Multivariate Growth Models for Older and Younger Sibling Risky Behavior

SAS Data Set-Up:

```
* Multivariate set-up;
DATA work.BivFamily; SET work.Family;
      risky=orisky; age=siboage; boy=Oboy; DV="SibO"; dvO=1; dvY=0; OUTPUT;
      risky=yrisky; age=sibyage; boy=Yboy; DV="SibY"; dvO=0; dvY=1; OUTPUT;
RUN;
* Centering older and younger at different place;
DATA work.BivFamily; SET work.BivFamily;
      IF DV="SibO" THEN endtime=age-17; * Up to 6 occasions (12-17);
      IF DV="SibY" THEN endtime=age-15; * Up to 3 occasions (13-15);
      LABEL endtime= "Time in Study (0=17 for O, 0=15 for Y)";
RUN;
```

With this coding we have created a time-in-study model (given that age=time within each) where age is centered at the last measurement occasion separately for each sibling (approximately age 17 for Sib O and 15 for Sib Y). Time is treated as continuous from these points (i.e., an exact age model). The differing centering points allows us to put each DV on its own metric, and to not assume age-based convergence. Thus, the intercept for dvO (Sib O) represents expected risky behaviors at age 17, and the intercept for dvY (Sib Y) represents expected risky behaviors at age 15. We can now examine the extent to which intercepts, slopes, and residuals are related across time of study.

```
TITLE1 "SAS Method #1 Direct Effects Unconditional Multivariate Growth Model for Risky Behavior";
PROC MIXED DATA=BivFamily COVTEST NOCLPRINT NOITPRINT MAXITER=1000 IC NAMELEN=100 METHOD=ML;
      CLASS family phase DV;
      MODEL risky = dvO dvY dvO*endtime dvY*endtime dvO*endtime*endtime dvY*endtime*endtime
                    / NOINT SOLUTION DDFM=Satterthwaite;
      RANDOM dvO dvY dvO*endtime dvY*endtime / G GCORR TYPE=UN SUBJECT=family; * Level 2 family;
      REPEATED DV / R RCORR TYPE=UN SUBJECT=phase*family;            * Level 1 crossed time*DV;
RUN;
```

```
            Dimensions
Covariance Parameters            13
Columns in X                      6
Columns in Z Per Subject          4
Subjects                        200 # Families
Max Obs Per Subject              12 # Obs per family (2 sibs * 6 times)

          Number of Observations
Number of Observations Read            2400
Number of Observations Used            1627
Number of Observations Not Used         773
```

R and RCORR matrices were not printed! I think it has to do with the unbalanced data...

Estimated G Matrix

Row	Effect	FAMILY	Col1	Col2	Col3	Col4
1	dvO	1	64.8158	24.5476	10.3524	4.6851
2	dvY	1	**24.5476**	50.0280	3.8068	10.2330
3	dvO*endtime	1	10.3524	3.8068	**1.7883**	0.7835
4	dvY*endtime	1	4.6851	10.2330	0.7835	2.0165

Estimated G Correlation Matrix

Row	Effect	FAMILY	Col1	Col2	Col3	Col4
1	dvO	1	1.0000	0.4311	0.9616	0.4098
2	dvY	1	**0.4311**	1.0000	0.4025	1.0000
3	dvO*endtime	1	0.9616	0.4025	1.0000	0.4126
4	dvY*endtime	1	0.4098	1.0000	**0.4126**	1.0000

```
                    Covariance Parameter Estimates
                                        Standard         Z
Cov Parm      Subject             Estimate       Error     Value        Pr Z
UN(1,1)       FAMILY               64.8158      7.4630      8.68      <.0001
UN(2,1)       FAMILY               24.5476      5.3952      4.55      <.0001 Intercept covariance
UN(2,2)       FAMILY               50.0280      6.8576      7.30      <.0001
UN(3,1)       FAMILY               10.3524      1.3195      7.85      <.0001
UN(3,2)       FAMILY                3.8068      0.9745      3.91      <.0001
UN(3,3)       FAMILY                1.7883      0.2555      7.00      <.0001
UN(4,1)       FAMILY                4.6851      1.7595      2.66      0.0078
UN(4,2)       FAMILY               10.2330      2.0845      4.91      <.0001
UN(4,3)       FAMILY                0.7835      0.3139      2.50      0.0126 Slope covariance
UN(4,4)       FAMILY                2.0165      0.7060      2.86      0.0021
UN(1,1)       FAMILY*phase         13.9353      0.7586     18.37      <.0001
UN(2,1)       FAMILY*phase         -0.7855      0.7540     -1.04      0.2975 Residual covariance
UN(2,2)       FAMILY*phase         14.1744      1.2213     11.61      <.0001

                           Information Criteria
Neg2LogLike     Parms        AIC        AICC        HQIC        BIC       CAIC
     9811.1        19     9849.1      9849.6      9874.5     9911.8     9930.8

                        Solution for Fixed Effects
                                         Standard
Effect                    Estimate          Error       DF    t Value    Pr > |t|
dvO                        28.8527         0.6107      191      47.25      <.0001 SibO int at age 17
dvO*endtime                 1.9504         0.1988      718       9.81      <.0001 SibO linear at age 17
dvO*endtime*endtime        0.08846        0.03297      911       2.68      0.0074 SibO quad at any age
dvY                        26.3777         0.5890      165      44.78      <.0001 SibY int at age 15
endtime*dvY                 1.5046         0.3218      267       4.68      <.0001 SibY linear at age 15
endtime*endtimevdvY        0.04982        0.08117      123       0.61      0.5405 SibY quad at any age
```

