Paper 285-2013

Growth Spline Modeling

Matthew Schuelke, Air Force Research Laboratory; Robert Terry, University of Oklahoma; Eric Day, University of Oklahoma

ABSTRACT

Although researchers in many fields have long theorized about how the contributions of antecedents to outcomes might change across time, only recently have statistical tools necessary to examine such relationships become commonplace. In these situations the flexibility to simultaneously model ever smaller groups of observations both cross-sectionally (i.e., how the relative contribution of the antecedents differs among individuals) and longitudinally (i.e., how the relative contribution of the antecedents differs through time) is desirable (Singer, 1998). The SAS® MIXED proceedure is a relatively new tool which provides this requisite flexibility. In this paper we will present an extensible, hybrid statistical approach comprised of spline modeling and growth modeling which allows for an examination of how the relative antecedent contributions to an outcome change through time while simultaneously controlling for past effects.

INTRODUCTION

Interest in the influence of individual differences when investigating the progression of a quantity over time is ubiquitous across many research disciplines. For instance, the effects of consumer differences are commonly investigated in marketing as they relate to product trial, adoption, and use; and patient factors are examined in clinical studies as they relate to dose response curves. The goal in such situations goes beyond examining influences on the overall, average response pattern because differences among increasingly minute groups of patterns can provide additional insight while examining the relationship between antecedent and outcome (Cudeck & Harring, 2007). To complicate matters, it is often desirable to model change at specific time points (e.g., time varying treatments) under a non-linear response curve. This makes finding an appropriate polynomial model difficult, and as a result, model misspecification is common (Jo, Gossett, & Simpson, 2007). However, growth models allow us to circumvent problems associated with correlated observations when shifting focus to increasingly individual level analyses by directly modeling the correlations, and spline models allow a more appropriate fit to the data in addition to allowing change to be modeled at points specified by the researcher (Hurley, Hussey, McKeown, & Addy, 2006; Luke, 2004; Marsh & Cormier, 2001; Preacher, Wichman, MacCallum, & Briggs, 2008; Smith, 1979). In this paper, we present an extensible, hybrid statistical approach comprised of spline modeling and growth modeling which allows an examination of dynamic antecedent-outcome relationships while properly controlling for past effects.

AN EXTENSIBLE EXAMPLE OF GROWTH SPLINE MODELING

DATA REQUIREMENTS

The purpose of the present family of analyses is to examine deviations from established response patterns while simultaneously controlling for those preceding patterns. Because the focus of this family of analyses is on patterns, and the simplest pattern is a line between two points, the elemental child of this family consists of two linear trends and requires three longitudinal observations per subject. The initial, base trend begins with the first observation, and the other trend begins with the adjacent, second observation. Both trends continue through the remaining observations from their relative starting points.

Of course, additional longitudinal observations are always desirable because an increased number of observations allows for the fitting of more complex models. Supplementary observations may also add variance to each trend in overidentified models and thus avoid fitting issues such as those that prevent full solution estimation (e.g., non-positive definite hessian matrices). Therefore, this family of analyses is better suited for data sets containing an increased number of longitudinal observations, one or more covariates, and a response pattern at least nominally composed of segments (e.g., phases, stages, or steps).

Although data sets which possess the minimally-required attributes for this family of analyses are abundant in private circles, publically available data sets on which to showcase these techniques are relatively scarce. Therefore, data from the first author's dissertation is presently utilized, and a brief background is necessary for an understanding of the application example.

EXAMPLE DATA SET BACKGROUND

Extant skill acquisition theory posits that both the relative and absolute contributions of abilities to skill acquisition change through time (Ackerman, 2007), but previous tests of theory inadequately controlled for past acquisition

(Schuelke, 2010). In order to better test existing theory, participants were trained on a complex and dynamic computer-based task. In this paper, performance observations from the dissertation are used to model skill acquisition using three additive, linear acquisition trends. Furthermore, the influence on each trend of a time-invariant, individual-level ability covariate (i.e., general mental ability) is modeled. The dissertation copy available on ProQuest contains additional information which is not essential for understanding the current example.

