Example 7: Multivariate Longitudinal Models for Older and Younger Sibling Risky Behavior
(complete syntax and output available for SAS—and SPSS with slightly different results—electronically)

These data come from the Family Relations Project, a study of family dynamics conducted by HDFS researchers at Penn State University. Data were collected from families with older and younger siblings separated by approximately 2 years of age. The outcome for this example is a measure of risky behavior, as measured annually from age 12–17 for the older sibling and concurrent age 13–15 for the younger sibling.

The code below creates a single risky behavior outcome across siblings (with indicators to keep track of who is who). It also creates a single continuous age variable centered at the last measurement occasion (age 17 for Sib O and 15 for Sib Y). 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 will now examine the extent to which intercepts, slopes, and residuals are related across time of study, and how gender moderates these effects.

SAS Code for Data Manipulation:

```
* SAS code to read data into work library and double-stack into multivariate;
DATA work.Example7; SET example.ICPSR_Example7;
       risky=orisknew; DV="SibO"; dvO=1; dvY=0; age=sibOage; endtime=age-17; boy=Oboy; OUTPUT;
       risky=yrisknew; DV="SibY"; dvO=0; dvY=1; age=sibYage; endtime=age-15; boy=Yboy; OUTPUT;
       LABEL  risky=       "risky: Risky Behavior for All Siblings"
              DV=          "DV: Categorical Sibling Indicator"
              dvO=         "dvO: Dummy Indicator for Older Sibling (ages 12-17)"
              dvY=         "dvY: Dummy Indicator for Younger Sibling (ages 13-15)"
              age=         "age: Time-Varying Age for All Siblings"
              endtime=     "endtime: Age centered at 17=O, 15=Y)"
              boy=         "boy: Gender (0=Girl, 1=Boy) for All Siblings";
RUN; * Sort double-stacked data;
PROC SORT DATA=work.Example7; BY family wave DV; RUN;
```

SPSS Code for Data Manipulation:

```
* SPSS code to double-stack into multivariate.
VARSTOCASES
  /MAKE risky "risky: Risky Behavior for All Siblings" FROM orisknew yrisknew
  /MAKE age "age: Time-Varying Age for All Siblings" FROM sibOage TO sibYage
  /MAKE boy "boy: Gender (0=Girl, 1=Boy) for All Siblings" FROM Oboy Yboy
  /INDEX =DV "DV: Categorical Sibling Indicator"
  /NULL = KEEP
  /KEEP = ALL.
 EXECUTE.
* Make other predictors.
DO IF (DV=1).
 COMPUTE endtime=age-17.
 COMPUTE dvO=1.
 COMPUTE dvY=0.
END IF.
DO IF (DV=2).
 COMPUTE endtime=age-15.
 COMPUTE dvO=0.
 COMPUTE dvY=1.
END IF.
VARIABLE LABELS
  dvO        "dvO: Dummy Indicator for Older Sibling (ages 12-17)"
  dvY        "dvY: Dummy Indicator for Younger Sibling (ages 13-15)"
  age        "age: Time-Varying Age for All Siblings"
  endtime    "endtime: Age centered at 17=O, 15=Y)".
EXECUTE.
* Sort double-stacked data.
SORT CASES BY family wave DV.
```

Model 1. Unconditional Multivariate Longitudinal Model for Risky Behavior (Direct Effects)

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

Level 2:

$$\beta_{0fO} = \gamma_{00O} + U_{0fO}$$
$$\beta_{1fO} = \gamma_{10O} + U_{1fO}$$
$$\beta_{2fO} = \gamma_{20O}$$

Model for older sibling

$$\beta_{0fY} = \gamma_{00Y} + U_{0fY}$$
$$\beta_{1fY} = \gamma_{10Y} + U_{1fY}$$
$$\beta_{2fY} = \gamma_{20Y}$$