Because the time metrics cover different ages, it is most appropriate to leave the growth terms separate for each Sib. We can remove the nonsignificant quadratic effect of time for Sib Y (not surprising given only 3 time points).

Examining own and unidirectional effects of sibling sex (O→Y) on growth in risky behavior:

```
TITLE1 "Adding Effects of Gender: Self and O-->Y";
PROC MIXED DATA=BivFamily COVTEST NOCLPRINT NOITPRINT MAXITER=1000 IC NAMELEN=100 METHOD=ML;
      CLASS family phase DV;
      MODEL risky = dvO dvO*endtime dvO*endtime*endtime dvY dvY*endtime
                    dvO*Oboy dvO*Oboy*endtime dvO*Oboy*endtime*endtime
                    dvY*Yboy dvY*Yboy*endtime dvY*Oboy dvY*Oboy*endtime
                          / NOINT SOLUTION DDFM=Satterthwaite;
      RANDOM dvO dvY dvO*endtime dvY*endtime / G GCORR TYPE=UN SUBJECT=family; * Level 2 family;
      REPEATED DV / R RCORR TYPE=UN SUBJECT=phase*family;          * Level 1 crossed time*DV;
RUN;
```

```
                         Estimated G Matrix
 Row     Effect          FAMILY         Col1         Col2         Col3         Col4
   1     dvO                  1      64.4248      22.1749      10.4961       4.2930
   2     dvY                  1      22.1749      49.3236       3.6434      10.2545
   3     dvO*endtime          1      10.4961       3.6434       1.7983       0.7672
   4     endtime*dvY          1       4.2930      10.2545       0.7672       1.9310

                    Estimated G Correlation Matrix
 Row     Effect          FAMILY         Col1         Col2         Col3         Col4
   1     dvO                  1       1.0000       0.3934       0.9751       0.3849
   2     dvY                  1       0.3934       1.0000       0.3869       1.0000
   3     dvO*endtime          1       0.9751       0.3869       1.0000       0.4117
   4     endtime*dvY          1       0.3849       1.0000       0.4117       1.0000

                           Information Criteria
Neg2LogLike     Parms        AIC        AICC        HQIC        BIC       CAIC
     9766.6        25     9816.6      9817.4      9850.0     9899.1     9924.1
```

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|	
dvO	28.2084	0.8537	194	33.04	<.0001	Sib O growth for girls
dvO*endtime	2.3724	0.2826	747	8.40	<.0001	
dvO*endtime*endtime	0.1580	0.04543	901	3.48	0.0005	
dvY	25.4502	0.9984	183	25.49	<.0001	Sib O growth: girls w/ SibO=G
endtime*dvY	1.4204	0.2972	59.3	4.78	<.0001	
dvO*Oboy	1.5007	1.2186	190	1.23	0.2197	Own gender on Sib O
dvO*endtime*Oboy	-0.7656	0.3955	717	-1.94	0.0533	
dvO*endtime*endtime*Oboy	-0.1281	0.06536	906	-1.96	0.0504	
dvY*Yboy	-0.9314	1.1041	166	-0.84	0.4001	Own gender on Sib Y
endtime*dvY*Yboy	-1.0286	0.3571	100	-2.88	0.0049	
dvY*Oboy	2.6318	1.1648	170	2.26	0.0251	Sib O gender on Sib Y
endtime*dvY*Oboy	0.7201	0.3710	138	1.94	0.0543	

Equation for final multivariate model with gender:

L1: $Risk_{tfs} = dvO\,[\,\beta_{0fO} + \beta_{1fO}(age_{tfO} - 17) + \beta_{2fO}(age_{tfO} - 17)^2 + e_{tfO}\,] + dvY\,[\,\beta_{0fY} + \beta_{1fO}(age_{tfY} - 15) + e_{tfY}\,]$

L2:
$\beta_{0fO} = \gamma_{00O} + \gamma_{01O}(OlderBoy_f) + U_{0fO}$
$\beta_{1fO} = \gamma_{10O} + \gamma_{11O}(OlderBoy_f) + U_{1fO}$
$\beta_{2fO} = \gamma_{20O} + \gamma_{21O}(OlderBoy_f)$
$\beta_{0fY} = \gamma_{00Y} + \gamma_{01Y}(OlderBoy_f) + \gamma_{02Y}(YoungerBoy_f) + U_{0fY}$
$\beta_{1fY} = \gamma_{10Y} + \gamma_{11Y}(OlderBoy_f) + \gamma_{12Y}(YoungerBoy_f) + U_{1fY}$

Creating fake families for all combinations of sibling gender in excel:
NOTE: YOU MUST HAVE VALUES FOR ALL VARIABLES IN THE MODEL TO DO THIS.

order	family	dv	dvO	dvY	phase	age	endtime	Oboy	Yboy
1	-99	SibO	1	0	2	12	-5	1	1
1	-99	SibO	1	0	3	13	-4	1	1
1	-99	SibO	1	0	4	14	-3	1	1
1	-99	SibO	1	0	6	15	-2	1	1
1	-99	SibO	1	0	7	16	-1	1	1
1	-99	SibO	1	0	8	17	0	1	1
2	-99	SibY	0	1	6	13	-2	1	1
2	-99	SibY	0	1	7	14	-1	1	1
2	-99	SibY	0	1	8	15	0	1	1
3	-98	SibO	1	0	2	12	-5	1	0
3	-98	SibO	1	0	3	13	-4	1	0
3	-98	SibO	1	0	4	14	-3	1	0
3	-98	SibO	1	0	6	15	-2	1	0
3	-98	SibO	1	0	7	16	-1	1	0
3	-98	SibO	1	0	8	17	0	1	0
4	-98	SibY	0	1	6	13	-2	1	0
4	-98	SibY	0	1	7	14	-1	1	0
4	-98	SibY	0	1	8	15	0	1	0
5	-97	SibO	1	0	2	12	-5	0	1
5	-97	SibO	1	0	3	13	-4	0	1
5	-97	SibO	1	0	4	14	-3	0	1
5	-97	SibO	1	0	6	15	-2	0	1
5	-97	SibO	1	0	7	16	-1	0	1
5	-97	SibO	1	0	8	17	0	0	1

