Latent Curve Model" of column (2) is not nested in the "Unconditional ALT Model" of column (3) because the ALT model treats y_1 as predetermined while the LCM model treats that variable as endogenous. Despite the nonnesting of some of these models, the other fit statistics are comparable for nonnested models. By all measures the AR model is inadequate. Considering all of the fit statistics, the "No Slope Unconditional ALT Model" appears to have the best fit among models (1) to (4).

Given that the "No Slope Unconditional ALT Model" was the best, we used it to estimate a conditional model that treats mother's education as a covariate. Though the chi-square for this model is marginally significant, the other fit statistics look excellent for this conditional model and we interpret the results of that model in detail. Table 5.2 shows the parameter estimates from the cALT model, which were taken from the Mplus 5.2 output for that model.

The first row of Table 5.2 shows the fixed relationships between the random intercepts (set at 1) and the observed repeated measures of self-esteem. The equal autoregressive effects of the self-esteem measure, the $\widehat{\rho}$ coefficients, are 0.192, showing a positive impact of past on current self-esteem. These effects are net of the random intercept effects. The residual variances ($\widehat{VAR(\varepsilon)}$) of the repeated measures are statistically significant; hence there is age-specific error in the repeated measures. They are similar in size, however, and could be constrained to be equal as another potential simplification to the model – the measurement error in the repeated measures is the same at all ages. The R-squares of all repeated measures but the first are moderate in size ranging from 0.305 to 0.369. This suggests that the random intercepts and the prior self-esteem variables explain roughly 30 to 37% of the variation in each self-esteem measure.

We turn now to the impact of mother's education on the random intercept. This is equivalent to a regression with the random intercept being the dependent variable and mother's education being the explanatory variable. There is a regression constant or fixed intercept ($\widehat{\mu}_\alpha$) and a slope ($\widehat{\gamma}_{\alpha 1}$). The slope ($\widehat{\gamma}_{\alpha 1}$) of mother's education is 0.005 so that each unit shift in education leads to an expected shift of 0.005 in the random intercept variable. The regression constant ($\widehat{\mu}_\alpha$) from this regression equation is 2.461 which is the predicted value of the random intercept when mother's education is zero, though a value of 0 for mother's education does not occur in our data. There is also significant variation of the regression residuals in the random intercepts equation of 0.025 ($= \widehat{VAR(\zeta_\alpha)}$) and an R-square

Table 5.2 ML parameter estimates and z-values in the No Slope Conditional ALT model for self-esteem, ages 15-30 (N = 5622)

Parameter	Model	SE 15-16	SE 17-18	SE 19-20	SE 21-22	SE 23-24	SE 25-26	SE 27-28	SE 29-30
λ_t of α	–	–	1.000 (–)	1.000 (–)	1.000 (–)	1.000 (–)	1.000 (–)	1.000 (–)	1.000 (–)
ρ	–	–	0.192 (7.67)	0.192 (7.67)	0.192 (7.67)	0.192 (7.67)	0.192 (7.67)	0.192 (7.67)	0.192 (7.67)
VAR (ε)	–	0.105 (47.37)	0.070 (25.50)	0.077 (23.51)	0.073 (18.96)	0.072 (16.86)	0.086 (14.37)	0.084 (10.02)	0.097 (6.86)
μ_α	2.461 (32.03)	–	–	–	–	–	–	–	–
VAR(ζ_α)	0.025 (9.66)	–	–	–	–	–	–	–	–
v	–	2.969 (159.76)	– –	– –	– –	– –	– –	– –	– –
COV(α, SE 15-16)	–	0.172 (18.85)	– –	– –	– –	– –	– –	– –	– –
$\gamma_{\alpha 1}$	0.005 (4.63)	– –	– –	– –	– –	– –	– –	– –	– –
$\gamma_{SE15-16,1}$	0.008 (5.065)	– –	– –	– –	– –	– –	– –	– –	– –
R^2	0.012 (α)	0.007	0.366	0.352	0.367	0.369	0.329	0.336	0.305

for this equation of only 0.012. Mother's education is a poor predictor of the random intercept for self-esteem. Turning to the effect of mother's education on initial self-esteem 15 to 16, we again find a low R-square (0.007), but a statistically significant intercept ($\widehat{v} = 2.969$) and slope ($\widehat{\gamma}_{SE15-16,1} = 0.008$). In addition, the random intercepts are significantly correlated with the self-esteem at 15 and 16 ($\widehat{COV}(\alpha, SE15-16) = 0.172$).