The example in this paper is applied to a person-period data set in which each subject has one record for every observation period. The data set contains ten observations from each of 131 participants for a total of 1,310 records.

Although there are many ways to code for time when fitting longitudinal models, time is coded as 0, 1, 2, 3. . . in this paper. Therefore, the intercept estimates the value of the outcome variable, in this case performance score (i.e., skill attainment), at occasion 0 (i.e., initial status or origin) while the slopes estimate rate of change in the outcome across occasions, in this case skill acquisition.

UNCONDITIONAL LINEAR GROWTH MODEL

In order to illustrate our melded method, we begin with a simple two-level model, in which the level-1 model is a linear individual growth model, and the level-2 model expresses variation in parameters from the growth model as random effects unrelated to any person-level covariates. By convention, we represent the parameters in the level-1 (i.e., within person) model using π and the parameters in the level-2 (i.e., between-person) model using β. Thus, we may write the level-1 and level-2 models as:

$$Y_{ij} = \pi_{0j} + \pi_{1j}(TIME)_{ij} + r_{ij}$$

where $r_{ij} \sim N(0, \sigma^2)$

and

$$\pi_{0j} = \beta_{00} + u_{0j}$$

$$\pi_{1j} = \beta_{10} + u_{1j}$$

where

$$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix}\right]$$

Which can be written in combined form as:

$$Y_{ij} = [\beta_{00} + \beta_{10}(TIME)_{ij}] + [u_{0j} + u_{1j}(TIME)_{ij} + r_{ij}]$$

This multilevel model is expressed as the sum of two parts: a fixed part, which contains two fixed effects (for the intercept β_{00} and for the slope β_{10} of TIME) and a random part, which contains three random effects (for the intercept u_{0j}, the slope u_{1j} of TIME, and the within person residual r_{ij}). The choice of this formulation treats both the intercept π_{0j} and slope π_{1j} as random effects, and there are no level-2 covariates. This model can be fit with PROC MIXED via the following code:

```
proc mixed noclprint covtest;
class id;
model y = time/solution ddfm=bw notest;
random intercept time/subject=id type=un;
```

In this code the CLASS variable on the RANDOM statement indicates that when the random effects are specified, we want to allow both intercepts (i.e., β_{00}) and slopes (i.e., β_{10}) to vary across individuals. By using the SUBJECT=ID and TYPE=UN commands, we are requesting estimates of the inter-individual variance in intercepts (i.e., τ_{00}) as well as slopes (i.e., τ_{11}) in addition to the covariance between intercepts and slopes (i.e., τ_{10}, which equals τ_{01} because of symmetry).

The MODEL statement indicates what type of growth model is to be fit. In the current model we use unconditional linear growth, but in the next model we will attempt to predict inter-individual differences in this growth. In our third model we build upon these two progressive models by breaking the linear growth into additive splines across time not only to find a better fitting model, but more so because such parameterization allows us to examine changes in the contributions of an antecedent to an outcome across time while properly controlling for past contributions.

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion

```
0               1      23449.26971685
1               3      22176.47983628      0.00000019
2               1      22176.47791494      0.00000000


                 Convergence criteria met.

                Covariance Parameter Estimates

                                   Standard         Z
Cov Parm     Subject     Estimate     Error     Value        Pr Z

UN(1,1)      id           1516408    229708      6.60      <.0001
UN(2,1)      id             99599     16566      6.01      <.0001
UN(2,2)      id           6085.53   2257.03      2.70      0.0035
Residual                   960634     41965     22.89      <.0001

                       Fit Statistics

              -2 Res Log Likelihood         22176.5
              AIC (smaller is better)       22184.5
              AICC (smaller is better)      22184.5
              BIC (smaller is better)       22196.0

               Null Model Likelihood Ratio Test

                 DF     Chi-Square      Pr > ChiSq

                  3        1272.79          <.0001

                  Solution for Fixed Effects

                             Standard
 Effect        Estimate        Error       DF     t Value     Pr > |t|

 Intercept      -208.05       118.78      130       -1.75       0.0822
 time            472.80      11.6336     1178       40.64       <.0001
```