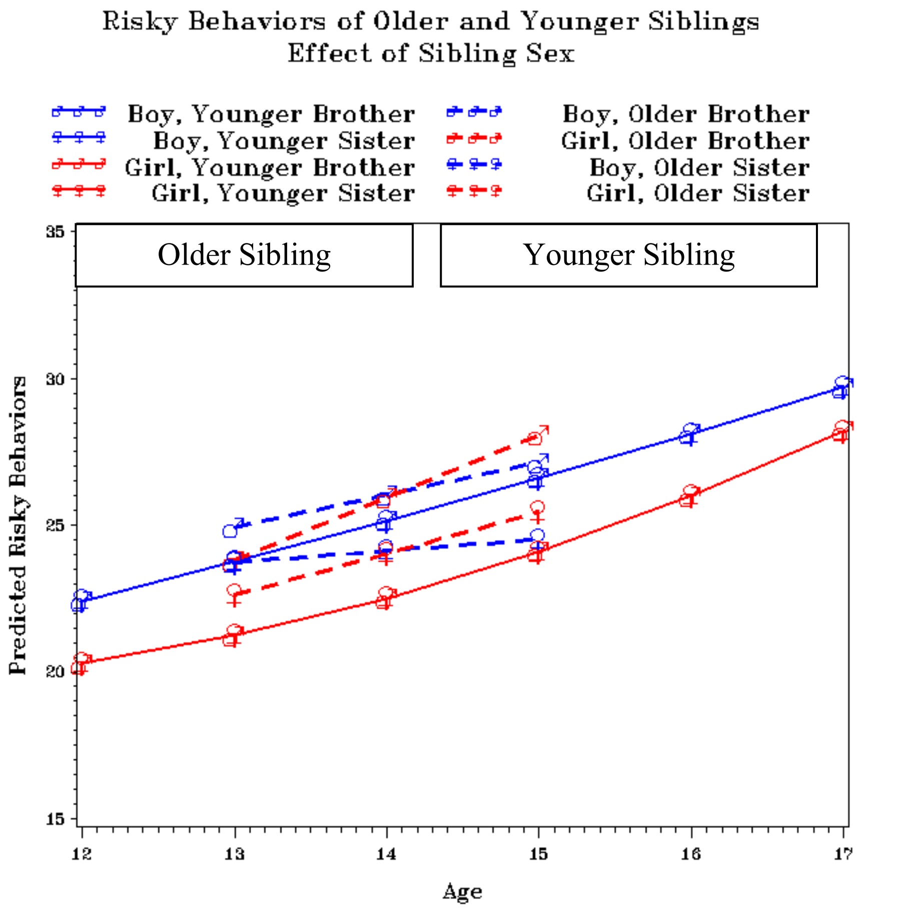
6	-97	SibY	0	1	6	13	-2	0	1
6	-97	SibY	0	1	7	14	-1	0	1
6	-97	SibY	0	1	8	15	0	0	1
7	-96	SibO	1	0	2	12	-5	0	0
7	-96	SibO	1	0	3	13	-4	0	0
7	-96	SibO	1	0	4	14	-3	0	0
7	-96	SibO	1	0	6	15	-2	0	0
7	-96	SibO	1	0	7	16	-1	0	0
7	-96	SibO	1	0	8	17	0	0	0
8	-96	SibY	0	1	6	13	-2	0	0
8	-96	SibY	0	1	7	14	-1	0	0
8	-96	SibY	0	1	8	15	0	0	0

```
* Where plot will go;
%LET filesave=F:\Example Data\Penn State Data\SibOSibY;
* Sort original datafile by ID variables for merging;
PROC SORT DATA=work.BivFamily; BY family phase; RUN;

* Importing fake people data from excel;
PROC IMPORT DATAFILE = "&filesave.\CPlots.xls"
       OUT=work.CPlots DBMS=EXCEL REPLACE; SHEET="Gender"; GETNAMES=YES; RUN;
* Sort fake people by ID variables for merging;
PROC SORT DATA=work.CPlots; BY family phase; RUN;
* Merging original datafile and fake people for prediction;
DATA work.Mergedbiv; MERGE work.BivFamily work.CPlots; BY family phase; RUN;

* Getting model-predicted values;
TITLE1 "Adding Effects of Gender: Self and O-->Y";
PROC MIXED DATA=Mergedbiv COVTEST NOCLPRINT NOITPRINT MAXITER=1000 iC NAMELEN=100 METHOD=ML;
       CLASS family phase DV;
       MODEL risky = dvO dvO*endtime dvO*endtime*endtime dvY dvY*endtime
                      dvO*Oboy dvO*Oboy*endtime dvO*Oboy*endtime*endtime
                       dvY*Yboy dvY*Yboy*endtime dvY*Oboy dvY*Oboy*endtime
                       / NOINT SOLUTION DDFM=Satterthwaite OUTPM=work.Gen_Plots;
       RANDOM dvO dvY dvO*endtime dvY*endtime / G GCORR TYPE=UN SUBJECT=family; * Level 2 family;
       REPEATED DV / R RCORR TYPE=UN SUBJECT=phase*family;         * Level 1 crossed time*DV; RUN;
* Renaming predicted outcome, keeping only fake people;
DATA work.Gen_Plots; SET work.Gen_Plots; WHERE Order IS NOT MISSING; run;
PROC SORT DATA=work.Gen_Plots; BY Order; run;

* Setting generic plotting options for text, size, and joining lines;
GOPTIONS NOBORDER FTEXT=Triplex FTITLE=Triplex VSIZE=7in HSIZE=7in INTERPOL=JOIN
GSFNAME=outgraph DEV=BMP GSFMODE=replace; run;
* Plot for Risky Behaviors by Sibling Gender;
FILENAME outgraph "&filesave.\Risky by Gender.bmp";
TITLE1 JUSTIFY=CENTER HEIGHT=1.6 "Risky Behaviors of Older and Younger Siblings";
TITLE2 JUSTIFY=CENTER HEIGHT=1.6 "Effect of Sibling Sex";
PROC GPLOT DATA=work.Gen_Plots;
        *Settings for Y-axis;
        AXIS1 LABEL=(HEIGHT=1.3 ANGLE=90 "Risky Behaviors") LENGTH=4.7in ORDER=(15 TO 35 BY 5);
        *Settings for X-axis;
        AXIS2 LABEL=(HEIGHT=1.3 "Age") LENGTH=6.0in ORDER=(12 TO 17 BY 1);
        *Setting for legend - place at top, no frame around it, 2 values across;
        *Labels for series in order of Order;
        LEGEND1 NOFRAME POSITION=(CENTER TOP) ACROSS=2 LABEL=NONE VALUE=(HEIGHT=1.5 COLOR=Black
        "Boy, Younger Brother"  "Boy, Older Brother"
        "Boy, Younger Sister"   "Girl, Older Brother"
        "Girl, Younger Brother" "Boy, Older Sister"
        "Girl, Younger Sister"  "Girl, Older Sister") ;
        SYMBOL1 COLOR=Blue   VALUE=> HEIGHT=2 LINE=1 WIDTH=2;  * > is boy symbol;
        SYMBOL2 COLOR=Blue   VALUE=> HEIGHT=2 LINE=3 WIDTH=2;  * line=3 is dashed;
        SYMBOL3 COLOR=Blue   VALUE=* HEIGHT=2 LINE=1 WIDTH=2;  * * is girl symbol;
        SYMBOL4 COLOR=Red    VALUE=> HEIGHT=2 LINE=3 WIDTH=2;  * line=3 is dashed;
        SYMBOL5 COLOR=Red    VALUE=> HEIGHT=2 LINE=1 WIDTH=2;
        SYMBOL6 COLOR=Blue   VALUE=* HEIGHT=2 LINE=3 WIDTH=2;
        SYMBOL7 COLOR=Red    VALUE=* HEIGHT=2 LINE=1 WIDTH=2;
        SYMBOL8 COLOR=Red    VALUE=* HEIGHT=2 LINE=3 WIDTH=2;
        PLOT pred*age=Order / VAXIS=AXIS1 HAXIS=AXIS2 LEGEND=LEGEND1;
RUN; QUIT;
```