So what have we learned about the trajectories of self-esteem from ages 15 and 16 to ages 29 and 30? First, we found an autoregressive process with prior self-esteem having a positive effect on current self-esteem, but this is combined with a random intercept term that provides for a different constant level of self-esteem for each child. In fact, including only an autoregressive term does not lead to a good fitting model. The autoregressive and the random intercept effects explained roughly 30% to 37% of the variation in the self-esteem variables. The random slope was not needed. This implies that once we control for the random intercept and the autoregressive relationship, there is no need to add a linear trend term in self-esteem for each child. There are differences in their levels of self-esteem that tend to be constant, but that are also affected by prior self-esteem. Our conditional model revealed statistically significant positive effects of mother's education on the random intercept and on the initial self-esteem 15 and 16, but the effects were small as was the R-square.

5.7 Conclusions

This paper reviewed the ALT model which synthesizes features of the autoregressive/cross-lagged and the latent growth curve models. It permits the lagged value of a repeated measure to influence the current value while at the same time permits there to be separate over-time trajectories for individuals in the sample. As such it provides a researcher added flexibility in capturing the nature of change exhibited in panel data. Furthermore, the ALT model yields evidence relevant to whether the synthesis is required or if a researcher can get by with only the autoregressive and cross-lagged model or only the latent curve model. Obvious generalizations of the ALT model include multiple repeated measures, autoregressive models beyond lag one (e.g., AR(p) models), nonlinear trajectories, or ALT models for latent variables with separate measurement models with multiple indicators. The ALT model already includes latent variables in that the random intercept and random slope variables are latent. However, in the case of a multiple indicator model for the repeated "measure," the ALT model would allow a model of the autoregressive relation and the trajectory of the latent variables that would control for the measurement error in the indicators of the latent variables. This also would provide an estimate of the amount of measurement error in the multiple indicators. In the conditional ALT model it also would be possible to include latent exogenous variables as predictors of the random intercepts, random slopes, and the initial value of the latent repeated variable.

Despite these desirable features, several cautionary notes are in order. First, the ALT model assumes that the repeated measure has a direct impact on itself at a later point in time. A researcher should have substantive reasons to believe that this is a reasonable hypothesis and should not use the ALT model as just a way to improve model fit. A second related point is that it is possible that the autoregressive relation resides in the disturbance rather than in the repeated measures. In this situation, the disturbances should be autoregressive rather than the repeated measures since this implies a model that generally differs from the ALT.[5] Third, our presentation assumes that the researcher has the correct functional form for the latent curve trajectory in the ALT model. If, for example, we assume a linear functional form when a trajectory is nonlinear, then the autoregressive part of the ALT model might be due to the researcher using the wrong functional form (Voelkle, 2008).[6] A related point

[5] Hamaker (2005) discusses the special cases where the ALT and autoregressive disturbance model can be made statistically equivalent.

[6] We explored nonlinearity in our empirical example by using the "freed loading" model (Bollen and Curran, 2006, pp. 98-103). There was no improvement to model fit and the autoregressive parameters were still significant suggesting that the linear functional form was an appropriate starting point.

is that extrapolating trends beyond the period of observation should only be done with great caution. A linear trend might be a good approximation of a trajectory within the time period of observation, but extrapolating too far out could lead to highly inaccurate predictions if the relation is really nonlinear. Finally, throughout our presentation we assume discrete time models are good approximations to continuous time models. Many processes occur in continuous time even when the data are available only at fixed times. If the waves of data collection are too spread out relative to the timing of the relationships, then our discrete time models could be misleading. For instance, the autoregressive or ALT model might lead to inaccurate estimates of relationships if the observation interval for the discrete time model is long. Delsing and Oud (2008) present an extension of the ALT model to continuous time modeling that enables researchers to use variables observed in panel data but allow continuous rather than discrete time.