Output 1. Output from an Unconditional Linear Growth Model

The first section of output contains information about the iteration history. PROC MIXED converged quickly because three iterations is only one more than the minimum necessary for an evaluation of convergence. Such rapid convergence is less likely under more sophisticated models and with data containing missing values or high degrees of collinearity.

The second section of output contains the covariance parameter estimates (i.e., random effects), which may be written in matrix form as follows:

$$\begin{pmatrix} \hat{\tau}_{00} & \hat{\tau}_{01} \\ \hat{\tau}_{10} & \hat{\tau}_{11} \end{pmatrix} = \begin{pmatrix} 1516408 & 99599 \\ 99599 & 6086 \end{pmatrix}$$

SAS also reports σ^2 as being 960,634, and produces accompanying standard errors, and hypothesis tests of the null hypotheses that these population variances and covariances are 0. All of the tests reject in this case, but we are most interested in the findings for τ_{00} and τ_{11}. These results indicate there is variance among individual intercepts, τ_{00} = 1,516,408, z = 6.60, p < .001, and slopes, τ_{11} = 6,085.53, z = 2.70, p < .01, which could potentially be explained by the addition of one or more person-level (i.e., level 2) covariates.

Next, the output contains some common goodness of fit statistics. These statistics are helpful for evaluating and comparing model fit with other nested models. Please consult that SAS manual for details on these statistics.

Finally, the output contains estimates and test information for the fixed effects. Because this is an individual growth model with no level-2 covariates, the effects can be interpreted in the typical manner. The average person began with a score of $ß_{00}$ = –208.05, $t(130) = –1.75$, $p < .10$, and gained $ß_{10}$ = 472.80 points per testing occasion, $t(1,178)$ = 40.64, $p < .01$.

LINEAR GROWTH MODEL WITH A PERSON-LEVEL COVARIATE

In our second example, we build upon the previous model by adding a standardized person-level covariate in order to explore whether variation in intercepts or slopes is related to the covariate (i.e., general mental ability).

$$Y_{ij} = \pi_{0j} + \pi_{1j}(TIME)_{ij} + r_{ij}$$

$$\text{where } r_{ij} \sim N(0, \sigma^2)$$

$$\text{and}$$

$$\pi_{0j} = ß_{00} + ß_{01}(z_{COVAR}) + u_{0j}$$

$$\pi_{1j} = ß_{10} + ß_{11}(z_{COVAR}) + u_{1j}$$

$$\text{where}$$

$$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix}\right]$$

Which can be written in combined form as:

$$Y_{ij} = [ß_{00} + ß_{01}(z_{COVAR}) + ß_{10}(TIME)_{ij} + ß_{11}(z_{COVAR})(TIME)_{ij}] + [u_{0j} + u_{1j}(TIME)_{ij} + r_{ij}]$$

We standardize the covariate by subtracting its mean and dividing by its standard deviation in order to aide in interpretation of the model. The interpretations of the fixed effects for $ß_{00}$ and $ß_{10}$ are therefore the overall average intercept and average slope as opposed to the intercept and slope of a case when COVAR is equal to 0. Standardizing the covariate also allows any effect to be more easily discussed in terms of standard deviations.

