STRUCTURAL EQUATION MODELING, *11*(1), 132–149

TEACHER'S CORNER

A SAS Macro for Estimating and Visualizing Individual Growth Curves

Madeline M. Carrig, R. J. Wirth, and Patrick J. Curran
University of North Carolina at Chapel Hill

Longitudinal data analyses can be usefully supplemented by the plotting of individual growth curves. Unfortunately, such graphics can be challenging and tedious to produce. This article presents and demonstrates a SAS macro designed to automate this task. The *OLStraj* macro graphically depicts ordinary least squares (OLS)-estimated individual trajectories, describes interindividual variability in OLS-estimated growth parameters, and identifies possible outlier observations. Analytical developments are briefly outlined, and the use of the macro is demonstrated, with particular attention paid to the potential utility of the macro as both a data screening and post hoc diagnostic device. Potential limitations of the macro and suggestions for future developments are discussed. It is hoped that the program will be of use to applied researchers who seek to maximize the effectiveness of growth curve models in answering questions about stability and change.

Structural equations-based latent growth curve, or latent trajectory, modeling (LTM) is a powerful and flexible analytic technique for the description and prediction of individual differences in stability and change over time. In LTM, the situation is similar to that encountered in any latent variable confirmatory factor analysis (CFA) or structural equation model (SEM); namely, although we believe individual-specific scores on the latent growth factors exist, it is the latent factor means, variances, and covariances that are estimated from the sample data and not the individual trajectories themselves. Indeed, when there is no missing data, the full LTM can be estimated entirely based on the sample mean vector and covariance matrix with no individual-specific information whatsoever. Frequently,

Requests for reprints should be sent to Patrick Curran, Department of Psychology, University of North Carolina, Chapel Hill, NC 27599–3270. E-mail: curran@unc.edu

the group-level estimates of factor means and variances resulting from the typical LTM analysis are sufficient to answer many substantive questions of interest. However, there are a number of situations in which individual-specific scores on latent growth factors (and hence estimates of individual growth trajectories) may be of use to the applied researcher.

For example, estimates of individual-specific scores allow for the examination of individual trajectories prior to model estimation, a procedure Rogosa (1994) and others argued should comprise part of any longitudinal data analysis. It is clear that computation of maximally informative LTM parameter estimates requires that a number of assumptions hold, including (a) selection of a valid functional form of growth to represent change in constructs of interest, (b) population homogeneity, (c) a linear relation between growth factors, and (d) absence of influential outliers. Estimates of individual growth trajectories can allow for the evaluation of whether the hypothesized functional form uniformly represents change for sampled individuals, and permit outlier diagnostics to be conducted. Still other uses for individual growth trajectory estimates are graphical visualization (e.g., Smith, Best, Stubbs, Archibald, & Roberson-Nay, 2002; Wainer & Velleman, 2000), post hoc exploration of model misfit, prediction of future observations, use of estimates as criterion variables in other analyses, and derivation of cut scores (e.g., as when selecting a subgroup of sampled individuals with a clinically significant cutoff score in a longitudinal study of depression; Biesanz, Curran, & Bollen, 2003). Taken together, these potential uses suggest that a procedure for estimating individual growth trajectories could represent a highly useful adjunct to LTM.

Unfortunately, because of the time-consuming and complicated nature of the programming work involved, it can be challenging for applied users of LTM technology to compute and graphically display estimates of individual growth trajectories. To capitalize on the estimation of individual trajectories in our own work, we have developed a SAS-based macro that estimates and plots ordinary least squares (OLS)-estimated individual growth trajectories for continuous repeated measures data across a variety of experimental settings. Given the time invested in developing this program, combined with the power of OLS trajectory estimation techniques when used in our own applied research, we thought it would be of some use to share this program publicly.

The goal of this article is to present the analytical developments that underlie OLS trajectory estimation, to present the macro itself, and to demonstrate the use of this macro on an empirical data set of antisocial behavior assessed on $N = 221$ children on four separate occasions. Although we focus on SEM-based approaches to trajectory analysis here, the macro may be used to supplement any growth curve modeling procedure. The macro and supporting documentation are available on a publicly accessible Web page (www.unc.edu/~curran/OLStraj.htm), and the user can replicate all analyses there. It is important to note that only minor modifications to the program are required to allow for the application of this macro across a wide variety of conditions that might be encountered in applied research. We conclude with a discussion of the known limitations of OLS trajectory estimation, potential limitations of the macro, and suggestions for future developments in this area.

THE UNCONDITIONAL RANDOM GROWTH CURVE MODEL

For the unconditional LTM that assumes linear growth, the general model is of the following form:

$$y_{it} = \alpha_i + \beta_i \lambda_t + \varepsilon_{it} \tag{1}$$

where y_{it} represents the value of the repeated measure for individual i at time t; α_i and β_i are the random intercept and slope, respectively, of the (linear) growth trajectory for individual i; λ_t represents the value of the time trend variable at time t; and ε_{it} represents the individual- and time-specific residual. If a quadratic functional form of growth is preferred, the general model assumes the following form:

$$y_{it} = \alpha_i + \beta_i \lambda_t + \beta_{2i} \lambda_t^2 + \varepsilon_{it} \tag{2}$$

where β_i represents the random slope, β_{2i} is the quadratic parameter describing the acceleration (or deceleration) of change over time for individual i, and all other parameters are as previously defined.