Keeping these limitations in mind, we believe that the ALT model provides a useful extension to some of the more commonly used models for panel data.

Acknowledgements Kenneth Bollen gratefully acknowledges the support from NSF SES 0617276 from NIDA 1-RO1-DA13148-01 & DA013148-05A2. We thank the editors and reviewers for valuable comments and thank Shawn Bauldry for research assistance.

References

Anderson, T. W. (1960). Some stochastic process models for intelligence test scores. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), *Mathematical methods in the social sciences.* Stanford, CA: Stanford University Press.

Bast, J. & Reitsma, P. (1997). Matthew effects in reading: A comparison of latent growth curve models and simplex models with structured means. *Multivariate Behavioral Research, 32,* 135-167.

Bohrnstedt, G. W. (1969). Observations on the measurement of change. *Sociological methodology, 1,* 113-133.

Bollen, K. A. (1989a). *Structural equation models with latent variables.* New York: Wiley.

Bollen, K. A. (1989b). A new incremental fit index for general structural equation models. *Sociological Methods & Research, 17,* 303-316.

Bollen, K. A. (2007). On the origins of latent curve models. In R. Cudeck & R. MacCallum (Eds.), *Factor analysis at 100* (pp. 79-98). Mahwah, NJ: Lawrence Erlbaum Associates.

Bollen, K. A. & Curran P. J. (2004). Autoregressive Latent Trajectory (ALT) models: A Synthesis of two traditions. *Sociological Methods & Research, 32,* 336-383.

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

Bollen, K. A., & Long, J. Scott. (1993). *Testing structural equation models.* Newbury Park, CA: Sage Publications.

Bollen, K. A., & Stine, R. A. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological Methodology, 20,* 115-140.

Bollen, K. A., & Stine, R. A. (1993). Bootstrapping goodness-of-fit measures in structural equation models. In K. A. Bollen & J. Scott Long (Eds.), *Testing structural equation models* (pp. 111-135). Newbury Park, CA: Sage Publications.

Box, G. E. P., & Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control.* San Francisco: Holden-Day.

Boyd, J. L. (2007). *Developmental and situational factors contributing to changes in eating behavior in first-year undergraduate women.* Master of Arts in Psychology. University of Waterloo, Canada.

Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical & Statistical Psychology, 37,* 62-83.

Browne, M. W., & du Toit, S. H. C. (1991). Models for learning data. In L. Collins & J. Horn (Eds.), *Best methods for the analysis of change: Recent advances, unanswered questions, future directions* (pp. 47-68). Washington, DC: APA.

Campbell, D. T. (1963). From description to experimentation: Interpreting trends as quasi-experiments. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 212-242). Madison, WI: University of Wisconsin Press.

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,* 452-59.

Curran, P. J., & Bollen, K. A. (2001). The bests of both worlds: Combining autoregressive and latent curve models. In Collins L. M. & Sayar, A.G. (Eds.), *New methods for the analysis of change* (pp. 105-136). Washington, D.C.: American Psychological Association.

Curran, P. J., & Willoughby, M. T. (2003). Implications of latent trajectory models for the study of developmental psychopathology. *Development and Psychopathology, 15,* 581-612.

Delsing, M. J. M. H., & Oud, J. H. L. (2008). Analyzing reciprocal relationships by means of the continuous-time autoregressive latent trajectory model. *Statistica Neerlandica, 62,* 58-82.

Diggle, P. J., Liang, K. Y., & Zeger, S. L. (1994). *Analysis of longitudinal data.* Oxford: Clarendon Press.

Duncan, O. D. (1969). Some linear models for two-wave, two-variable panel analysis. *Psychological Bulletin, 72,* 177-182.