Solution for Fixed Effects - Revisited

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Sib O growth for girls					
dvO	28.2084	0.8537	194	33.04	<.0001
dvO*endtime	2.3724	0.2826	747	8.40	<.0001
dvO*endtime*endtime	0.1580	0.04543	901	3.48	0.0005
Own gender on Sib O					
dvO*Oboy	1.5007	1.2186	190	1.23	0.2197
dvO*endtime*Oboy	-0.7656	0.3955	717	-1.94	0.0533
dvO*endtime*endtime*Oboy	-0.1281	0.06536	906	-1.96	0.0504
Sib O growth: girls w/ SibO=G					
dvY	25.4502	0.9984	183	25.49	<.0001
endtime*dvY	1.4204	0.2972	59.3	4.78	<.0001
Own gender on Sib Y					
dvY*Yboy	-0.9314	1.1041	166	-0.84	0.4001
endtime*dvY*Yboy	-1.0286	0.3571	100	-2.88	0.0049
Sib O gender on Sib Y					
dvY*Oboy	2.6318	1.1648	170	2.26	0.0251
endtime*dvY*Oboy	0.7201	0.3710	138	1.94	0.0543

Example Results Section: The extent to which gender predicted risky behavior in older and younger siblings was examined in a multivariate multilevel model in which time and sibling were crossed and nested within families. Based on the design of the study, exact age at each occasion was centered at age 17 for older siblings and age 15 for younger siblings. Preliminary analyses suggested a fixed quadratic, random linear age model was the best unconditional growth model for older siblings, whereas a random linear age model was best for younger siblings. The intercepts were significantly related across siblings (r = .43), indicating that in families in which older siblings engaged in more risky behavior at age 17 than their peers, younger siblings were also more likely to engage in more risky behavior at age 15 than their peers. Similarly, the linear age slopes were significantly related across siblings (r = .42), indicating that older siblings who increased in risky behavior more across adolescence than their peers were more likely to have younger siblings who did the same. However, there was no significant relationship among the time-specific residuals, indicating that after controlling for growth, on occasions where older siblings were engaging in more risky behavior than predicted, their younger siblings were not significantly more likely to do so as well. The effects of one's own gender and the gender of the older sibling on the younger sibling were then examined. As seen in Table 1 and Figure 1, although there was no significant effect of gender for the older siblings at age 17, older girls had marginally greater linear and quadratic rates of increase across age (i.e., greater acceleration). Similarly, although there was no significant effect of gender for the younger siblings at age 15, younger girls had significantly greater linear rates of increase across age. Finally, having an older brother was related to significantly greater risky behavior for the younger sibling at age 15, and a marginally greater linear rate of increase across age.