This model can be fit with PROC MIXED via the following code.

```
proc mixed noclprint covtest;
class id;
model y = time scovar time*scovar/solution ddfm=bw notest;
random intercept time/subject=id type=un;
```

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	23230.47392614	
1	3	22131.31246104	0.00000022
2	1	22131.31029556	0.00000000

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	id	1237669	195966	6.32	<.0001
UN(2,1)	id	79986	14558	5.49	<.0001
UN(2,2)	id	4860.83	2116.98	2.30	0.0108
Residual		960634	41966	22.89	<.0001

Fit Statistics

-2 Res Log Likelihood	22131.3
AIC (smaller is better)	22139.3
AICC (smaller is better)	22139.3
BIC (smaller is better)	22150.8

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
3	1099.16	<.0001

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-208.05	109.46	129	-1.90	0.0596
time	472.80	11.2246	1177	42.12	<.0001
scovar	539.27	109.88	129	4.91	<.0001
time*scovar	36.7636	11.2677	1177	3.26	0.0011

Output 2. Output from a Linear Growth Model with a Person-Level Covariate

Although the output of this model is identical in layout to that of the previous model, there are a few similarities and differences of note within some sections:

First, because we standardized the level-2 covariate, the estimates for the intercept and for time (i.e., $ß_{00}$ and $ß_{10}$) are identical to what they were in the unconditional model of the previous section. The interpretation is also similar except that one must now append "controlling for the covariate".

Second, this model adds estimates and tests for the covariate and its interaction with time. The estimate of $ß_{01}$ = 539.27, $t(129) = 4.91$, $p < .001$, indicates that individuals who differ by a standard deviation with respect to the covariate have performance scores which differ by 539.27 points on average. The estimate of $ß_{11}$ = 36.76, $t(1177) = 3.26$, $p < .01$, indicates that individuals who differ by a standard deviation with respect to the covariate have growth (i.e., acquisition) rates which differ by 36.76 points.

Third, the estimate for σ^2 has remained unchanged at 960,634, $z = 22.89$, $p < .001$. But the estimates in the variance-covariance matrix of the intercepts and slopes have changed to:

$$\begin{pmatrix} \hat{\tau}_{00} & \hat{\tau}_{01} \\ \hat{\tau}_{10} & \hat{\tau}_{11} \end{pmatrix} = \begin{pmatrix} 1237669 & 79986 \\ 79986 & 4861 \end{pmatrix}$$

When comparing these new estimates to those from the unconditional model of the previous section, we see that inclusion of the covariate helped predict both initial status and growth rates because the intercept and slope variance estimates (i.e., τ_{00} and τ_{11}, respectively) have decreased. Specifically, inclusion of the covariate reduced intercept variance by [(1,516,408 – 1,237,669) / 1,516,408 = 0.1838] 18% and the growth rate variance by [(6085.53 – 4860.83) / 6085.53 = 0.2012] 20%. In other words, the covariate accounts for 18% of the explainable variance in initial status and 20% of the explainable variance in growth rates.

LINEAR GROWTH SPLINE MODEL WITH A PERSON-LEVEL COVARIATE

In our final example, we build upon the previous two examples and break the linear growth parameter (π_{1j}) into additive splines.

$$Y_{ij} = \pi_{0j} + \pi_{1j}(TIME)_{ij} + \pi_{2j}D_{1i}(TIME - TIME_1)_{ij} + \pi_{3j}D_{2i}(TIME - TIME_2)_{ij} + r_{ij}$$

where $r_{ij} \sim N(0, \sigma^2)$

and

$$\pi_{0j} = ß_{00} + ß_{01}(z_{COVAR}) + u_{0j}$$

$$\pi_{1j} = ß_{10} + ß_{11}(z_{COVAR}) + u_{1j}$$

$$\pi_{2j} = ß_{20} + ß_{21}(z_{COVAR}) + u_{2j}$$

$$\pi_{3j} = ß_{30} + ß_{31}(z_{COVAR}) + u_{3j}$$

where

$$\begin{pmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \\ u_{3j} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} & \tau_{02} & \tau_{03} \\ \tau_{10} & \tau_{11} & \tau_{12} & \tau_{13} \\ \tau_{20} & \tau_{21} & \tau_{22} & \tau_{23} \\ \tau_{30} & \tau_{31} & \tau_{32} & \tau_{33} \end{pmatrix}\right]$$