Model for younger sibling

```
TITLE1 "SAS Unconditional Multivariate Longitudinal Model for Risky Behavior";
TITLE2 "Using direct effects of linear and quadratic age per sibling";
PROC MIXED DATA=work.Example7 COVTEST NOCLPRINT NOITPRINT IC NAMELEN=100 METHOD=ML;
  CLASS family wave DV;
  MODEL risky = dvO dvO*endtime dvO*endtime*endtime dvY dvY*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=4 RCORR=4 TYPE=UN SUBJECT=wave*family;        * Level 1 crossed time*DV;
  * Testing differences in effects;
  ESTIMATE "Intercept Difference at Age 17/15?"      dvO -1 dvY 1;
  ESTIMATE "Linear Slope Difference at Age 17/15?"   dvO*endtime -1 dvY*endtime 1;
  ESTIMATE "Quadratic Slope Difference?"             dvO*endtime*endtime -1 dvY*endtime*endtime 1;
RUN; TITLE1; TITLE2;


TITLE "SPSS Unconditional Multivariate Longitudinal Model for Risky Behavior".
SUBTITLE "Using direct effects of linear and quadratic age per sibling".
MIXED risky BY family wave DV WITH dvO dvY endtime
   /METHOD = ML
   /PRINT  = SOLUTION TESTCOV G R
   /FIXED  = dvO dvO*endtime dvO*endtime*endtime dvY dvY*endtime dvY*endtime*endtime | NOINT
   /RANDOM = dvO dvY dvO*endtime dvY*endtime | SUBJECT(family) COVTYPE(UN)
   /REPEATED = DV | SUBJECT(wave*family) COVTYPE(UN)
   /TEST = "Intercept Difference at Age 17/15?"      dvO -1 dvY 1
   /TEST = "Linear Slope Difference at Age 17/15?"   dvO*endtime -1 dvY*endtime 1
   /TEST = "Quadratic Slope Difference?"             dvO*endtime*endtime -1 dvY*endtime*endtime 1.
TITLE.
SUBTITLE.
```

SAS output:

Estimated R Matrix for FAMILY*wave 1 6

Row	Col1	Col2
1	13.9353	-0.7855
2	-0.7855	14.1744

Residual variance and covariance across sibs

Estimated R Correlation Matrix for FAMILY*wave 1 6

Row	Col1	Col2
1	1.0000	-0.05589
2	**-0.05589**	1.0000

Residual correlation across sibs

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	endtime*dvY	1	4.6851	10.2330	0.7835	2.0165

Random effects variance and covariance across sibs

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	endtime*dvY	1	0.4098	1.0000	**0.4126**	1.0000

Random effects correlations across sibs

NOTE: Estimated G matrix is not positive definite.

This note tells us that there is something wrong with the **G** matrix: it appears to be $r = 1.0$ at 4,2.

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z 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**	**Linear slope covariance**
UN(4,4)	FAMILY	2.0165	0.7060	2.86	0.0021	
UN(1,1)	FAMILY*wave	13.9353	0.7586	18.37	<.0001	
UN(2,1)	FAMILY*wave	-0.7855	0.7540	-1.04	**0.2975**	**Residual covariance**
UN(2,2)	FAMILY*wave	14.1744	1.2213	11.61	<.0001	

Only **G** and **R** covariances are directly tested for significance.

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

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

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept Difference at Age 17/15?	-2.4749	0.6863	169	-3.61	0.0004
Linear Slope Difference at Age 17/15?	-0.4458	0.3677	358	-1.21	0.2261
Quadratic Slope Difference?	-0.03864	0.08750	162	-0.44	0.6594

Because the siblings have different age ranges, we will leave the growth terms separate for each Sib.
We will remove the nonsignificant quadratic effect of time for Sib Y (not surprising given only 3 occasions).

Model 2. Examining own and unidirectional effects of sibling sex (O→Y) on growth in risky behavior

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

Level 2:

$$\beta_{0fO} = \gamma_{00O} + \gamma_{01O}(OlderBoy_f) + U_{0fO}$$
$$\beta_{1fO} = \gamma_{10O} + \gamma_{11O}(OlderBoy_f) + U_{1fO}$$
$$\beta_{2fO} = \gamma_{20O} + \gamma_{21O}(OlderBoy_f)$$

Model for older sibling

$$\beta_{0fY} = \gamma_{00Y} + \gamma_{01Y}(OlderBoy_f) + \gamma_{02Y}(YoungerBoy_f) + U_{0fY}$$
$$\beta_{1fY} = \gamma_{10Y} + \gamma_{11Y}(OlderBoy_f) + \gamma_{12Y}(YoungerBoy_f) + U_{1fY}$$