Because the intercepts and slopes in Equations 1 and 2 are random, they can be expressed as functions of a mean and deviation in the Level 2 equations:

$$\alpha_i = \mu_\alpha + \zeta_{\alpha i} \tag{3}$$

$$\beta_i = \mu_\beta + \zeta_{\beta i} \tag{4}$$

such that the intercept of the (linear) latent trajectory for individual i is held to be a function of the mean of the intercepts across all cases (μ_α) and an individual-specific intercept disturbance deviation ($\zeta_{\alpha i}$), and the slope of the (linear) latent trajectory for individual i is held to be a function of the mean of the slopes across all cases (μ_β) and an individual-specific slope disturbance deviation ($\zeta_{\beta i}$). The combined, or reduced form, equation for the unconditional linear growth model is

$$y_{it} = [\mu_\alpha + \lambda_t \mu_\beta] + [\zeta_{\alpha i} + \lambda_t \zeta_{\beta i} + \varepsilon_{it}] \tag{5}$$

From this equation, it is readily apparent that the values of the repeated measures y_{it} are hypothesized to be a function of latent growth factors and latent individual-level disturbance factors. In most applications the λ_t are fixed, but it is possible to estimate them from the sample data (McArdle & Epstein, 1987; Meredith & Tisak, 1984, 1990).

Therefore, in the standard unconditional linear growth curve model, the estimated parameters are the means of the latent trajectory factors $(\hat{\mu}_\alpha, \hat{\mu}_\beta)$, the variances of the latent trajectory factors (denoted $\hat{\psi}_\alpha$ and $\hat{\psi}_\beta$), the covariance between the latent trajectory factors (denoted $\hat{\psi}_{\alpha\beta}$), and the residual variance (denoted $\hat{\sigma}^2$). This fixed- and random-effects growth curve model may be estimated using a multilevel modeling procedure such as hierarchical linear modeling (HLM; Bryk & Raudenbush, 1987) or using a structural equations-based latent growth curve modeling approach (LTM; McArdle, 1988; Meredith & Tisak, 1984, 1990). Although we focus on the LTM approach here, the *OLStraj* macro may be used to supplement any growth curve modeling approach.

LTM ESTIMATION OF THE RANDOM GROWTH CURVE MODEL

Using the traditional SEM matrix terminology, the general expression for the LTM is

$$\mathbf{y} = \Lambda\eta + \varepsilon \tag{6}$$

where $\mathbf{y}$ is a $T \times 1$ vector of repeated measures, Λ is a $T \times k$ matrix of factor loadings (i.e., time trend variable values), η is a $k \times 1$ vector of latent factors, and ε is a $T \times 1$ vector of residuals. Equation 6 is analogous to the Level 1 model presented in Equation 1. In a fashion analogous to the Level 2 growth curve equation, η can be expressed in terms of a mean and deviation as follows:

$$\eta = \mu_\eta + \zeta \tag{7}$$

where μ_η is a $k \times 1$ vector of growth factor means and ζ is a $k \times 1$ vector of residuals. Equation 7 can be substituted into Equation 6 to result in the reduced-form matrix equation for the LTM:

$$\mathbf{y} = \Lambda\,(\mu_\eta + \zeta) + \varepsilon \tag{8}$$

This model is presented in Figure 1.

The model-implied variances of the observed repeated measures are

$$\mathrm{VAR}(\mathbf{y}) = \Lambda\Psi\Lambda' + \Theta_\varepsilon \tag{9}$$

where Θ_ε represents the covariance structure of the residuals for the T-repeated measures of $\mathbf{y}$ and Ψ represents the covariance matrix of the deviations ζ. Finally, the model-implied mean structure of the observed repeated measures is represented in the following matrix expression:

$$\mathbf{E}(\mathbf{y}) = \Lambda\mu_\eta \tag{10}$$

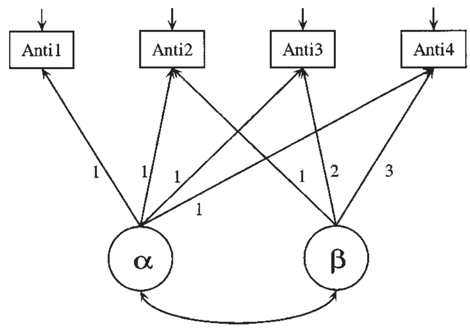


FIGURE 1 A path diagram of the four timepoint linear trajectory model of antisocial behavior (Anti1–Anti4).

where Λ and μ_η are defined as before.

As in any SEM model, the optimal parameter estimates in LTM are typically viewed as those that minimize the difference between the observed and the model-implied mean vectors and covariance matrices. Accordingly, in most standard applications that do not consider missing data, the sample mean and covariance matrices, and not individual-level data, are utilized in model estimation. Perhaps the most widely used approach to computation of parameter estimates is minimization of the maximum likelihood fitting function. However, other estimators (e.g., arbitrary distribution function, two-stage least squares) are also available.

Therefore, in latent trajectory analysis, a set of observed repeated measures is used to estimate an unobserved, or latent, growth trajectory that is hypothesized to have given rise to the observed data. The parameters estimated $(\hat{\Psi}, \hat{\mu}_\eta, \text{and}\, \hat{\Theta}_\varepsilon)$ can be computed using the sample mean vector and covariance matrix alone, and represent *summary* statistics for latent scores on growth factors and disturbances. Just as factor scores are not routinely estimated in the standard CFA or SEM, individual-level values of latent growth factors are neither directly assessed, nor estimated, in the typical LTM. Fortunately, it is possible to estimate individual-level scores on latent growth factors. It is to OLS estimation of individual growth trajectories, in particular, that we now turn.

OLS ESTIMATION OF INDIVIDUAL TRAJECTORIES

There are multiple approaches to the estimation of latent growth factor scores, including generalized least squares, factor regression (or empirical Bayes), and constrained covariance. Biesanz et al. (2003) concluded from their examination of the analytic properties of factor score estimators that no one estimator is ideal for all potential uses. Because our priorities in developing the *OLStraj* macro were trajectory visualization and accessibility to the user, we selected the OLS estimator (which is equivalent to the generalized least squares estimator under certain restrictive conditions) for use in the program.