Which can be written in combined form as:

$$Y_{ij} = [\beta_{00} + \beta_{01}(z_{COVAR}) + \beta_{10}(TIME)_{ij} + \beta_{11}(z_{COVAR})(TIME)_{ij} + \beta_{20}D_{1i}(TIME - TIME_1)_{ij} + \beta_{21}(z_{COVAR})D_{1i}(TIME - TIME_1)_{ij} + \beta_{30}D_{2i}(TIME - TIME_2)_{ij} + \beta_{31}(z_{COVAR})D_{2i}(TIME - TIME_2)_{ij}] + [u_{0j} + u_{1j}(TIME)_{ij} + u_{2j}D_{1i}(TIME - TIME_1)_{ij} + u_{3j}D_{2i}(TIME - TIME_2)_{ij} + r_{ij}]$$

In this model Y_{ij} is the response of an individual at a given time point while π_{0j} is an intercept coded as response at the origin. The parameter π_{1j} represents the initial spline and the underlying linear trend, conditional on the other splines, throughout the response period. The $TIME_1$ and $TIME_2$ variables are set times since the origin denoting spline starting points (e.g., knots). The D variables (i.e., D_{1i} and D_{2i}) are dummy variables equal to 0 when the amount of time from the origin, is less than $TIME_1$ and $TIME_2$ respectively, and equal to 1 when $TIME$ is greater than i_1 and i_2 respectively. The remaining parameters (π_{2j} and π_{3j}) become summative deviations to the underlying trend when enough time passes according to the D dummy variables.

Individually the splines represent response in a particular segment of time while controlling for past effects, but collectively they can capture the nature of the overall response trend. When applied to data with a decelerating logarithmic response pattern as in the current example, the resulting model is akin to that represented in Figure 1.