Model for younger sibling

```
TITLE1 "SAS Adding Effects of Gender: Self and O-->Y";
PROC MIXED DATA=work.Example7 COVTEST NOCLPRINT NOITPRINT IC NAMELEN=100 METHOD=ML;
  CLASS family wave DV;
  MODEL risky = dvO dvY dvO*endtime dvY*endtime dvO*endtime*endtime
                dvO*Oboy dvY*Yboy dvO*Oboy*endtime dvY*Yboy*endtime dvO*Oboy*endtime*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=4 RCORR=4 TYPE=UN SUBJECT=wave*family;          * Level 1 crossed time*DV;
  ESTIMATE "Diff Effect of Own Gender on Intercept?"        dvO*Oboy -1 dvY*Yboy 1;
  ESTIMATE "Diff Effect of Own Gender on Linear Slope?"     dvO*Oboy*endtime -1 dvY*Yboy*endtime 1;
  ESTIMATE "Diff Effect of SibO Gender on Intercept?"       dvO*Oboy -1 dvY*Oboy 1;
  ESTIMATE "Diff Effect of SibO Gender on Intercept?"       dvO*Oboy*endtime -1 dvY*Oboy*endtime 1;
RUN; TITLE1;
```

```
TITLE " SPSS Adding Effects of Gender: Self and O-->Y".
MIXED risky BY family wave DV WITH dvO dvY endtime Oboy Yboy
   /METHOD = ML
   /PRINT  = SOLUTION TESTCOV G R
   /FIXED  = dvO dvY dvO*endtime dvY*endtime dvO*endtime*endtime
             dvO*Oboy dvY*Yboy dvO*Oboy*endtime dvY*Yboy*endtime dvO*Oboy*endtime*endtime
             dvY*Oboy dvY*Oboy*endtime | NOINT
   /RANDOM = dvO dvY dvO*endtime dvY*endtime | SUBJECT(family) COVTYPE(UN)
   /REPEATED = DV | SUBJECT(wave*family) COVTYPE(UN)
   /TEST = "Diff Effect of Own Gender on Intercept?"        dvO*Oboy -1 dvY*Yboy 1
   /TEST = "Diff Effect of Own Gender on Linear Slope?"     dvO*Oboy*endtime -1 dvY*Yboy*endtime 1
   /TEST = "Diff Effect of SibO Gender on Intercept?"       dvO*Oboy -1 dvY*Oboy 1
   /TEST = "Diff Effect of SibO Gender on Intercept?"       dvO*Oboy*endtime -1 dvY*Oboy*endtime 1.
TITLE.
```

SAS output:

```
    Estimated R Matrix
    for FAMILY*wave 1 6
Row        Col1        Col2
  1     13.8872     -1.0834
  2     -1.0834     13.9522
```

Residual variance and covariance across sibs after controlling for gender

```
  Estimated R Correlation
Matrix for FAMILY*wave 1 6
Row        Col1        Col2
  1      1.0000    -0.07784
  2   -0.07784      1.0000
```

Residual correlation across sibs after controlling for gender

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	dvY*endtime	1	4.2930	10.2545	0.7672	1.9310

Random effects variance and covariance across sibs after controlling for gender

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	dvY*endtime	1	0.3849	1.0000	**0.4117**	1.0000

Random effects correlations across sibs after controlling for gender

NOTE: Estimated G matrix is not positive definite.

This note tells us that there is something wrong with the **G** matrix: it appears to be $r = 1.0$ at 4,2.

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z	
UN(1,1)	FAMILY	64.4248	7.4492	8.65	<.0001	
UN(2,1)	FAMILY	22.1749	5.3428	4.15	**<.0001**	**Intercept covariance**
UN(2,2)	FAMILY	49.3236	6.7433	7.31	<.0001	
UN(3,1)	FAMILY	10.4961	1.3293	7.90	<.0001	
UN(3,2)	FAMILY	3.6434	0.9720	3.75	0.0002	
UN(3,3)	FAMILY	1.7983	0.2571	6.99	<.0001	
UN(4,1)	FAMILY	4.2930	1.7129	2.51	0.0122	
UN(4,2)	FAMILY	10.2545	2.0239	5.07	<.0001	
UN(4,3)	FAMILY	0.7672	0.3085	2.49	**0.0129**	**Linear slope covariance**
UN(4,4)	FAMILY	1.9310	0.6570	2.94	0.0016	
UN(1,1)	FAMILY*wave	13.8872	0.7585	18.31	<.0001	
UN(2,1)	FAMILY*wave	-1.0834	0.7357	-1.47	**0.1408**	**Residual covariance**
UN(2,2)	FAMILY*wave	13.9522	1.1564	12.07	<.0001	

Only **G** and **R** covariances are directly tested for significance.