OLS estimation uses the OLS fitting function to estimate a regression model for each case in the sample. In the case of a linear functional form of growth, the model estimated for each case is of the following form:

$$y_{it} = \alpha_i + \beta_i \lambda_t + \varepsilon_{it} \tag{11}$$

where y_{it} is the value of the repeated measure for individual i at time t, α_i is the regression intercept for individual i, λ_t is the value of the user specified coding of time at time t, β_i describes linear change over time in y for individual i, and ε_{it} is the time-specific regression residual for individual i. The OLS estimator of slope for each case reduces to

$$\hat{\beta}_i = \frac{\sum_{t=1}^{T}(\lambda_t - \bar{\lambda})(y_{it} - \bar{y}_i)}{\sum_{t=1}^{T}(\lambda_t - \bar{\lambda})^2} \tag{12}$$

and the OLS estimator of intercept for each case reduces to

$$\hat{\alpha}_i = \bar{y}_{it} - \hat{\beta}_i \bar{\lambda} \tag{13}$$

where $\bar{\lambda}$ is the mean of the time trend variable and $\bar{y}_{it}$ is the mean of y for individual i across the T time points.

The OLS estimator possesses several advantages over other point estimation techniques; most notably, case-by-case OLS estimates are readily computed in many widely available statistical packages and possess intuitive appeal for researchers trained in regression. In addition, the mean of the case-by-case OLS intercepts

$$\bar{\hat{\alpha}} = \frac{\sum_{i=1}^{I}\hat{\alpha}_i}{N} \tag{14}$$

and the mean of the case-by-case OLS slopes

$$\bar{\hat{\beta}} = \frac{\sum_{i=1}^{I}\hat{\beta}_i}{N} \tag{15}$$

serve as unbiased estimators of the mean population intercept (μ_α) and mean population slope (μ_β), respectively (Rogosa, Brandt, & Zimowski, 1982). Although the sample variances and covariance of $\hat{\alpha}_i$ and $\hat{\beta}_i$ are biased, OLS estimate-based formulae for corrected (unbiased) variance–covariance estimates are available (e.g., see Rogosa & Saner, 1995).

We now turn to a brief description of the trajectory estimation functions and features of the *OLStraj* macro.

THE *OLStraj* MACRO

The individual trajectories software program and program documentation are available at www.unc.edu/~curran/OLStraj.htm. The program was written using the SAS macro language and incorporates both procedures included in SAS/Base and SAS/STAT software and high-resolution graphics provided by SAS/GRAPH software. The main module of the program is the macro *OLStraj,* which is embedded and invoked within the program. The user may alternatively choose to store the *OLStraj* macro in a user defined macro library.

As described in the program documentation, invocation of *OLStraj* presumes that a SAS data set INITIAL has previously been created by the user. Prior to using *OLStraj,* the user must also (a) create an identification (ID) variable on data set INITIAL, (b) identify the names of the variables on data set INITIAL that represent the repeated measures of interest, and (c) specify the coding of time for each of the repeated measures.

On invocation, the macro *OLStraj* performs the following procedures:

1. *Listwise deletion.* If specified by the user, *OLStraj* will delete any observation in INITIAL having a missing value for one or more of the repeated measures variables. Alternatively, OLS estimates can be computed using all available information for each case (i.e., partially missing data).
2. *Creation of subsample. OLStraj* allows the user to specify that a subset of observations be evaluated. (This feature is particularly useful for larger data sets, when a closer examination of a subsample of trajectories is desired.)
3. *OLS case-by-case regressions. OLStraj* next uses the OLS fitting function to estimate a regression model for each case in the sample (or subsample). The user specifies whether a linear, a quadratic, or both linear and quadratic case-by-case regressions should be performed.
4. *Creation of group-level plots.* If specified by the user, *OLStraj* produces two group-level plots using high-resolution graphics. The first plot displays overlaid simple-joined (noninterpolated) trajectories for all cases in the sample (or subsample). The second plot displays overlaid OLS-estimated trajectories for all cases in the sample (or subsample). If the user specifies a quadratic functional form of growth for OLS case-by-case regressions, the OLS-estimated trajectories displayed on this plot will be quadratic in functional form. Otherwise,

the OLS-estimated trajectories displayed on the plot will be linear in functional form.

5. *Creation of individual-level plots.* If desired by the user, *OLStraj* produces a set of multiple individual-level plots using high-resolution graphics. These individual graphs plot the value of each individual's repeated measures against time (as coded by the user). Linear, quadratic, or both OLS-estimated trajectories are superimposed on the observed repeated measures.

6. *Creation of histograms and box plots.* If specified by the user, *OLStraj* produces a set of three group-level histograms and one multivariable group-level box-and-whisker plot using high-resolution graphics.

The group-level histograms display frequency counts for ranges of values of the estimated intercepts, slopes, and (if quadratic regressions are specified by the user) quadratic terms resulting from OLS case-by-case regressions. *OLStraj* automatically computes 18 data ranges of equal width, which extend from the minimum to the maximum value of the parameter of interest; some of these data ranges may be empty, and this will be reflected on the histograms produced. The box-and-whisker plot provides a visual display of descriptive summary statistics for estimated intercepts, slopes, and (if quadratic regressions are specified) quadratic terms. The type of box-and-whisker plot displayed by *OLStraj* corresponds to the schematic box-and-whisker plot detailed in Tukey (1977). Box plots for estimated parameters are presented side-by-side; each plot displays the mean, quartiles, and minimum and maximum observations for an individual-level estimated parameter. Observations located outside lower and upper "fences" are labeled with the value of the ID variable for that case.

7. *Creation of output data set.* If desired by the user, *OLStraj* creates an output data set that contains parameter estimates resulting from the OLS case-by-case regressions. This data set can be submitted to other SAS procedures or exported for use in other statistical or graphics programs.