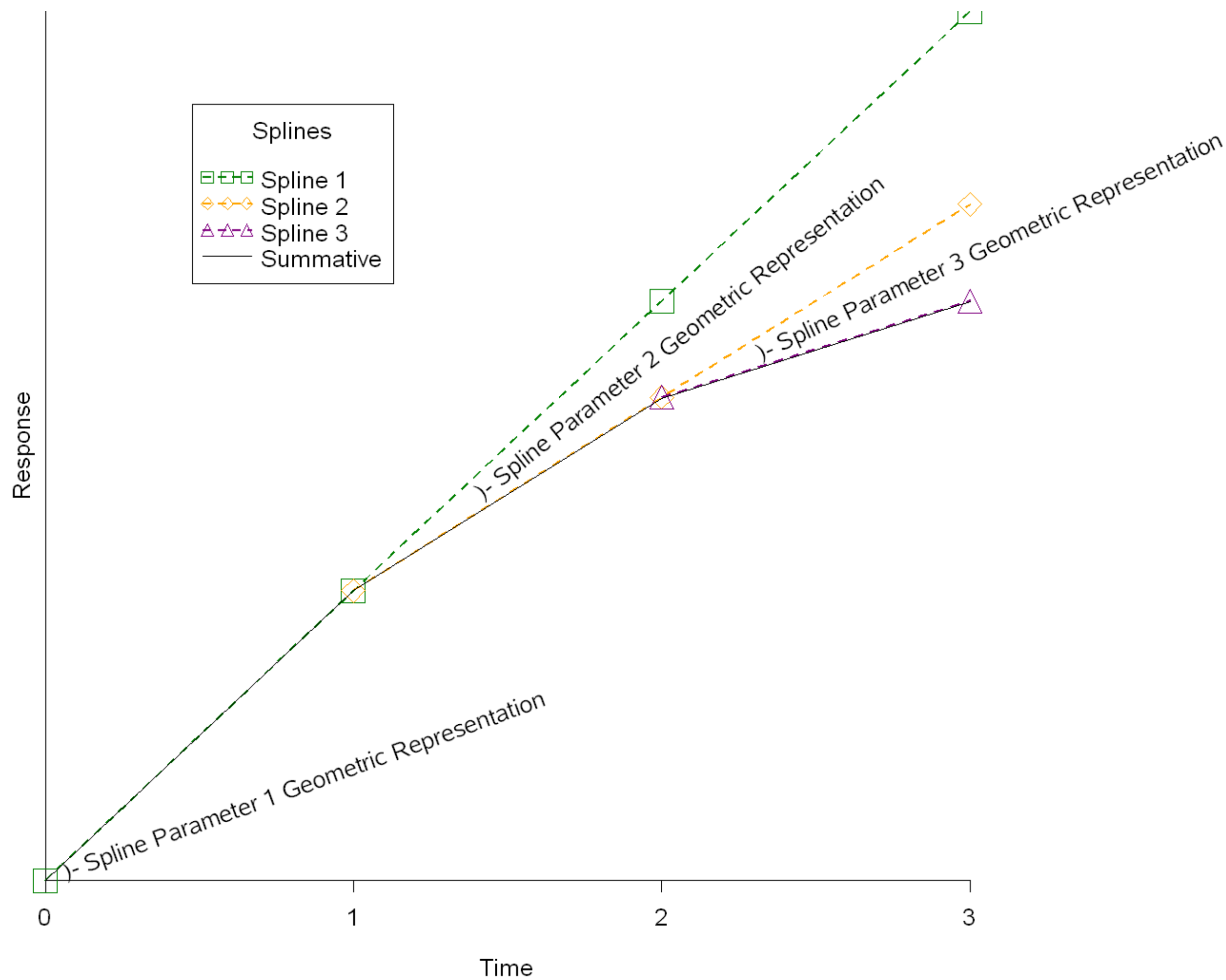

Figure 1. A Basic Hypothetical Example of Additive Splines

In the first time period of Figure 1, only the first spline is present, but this effect continues throughout the remaining time periods. In the second time period, the second spline begins—as per the time-based dummy-coding—in addition to the continuation of the first spline. This model parameterization method allows for the estimation of antecedent contributions to outcome at the second time period while controlling for previous contributions from the first time period. Additionally, the total antecedent contribution during the second time period can be estimated via summation of the first two spline parameter values (i.e., calculating overall model slope in the second time period). In the third time period, the third spline begins in addition to the continuation of the first and second splines. Similar to the second time period, the antecedent contributions to outcome at the third time period can be estimated while controlling for previous antecedent contributions from the previous two time periods, or the total antecedent contribution during the third time period can be estimated via summation of all currently present spline parameter values.

Table 1 presents the coding for the time-based spline variables in the hypothetical example.

	Measurement Occasion				
Variable	**1**	**2**	**3**	**4**	**Interpretation**
Spline 1	0	1	2	3	Underlying linear change (e.g., base acquisition rate)
Spline 2	0	0	1	2	Linear deviation to the underlying linear change (e.g., change from base acquisition starting in second time period)
Spline 3	0	0	0	1	Deviation to the linear deviation to the underlying linear change (e.g., change in acquisition rate starting in the third time period)

Table 1. Coding and Interpretation of Hypothetical Spline Variables

Once values for a time variable have been established, values for the spline variables can be arrived at in a DATA step with code similar to the following:

```
spline1 = time;
spline2 = (time > 1)*(time – 1);
spline3 = (time > 2)*(time – 2);
```