Information Criteria

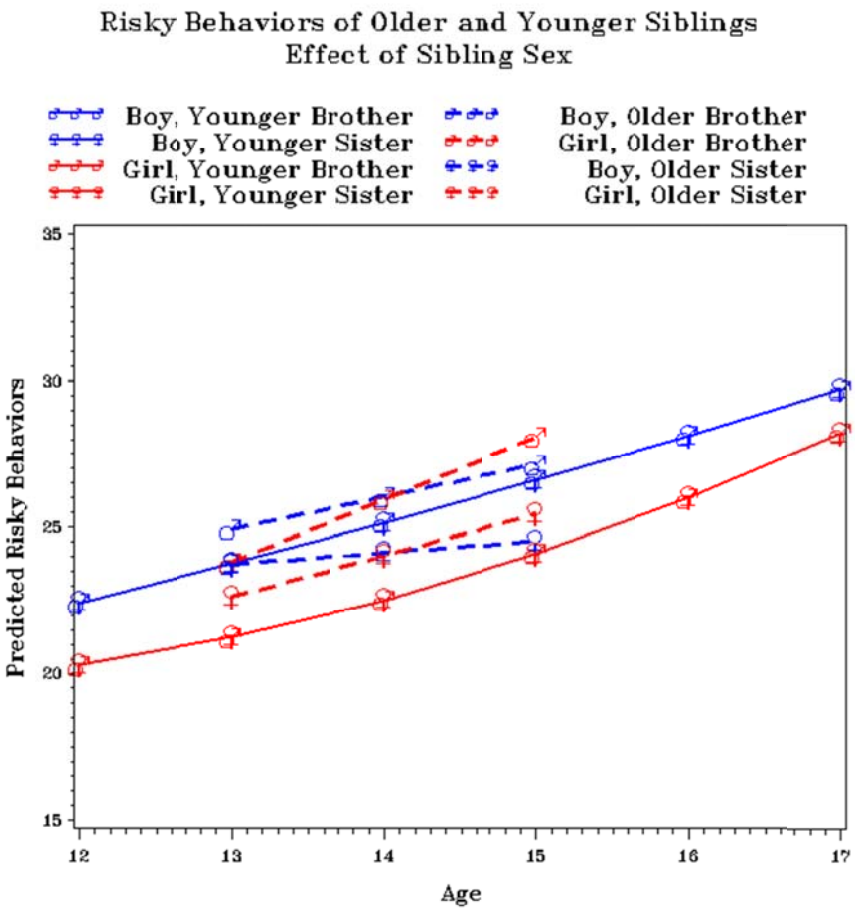
Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
9766.6	25	9816.6	9817.4	9850.0	9899.1	9924.1

Solution for Fixed Effects (re-arranged for convenience)

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	
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	25.4502	0.9984	183	25.49	<.0001	Sib Y growth: girls w/ SibO=G
dvY*endtime	1.4204	0.2972	59.3	4.78	<.0001	
dvY*Yboy	-0.9314	1.1041	166	-0.84	0.4001	Own gender on Sib O
dvY*endtime*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
dvY*endtime*Oboy	0.7201	0.3710	138	1.94	0.0543	

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
Diff Effect of Own Gender on Intercept?	-2.4321	1.6444	356	-1.48	0.1400
Diff Effect of Own Gender on Linear Slope?	-0.2630	0.5326	403	-0.49	0.6217
Diff Effect of SibO Gender on Intercept?	1.1311	1.3962	173	0.81	0.4189
Diff Effect of SibO Gender on Linear Slope?	1.4857	0.5166	341	2.88	0.0043



dvO	28.2084
dvO*endtime	2.3724
dvO*endtime*endtime	0.1580
dvO*Oboy	1.5007
dvO*endtime*Oboy	-0.7656
dvO*endtime*endtime*Oboy	-0.1281
dvY	25.4502
dvY*endtime	1.4204
dvY*Yboy	-0.9314
dvY*endtime*Yboy	-1.0286
dvY*Oboy	2.6318
dvY*endtime*Oboy	0.7201

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.