EMPIRICAL EXAMPLE

A subset of 221 individuals was selected from the National Longitudinal Survey of Labor Marketing Experience in Youth (NLSY), a study initiated in 1979 by the U.S. Department of Labor. The repeated measure used in this example, the antisocial behavior subtest of the Behavior Problems Index (BPI), was drawn from a larger battery of instruments administered to NLSY mothers and children. The antisocial behavior subtest of the BPI is one of six subtests developed by Zill and Peterson (Baker, Keck, Mott, & Quinlan, 1993). Six 3-point Likert-type items make up the subscale, ranging from 0 (*not true*) to 1 (*sometimes true*) to 2 (*often true*). For this example, responses were summed to compute an overall measure of antisocial behavior; scores could range from 0 to 12. The overall measure of antisocial behavior was computed at each of the four time periods (1986, 1988, 1990, and 1992; denoted as Time 1, Time 2, Time 3, and Time 4, respectively). The resulting variables constructed for Time 1 through Time 4 were named Anti1, Anti2, Anti3, and Anti4, respectively. The individuals selected for this example had complete data for all repeated measures. See Curran and Bollen (2001) for further details.

An Unconditional LTM

To examine the fixed and random components of growth in antisocial behavior, we used PROC CALIS (SAS Institute, 1990) to estimate an unconditional linear LTM for the repeated measures of antisocial behavior collected at Times 1, 2, 3, and 4. Two latent factors were estimated: The first factor defined the intercept of the developmental trajectory of antisocial behavior (with all factor loadings fixed to 1.0) and the second defined the linear slope of the trajectory (with factor loadings set to 0, 1, 2, and 3 to define an annual metric of time). The model estimated is presented in Figure 1. Means were estimated for both the intercept and slope factors, with these values representing the model-implied mean developmental trajectory pooled over all individuals. Variances were also estimated for the intercept and slope factors, representing the degree of individual variability in trajectories around the group mean values. The covariance between the two factors represented the covariation between initial level and rate of change. Finally, residual variances were estimated for each repeated measure and represent the variability in the time-specific measures not accounted for by the underlying random trajectories. It was assumed that measurement error remained constant over time and that residual variances were uncorrelated.

The initial model fit the data well according to the chi-square test of omnibus model fit: $\chi^2(8, N = 221) = 5.56$, $p = .70$. Numerous fit indexes also suggested exemplary model fit to the data: comparative fit index = 1.0, Nonnormed Fit Index = 1.0, incremental fit index = 1.0, and root mean square error of approximation = 0. The estimated means of the latent factors suggested that the model-implied mean trajectory for the sample was characterized by a significant mean BPI score of 1.55 at the first time period ($p < .001$), and a significantly increasing slope of .18 units per year during the study window ($p < .001$). Further, significant variance estimates for both the intercept ($\hat{\psi} = .97$, $p < .001$) and slope ($\hat{\psi} = .10$, $p = .03$) factors indicated substantial interindividual variability in both initial level and rate of change in antisocial behavior. Finally, the estimated correlation between the intercept and slope factors ($r = .49$, $p = .04$) suggested a significant and positive association between antisocial behavior at Time 1 and the rate of change in antisocial behavior over time. Again, note that the $\hat{\mu}_\alpha, \hat{\mu}_\beta, \hat{\psi}_\alpha, \hat{\psi}_\beta$, and $\hat{\psi}_{\alpha\beta}$ resulting from the unconditional LTM represent summary statistics for the latent growth factors and were estimated without direct calculation of the individual trajectories for each case in the sample.

Next, we demonstrate how the *OLStraj* macro may be used to supplement the findings from the LTM of antisocial behavior.

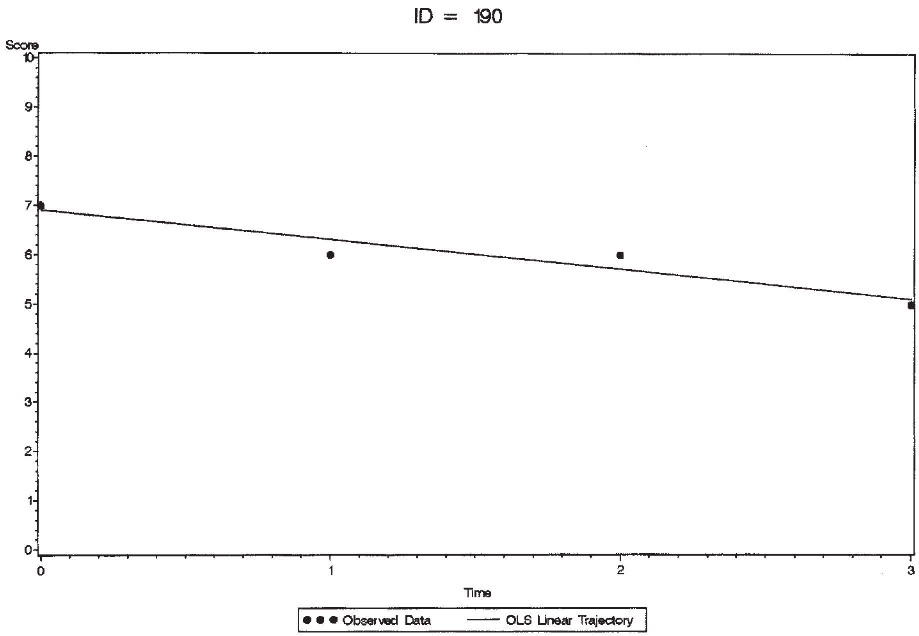
Application of the *OLStraj* Macro

Selected output from the *OLStraj* macro is displayed in Figures 2 through 8. Our initial step was to use *OLStraj* to estimate and plot individual trajectories for each observation in the sample. We first requested plots that superimpose estimated linear trajectories on the observed repeated measures for an evaluation of individual fit to the data. These plots reveal that for many children in the sample, a linear functional form appears to represent change in antisocial behavior quite well (see Figure 2 for representative plots). This finding is consistent with the excellent fit of the unconditional linear LTM to the antisocial behavior data. Next, we requested that *OLStraj* add estimated quadratic trajectories to the individual trajectory plots. As shown in Figure 3, which presents a subset of the relevant output, it appears that for other children a quadratic functional form of growth better accounts for change in antisocial behavior than a linear function. This finding suggests that if theoretically reasonable, it may be beneficial to fit a quadratic LTM to the observed data and evaluate the significance of improvement in model fit.