Returning to our discussion of the linear growth spline model with a person-level covariate, the model can be fit with PROC MIXED via the following code.

```
proc mixed noclprint covtest;
class id;
model y = scovar spline1 spline2 spline3
          scovar*spline1 scovar*spline2 scovar*spline3/solution ddfm=bw notest;
random intercept spline1 spline2 spline3/subject=id type=un;
```

```
                         Iteration History

     Iteration    Evaluations    -2 Res Log Like      Criterion

             0              1     22988.96993683
             1              1     21016.85219819     0.00000000

                      Convergence criteria met.

                   Covariance Parameter Estimates

                                             Standard         Z
  Cov Parm     Subject     Estimate       Error       Value        Pr Z

  UN(1,1)      id            588779      105807        5.56      <.0001
  UN(2,1)      id            228106       54421        4.19      <.0001
  UN(2,2)      id            297471       54852        5.42      <.0001
```

```
        UN(3,1)      id          -264216       69585     -3.80      0.0001
        UN(3,2)      id          -289927       67092     -4.32      <.0001
        UN(3,3)      id           331785       91639      3.62      0.0001
        UN(4,1)      id            34109       37747      0.90      0.3662
        UN(4,2)      id           -44341       27565     -1.61      0.1077
        UN(4,3)      id          2410.34       40355      0.06      0.9524
        UN(4,4)      id            51734       28354      1.82      0.0340
        Residual                  307539       15513     19.82      <.0001

                               Fit Statistics

                     -2 Res Log Likelihood         21016.9
                     AIC (smaller is better)       21038.9
                     AICC (smaller is better)      21039.1
                     BIC (smaller is better)       21070.5

                       Null Model Likelihood Ratio Test

                         DF     Chi-Square      Pr > ChiSq

                         10        1972.12          <.0001

                         Solution for Fixed Effects

                                         Standard
      Effect              Estimate          Error       DF    t Value    Pr > |t|

      Intercept           -1598.83        80.2414      129     -19.93      <.0001
      ccovar                402.17        80.5494      129       4.99      <.0001
      spline1              1606.33        57.7497     1173      27.82      <.0001
      spline2             -1284.13        74.0876     1173     -17.33      <.0001
      spline3             -76.3486        40.7541     1173      -1.87      0.0613
      scovar*spline1        146.06        57.9714     1173       2.52      0.0119
      scovar*spline2       -118.73        74.3720     1173      -1.60      0.1107
      scovar*spline3      -14.8630        40.9105     1173      -0.36      0.7164
```

Output 3. Output from a Linear Growth Spline Model with a Person-Level Covariate

Similar to the two preceding models, this model converged quickly. However, there are now many more interesting random and fixed effects.

As with the estimated random effects of the previous models, we are interested in the covariance parameter estimates, which can be written in matrix form as follows:

$$\begin{pmatrix} \hat{\tau}_{00} & \hat{\tau}_{01} & \hat{\tau}_{02} & \hat{\tau}_{03} \\ \hat{\tau}_{10} & \hat{\tau}_{11} & \hat{\tau}_{12} & \hat{\tau}_{13} \\ \hat{\tau}_{20} & \hat{\tau}_{21} & \hat{\tau}_{22} & \hat{\tau}_{23} \\ \hat{\tau}_{30} & \hat{\tau}_{31} & \hat{\tau}_{32} & \hat{\tau}_{33} \end{pmatrix} = \begin{pmatrix} 588779 & 228106 & -264216 & 34109 \\ 228106 & 297471 & -289927 & -44341 \\ -264216 & -289927 & 331785 & 2410.34 \\ 34109 & -44341 & 2410.34 & 51734 \end{pmatrix}$$

SAS reports σ^2 as being 307,539, which is a substantial decrease when compared to the previous models (c.f., 960,634). However, we must remember that this reduction in variance has mostly been divvied up among the new linear trends. There appears to be significant variance between people in their initial skill attainment, τ_{00} = 588,779, z = 5.56, p < .001, base skill acquisition, τ_{11} = 297,471, z = 5.42, p < .001, linear deviation from base acquisition, τ_{22} = 331,785, z = 3.62, p < .001, and final deviation from previously established acquisition, τ_{33} = 51,734, z = 19.82, p < .001, which might all be predicted with additional person-level covariates.