Goldstein, H., Healy, M. J. R., & Rasbash, J. (1994). Multilevel time series models with applications to repeated measures data. *Statistics in Medicine, 13,* 1643-1655.

Hamaker, E. (2005). Conditions for the equivalence of the autoregressive latent trajectory model and a latent growth curve model with autoregressive disturbances. *Sociological Methods & Research, 33,* 404-416.

Heise, D. R. (1969). Separating reliability and stability in test-retest correlation. *American Sociological Review, 34,* 93-101.

Humphreys, L. G. (1960). Investigations of the simplex. *Psychometrika, 25,* 313-323.

Jöreskog, K. G. (1970). A general method for analysis of covariance structures. *Biometrika, 57,* 239-251.

Jöreskog, K. G. (1979). Statistical models and methods for analysis of longitudinal data. In K. G. Jöreskog & D. Sörbom (Eds.), *Advances in factor analysis and structural equation models.* Cambridge, Mass: Abt.

Kenny, D. A., & Campbell, D. T. (1989). On the measurement of stability in over-time data. *Journal of Personality, 57,* 445-481.

Kessler, R. C., & Greenberg, D. F. (1981). *Linear Panel Analysis.* New York: Academic Press.

Lawley, D. N., & Maxwell, A. E. (1971). *Factor analysis as a statistical method.* London: Butterworth.

Meredith, W., & Tisak, J. (1984). *Tuckerizing curves.* Paper presented at the annual meeting of the Psychometric Society, Santa Barbara, CA.

Muthén, L. K. & Muthén, B. O. (1998-2007). *Mplus User's Guide* (5th ed.). Los Angeles: Muthén & Muthén.

Rao, C. R. (1958). Some statistical methods for comparison of growth curves. *Biometrika, 51,* 83-90.

Raftery, A. E. (1995). Bayesian model selection in social research (with discussion). *Sociological Methodology, 25,* 111-163.

Rodebaugh, T. L., Curran, P. J., & Chambless, D. L. (2002). Expectancy of panic in the maintenance of daily anxiety in panic disorder with agoraphobia: A longitudinal test of competing models. *Behavior Therapy, 33,* 315-336.

Rogosa, D., & Willett, J. B. (1985). Satisfying simplex structure is simpler than it should be. *Journal of Educational Statistics, 10,* 99-107.

Rogosa, D. R., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin, 92,* 726-748.

Satorra, A. (1990). Robustness issues in structural equation modeling: A review of recent developments. *Quality & Quantity, 24,* 367-386.

Satorra, A., & Bentler, P. M. (1988). Scaling corrections for chi-square statistics in covariance structure analysis. *Proceedings of the American Statistical Association,* 308-313.

Schwarz, G. (1978). Estimating the dimensions of a model. *Annals of Statistics, 6,* 461-464.

Simons-Morton, B., & Chen, R. S. (2006). Over time relationships between early adolescent and peer substance abuse. *Addictive Behaviors, 31,* 1211-1223.

Steiger, J. H., & Lind, J. M. (1980). *Statistically based tests for the number of common factors.* Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.

Tucker, L. R. (1958). Determination of parameters of a functional relation by factor analysis. *Psychometrika, 23,* 19-23.

Voelkle, M. C. (2008). Reconsidering the use of Autoregressive Latent Trajectory (ALT) models. *Multivariate Behavioral Research, 43,* 564-591.

Wan, T. T. H., Zhang, N. J., & Unruh, L. (2006). Predictors of resident outcome improvement in nursing homes. *Western Journal of Nursing Research, 28,* 974-993.

Werts, C. E., Jöreskog, K. G., & Linn, R. L. (1971). Comment on the estimation of measurement error in panel data. *American Sociological Review, 36,* 110-112.

Wiley, D. E., & Wiley, J. A. (1970). The estimation of measurement error in panel data. *American Sociological Review, 35,* 112-117.

Zyphur, M. J., Chaturvedi, S., & Arvey, R. D. (2008). Job performance over time is a function of latent trajectories and previous performance. *Journal of Applied Psychology, 93,* 217-224.