The *OLStraj* macro was then employed to visually examine all 221 OLS-estimated individual trajectories simultaneously. The graphs provided by the macro are (a) a plot of overlaid simple trajectories and (b) a plot of overlaid OLS-estimated trajectories. As can be seen in the overlaid OLS-estimated trajectories plot (included as Figure 4), consistent with findings from the unconditional linear LTM, substantial variability in both initial level and rate of change in antisocial behavior appears to exist within the sample. From examination of the plot, it is clear that a sizable number of negative linear slopes, positive linear slopes, and possibly a number of linear slopes very close to zero exist in the sample. Potentially outlying or otherwise problematic observations (e.g., observations with negative OLS-estimated intercepts, which are theoretically impossible) can clearly be discerned from the plot and may be identified for further evaluation.

From Figure 4, it is apparent that when a large number of individuals is incorporated into a single plot, certain individual trajectories become difficult to discern. Accordingly, we next used the macro to visually examine OLS-estimated individual trajectories for a small subset ($n = 25$) of the total sample ($N = 221$; Figure 5). Figure 5 reveals more clearly that a positive rate of change in antisocial behavior over time does not hold for all individuals. In fact, a number of the individuals ($n = 6$ of 25) in the subset actually appear to experience decreasing trajectories of antisocial behavior (despite the positive mean trajectory for the entire group).

Distributional characteristics were further explored using the *OLStraj* macro's histogram and box-and-whisker plot option. Examination of a histogram analyzing all $N = 221$ individual slopes (Figure 6) indicates that the distribution of



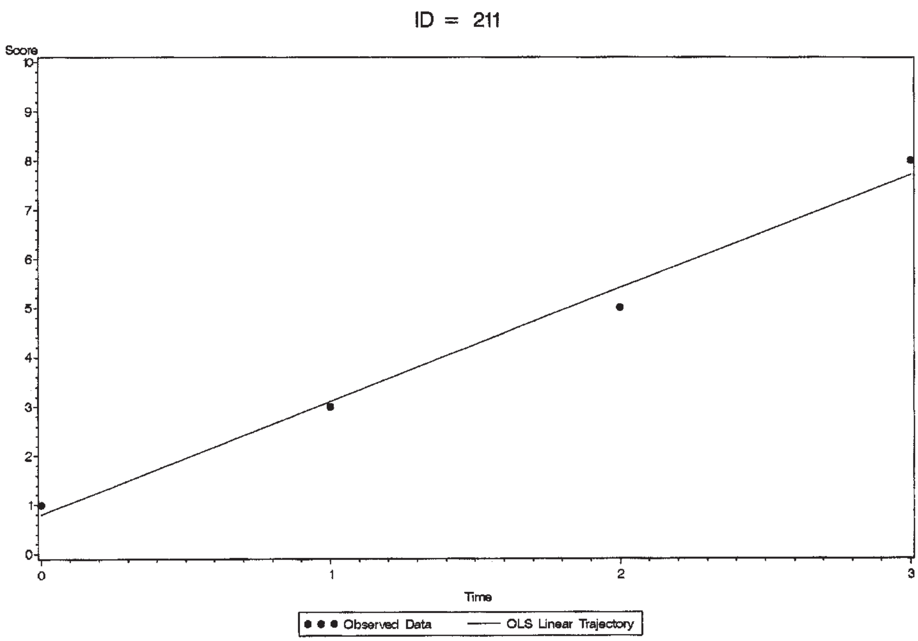


FIGURE 2 Two children for whom a linear functional form appears to represent growth well. ID = identification variable; OLS = ordinary least squares.