Finally, we are also interested in the fixed effects estimates. The average individual has an initial skill attainment score of $ß_{00}$ = –1598.83, t(129) = –19.93, p < .001, and starts acquiring skill at a rate of $ß_{10}$ = 1606.33, t(1173) = 27.82, p < .001, per testing occasion. The average individual then experiences an acquisition deceleration of $ß_{20}$ = –1284.13, t(1173) = –17.33, p < .001, per testing occasion and then another possible deceleration of $ß_{30}$ = –76.35, t(1173) = –1.87, p < .10. With respect to the covariate, individuals who differ by a standard deviation in general mental ability have performance scores which differ by $ß_{01}$ = 402.17, t(129) = 4.99, p < .001, points on average and have initial skill acquisition rates which differ by $ß_{11}$ = 146.06, t(1173) = 2.52, p < .05, points on average. However, general mental ability appears to have little effect on later acquisition decelerations, $ß_{21}$ = –118.73, t(1173) = –1.60, p = .11, and $ß_{31}$ = –14.86, t(1173) = –0.36, p < .72. Note, although not done so here, the covariate-spline estimates can be statistically compared via CONTRAST statements, and an acquisition rate (i.e., model tangent) can be estimated for any desired time point by adding all previous acquisiton rates (i.e., splines) together in an ESTIMATE statement.

CONCLUSION

As we have demonstrated, the combination of growth modeling and spline modeling is a powerful statistical technique for the investigation of dynamic antecedent-outcome relationships which is easily implemented in PROC MIXED. This parameterization method allows for the comparison of correctly specified and controlled for antecedent contributions to outcome both within and across time periods and is easily extensible in that can be extended to many other, more complicated situations including additional time periods with many antecedents. Furthermore, the splines themselves can be specified to be more sophisticated than simple linear trends, and other overall trends are easily accommodated. In short, the possibilities are endless.

REFERENCES

- Ackerman, P. L. (2007). New developments in understanding skilled performance. *Current Directions in Psychological Science, 16*(5), 235-239.
- Cudeck, R., & Harring, J. R. (2007). Analysis of nonlinear patterns of change with random coefficient models. *Annual Review of Psychology, 58*, 615-637.
- Hurley, D., Hussey, J., McKeown, R., & Addy, C. (2006). *An evaluation of splines in linear regression*. Paper presented at SUGI 31, San Francisco, CA.
- Jo, C. H., Gossett, J., & Simpson, P. (2007). *Regression splines with longitudinal data*. Poster presented at SAS Global Forum 2007, Orlando, FL.
- Luke, D. A. (2004). *Multilevel modeling*. Sage University Paper series on Quantitative Applications in the Social Sciences, 07-143. Thousand Oaks, CA: Sage.
- Marsh, L., & Cormier, D. R. (2001). *Spline regression models*. Sage University Paper series on Quantitative Applications in the Social Sciences, 07-137. Thousand Oaks, CA: Sage.
- Preacher, K. J., Wichman, A. L., MacCallum, R. C., & Briggs, N. E. (2008). *Latent growth curve modeling*. Sage University Paper series on Quantitative Applications in the Social Sciences, 07-157. Thousand Oaks, CA: Sage.
- Schuelke, M. J. (2010). *Ability-growth interactions in the acquisition of a complex skill: A spine-modeling approach*. (Doctoral Dissertation). Retrieved from ProQuest Dissertations & Theses. (3439070).
- Singer, J. D. (1998). Using SAS Proc Mixed to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics, 24*(4), 323-255.
- Smith, P. L. (1979). Splines as a useful and convenient statistical tool. *The American Statistician, 33*(2), 57-62.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the first author at:

Matthew Schuelke
Air Force Research Laboratory
2215 First Street, Building 33
Wright-Patterson Air Force Base, OH 45433
Work Phone: (937) 255-2240
Fax: (937) 785-8752
E-mail: matthew.schuelke.ctr@wpafb.af.mil