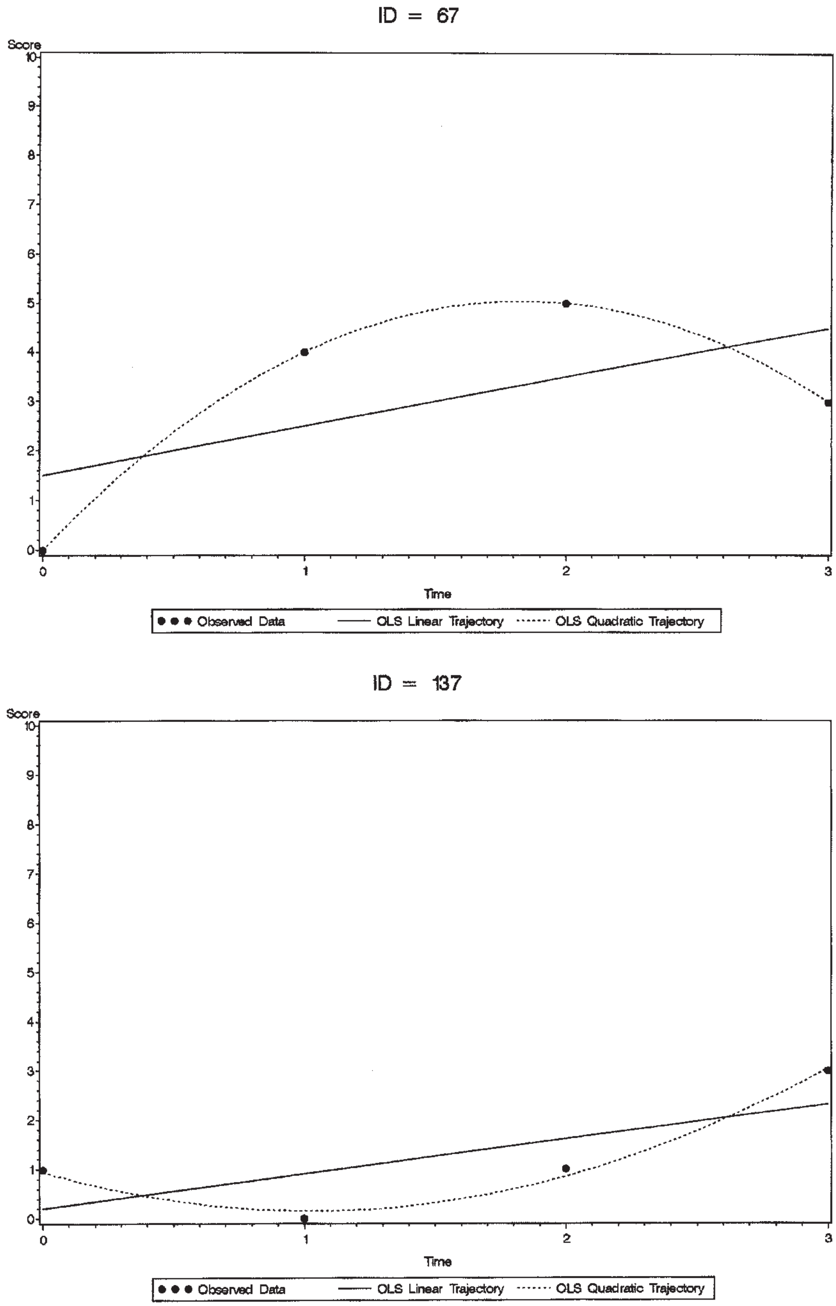


FIGURE 3 Two children for whom a quadratic functional form of growth appears to better fit the data than a linear functional form. ID = identification variable; OLS = ordinary least squares.

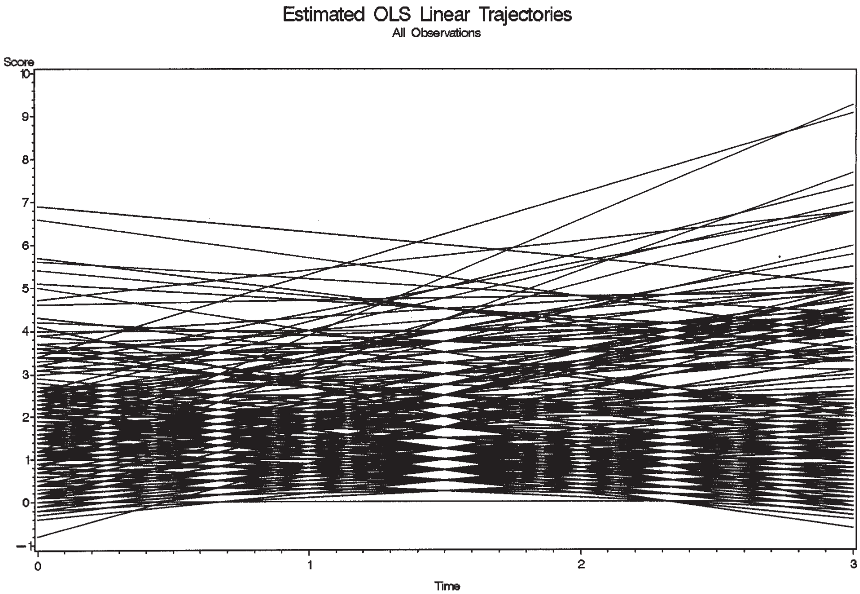


FIGURE 4 Overlaid ordinary least squares (OLS)-estimated trajectories plot (N = 221).

OLS-estimated latent slopes is centered close to zero (M = .18), has a slight positive skew, and represents a relatively poor approximation to the normal distribution. As shown in Figure 7, the box-and-whisker plot demonstrates that extreme values for estimated growth parameters exist, suggesting the possible presence of influential outliers. One of the extreme slopes identified and labeled on the box-and-whisker plot was estimated by *OLStraj* for ID number 69. As shown in Figure 8, the pattern of observed and OLS-estimated antisocial behavior scores for ID 69 is indeed unusual. This observation warrants further attention (e.g., this observation might be checked for keypunch errors, or if proven to exert an undue influence on model results, considered for removal from the sample on this basis). Because evaluation of outliers is beyond the scope of this article, we do not address this further.

DISCUSSION

As is hopefully apparent from the brief example presented here, we believe that *OLStraj* possesses many potential uses as a diagnostic device. For example, *OLStraj* can be used to evaluate model assumptions. The visually presented OLS

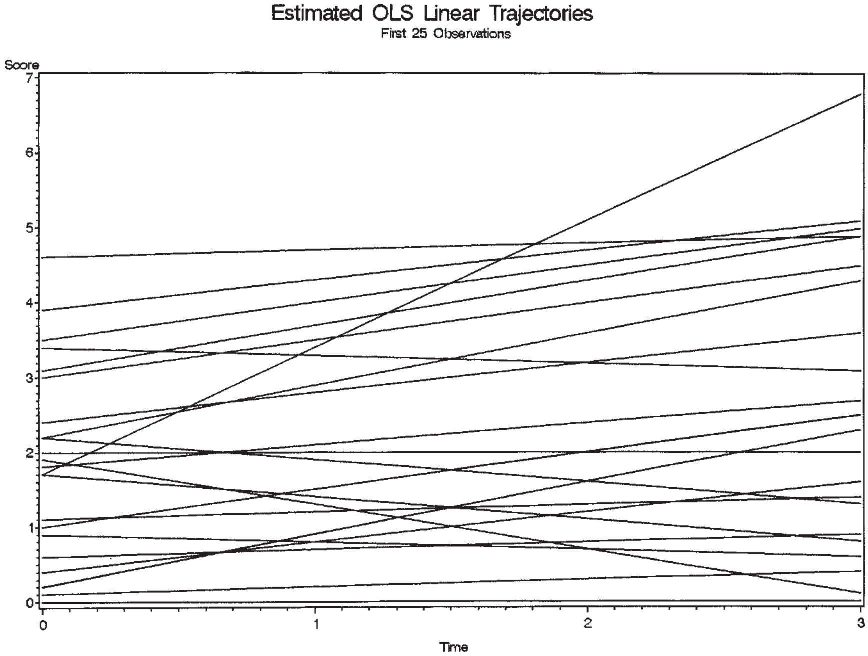


FIGURE 5 Overlaid ordinary least squares (OLS)-estimated trajectories plot (n = 25).

estimates of individual growth trajectories computed by *OLStraj* can allow for the identification of cases for which a simple linear functional form of growth does not characterize the observed data well; further examination of these cases could reveal that these cases are outliers, or perhaps alternatively, that a quadratic or otherwise nonlinear model of growth might be more appropriate for the sampled population. In addition, *OLStraj*'s individual growth trajectory estimates can allow for the assessment of whether sampled individuals are homogenous with respect to functional form, and may in some cases indicate that a growth mixture model should be considered (e.g., see Carrig & Bauer, 2001; Muthén & Shedden, 1999). Further, *OLStraj* can be used prior to model estimation to screen for extreme values of OLS-estimated individual growth parameters. As previously suggested, if such values are identified, associated observations might be checked for data entry errors or otherwise considered for removal from the sample. The *OLStraj* macro can also be used profitably as a post hoc device for the exploration of model misfit.

In addition, we have found that even after the estimation of a final model, the *OLStraj* macro can represent a highly useful adjunct to HLM-based or LTM-based growth curve modeling. The OLS estimates computed by the program, which can be saved by the program to an output data set, can be used as criterion variables in other analyses. They may also be utilized for the prediction of future observations

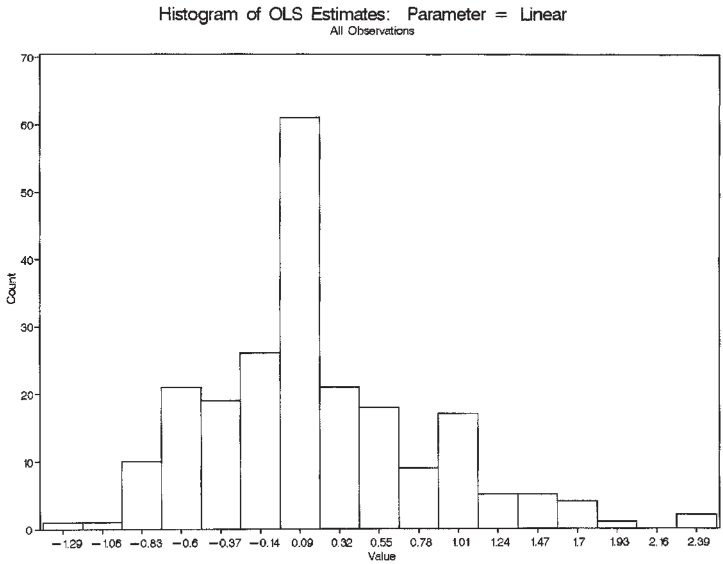


FIGURE 6 Histogram of ordinary least squares (OLS)-estimated slopes ($N = 221$).

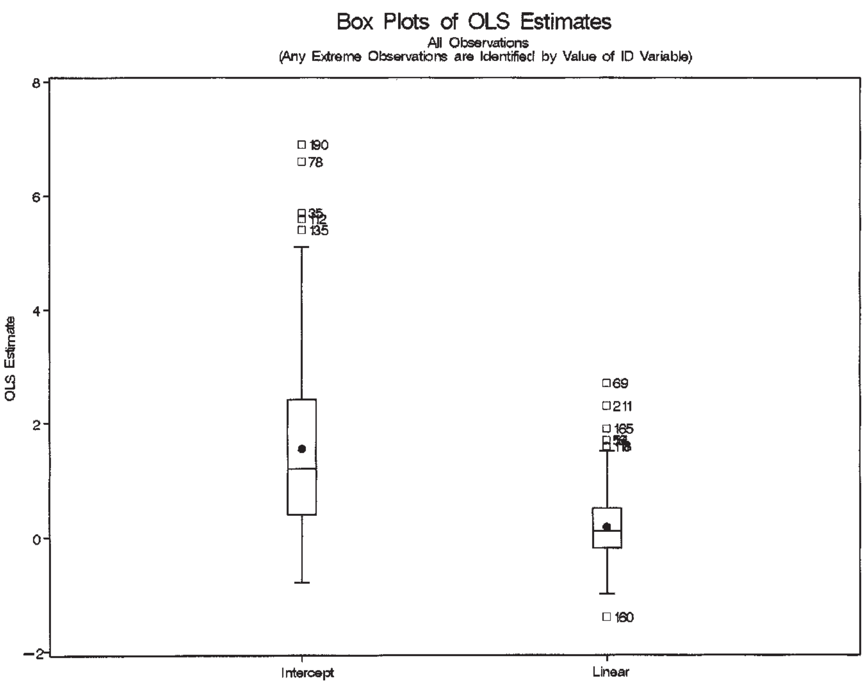


FIGURE 7 Box-and-whisker plots of ordinary least squares (OLS)-estimated intercepts and slopes ($N = 221$). ID = identification variable.

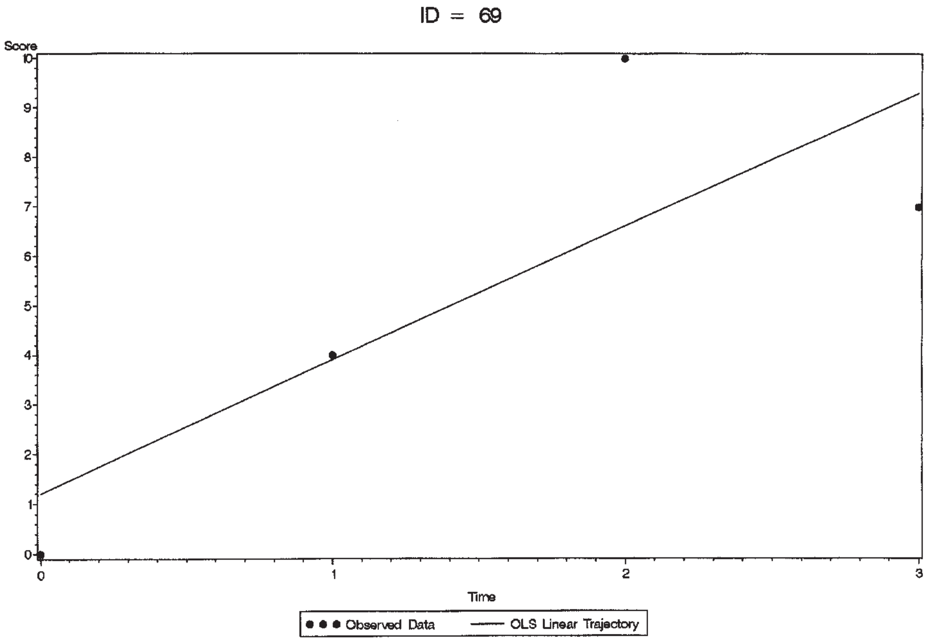


FIGURE 8 Estimated individual trajectory plot for identification variable (ID) value = 69, identified as a possible outlier by its extreme ordinary least squares (OLS)-estimated slope.

and for the derivation of cut scores. Finally, the plots produced by *OLStraj* can be directly exported to other applications for use in the preparation of manuscripts (or alternatively, the data set resulting from the program can be imported into the investigator's graphics program of choice). Depending on the investigator's particular application, reporting of results might include *OLStraj* plots that demonstrate careful data screening and evaluation of assumptions prior to the acceptance and interpretation of findings from complex growth curve models.

It should be noted that the case-by-case OLS estimator employed by *OLStraj* possesses important limitations (Biesanz et al., 2003). First, OLS trajectory estimation places restrictions on the structure of error variances, such that the variance of ε_{it} is assumed to equal the variance of ε_i and the variances $\varepsilon_{i1}, \varepsilon_{i2}, \ldots, \varepsilon_{it}$ are assumed to be uncorrelated. The assumption that the error variances within individuals are equal across time and uncorrelated may be inconsistent with the researcher's hypotheses regarding the true error structure within his or her data. Second, the OLS framework does not lend itself well to more complex models of change, such as the autoregressive, time-varying covariate, or multivariate LTM, or to the use of noncontinuous (e.g., categorical) observed data. Third, omnibus tests of model fit are not readily available for case-by-case OLS estimation of growth trajectories, thus rendering overall judgments about model evaluation difficult. A fourth limitation is that compared with constrained covariance estimates of individual trajectories, OLS estimates of individual intercept and slope are overdispersed and suggestive of a weaker relation between intercept and slope than actually exists (Biesanz et al., 2003). Therefore, future development on the *OLStraj* macro should emphasize the inclusion of additional approaches to individual trajectory estimation. An option to compute constrained covariance-estimated trajectories, in particular, could be a beneficial addition. Finally, future work on the *OLStraj* macro could incorporate the estimation of nonlinear functional forms of growth that are not members of the polynomial family (e.g., monomolecular or Gompertz functions estimated in SAS PROC NLIN), which would represent useful supplements to the linear and quadratic forms currently available in the program.

In conclusion, in our own work we have found the *OLStraj* macro to be a highly useful companion program to standard growth curve modeling software. Although we have focused on an SEM-based approach to growth curve estimation in this article, the *OLStraj* macro lends itself equally well to multilevel modeling approaches to growth curve modeling. It is hoped that the program will be of use to applied researchers who seek to maximize the effectiveness of LTMs in answering questions about stability and change.

ACKNOWLEDGMENTS

This work was supported in part by Grant DA13148 awarded to Patrick J. Curran and Kenneth A. Bollen.

The SAS macro, online manual, and empirical data can be downloaded from *www.unc.edu/~curran/OLStraj.htm*.

We would like to thank Ken Bollen and the Carolina Structural Equation Modeling Group at the University of North Carolina at Chapel Hill for their valuable input.

REFERENCES

Baker, P. C., Keck, C. K., Mott, F. L., & Quinlan, S. V. (1993). *NLSY child handbook: A guide to the 1986–1990 National Longitudinal Survey of Youth Child Data.* Columbus, OH: Center for Human Resource Research.

Biesanz, J. C., Curran, P. J., & Bollen, K. A. (2003). *Estimating individual trajectories in latent curve models.* Manuscript submitted for publication.

Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin, 101,* 147–158.

Carrig, M. M., & Bauer, D. J. (2001, April). *A comparison of cluster analysis and growth mixture modeling in the recovery of developmental trajectory classes.* Poster presented at the biannual meeting of the Society for Research in Child Development, Minneapolis, MN.

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

McArdle, J. J. (1988). Dynamic but structural modeling of repeated measures data. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (pp. 561–614). New York: Plenum.

McArdle, J. J., & Epstein, D. (1987). Latent growth curves within developmental structural equation models. *Child Development, 58,* 110–133.

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

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55,* 107–122.

Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics, 55,* 463–469.

Rogosa, D. R. (1994). Individual trajectories as the starting point for longitudinal data analysis. *Alzheimer Disease and Associated Disorders, 8,* 302–307.

Rogosa, D. R., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin, 90,* 726–748.

Rogosa, D., & Saner, H. (1995). Longitudinal data analysis examples with random coefficient models. *Journal of Educational and Behavioral Statistics, 20,* 149–170.

SAS Institute. (1990). *SAS/STAT user's guide: Vol. 1.* Cary, NC: Author.

Smith, L. D., Best, L. A., Stubbs, D. A., Archibald, A. B., & Roberson-Nay, R. (2002). Constructing knowledge: The role of graphs and tables in hard and soft psychology. *American Psychologist, 57,* 749–761.

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, MA: Addison-Wesley.

Wainer, H., & Velleman, P. F. (2000). Statistical graphics: Mapping the pathways of science. *Annual Review of Psychology, 12,* 201–228.