

**Example 5: Multivariate General Linear Models for Family (Triadic) Data**  
**Part 1 using Univariate Software: STATA MIXED, R GLS, and SAS MIXED**  
**Part 2 using Path Analysis Software: Mplus, STATA SEM, and R LAVAAN**  
*(complete syntax and output available for STATA, R, and SAS electronically)*

These data were collected as part of a study of family dynamics conducted at Penn State University. The sample for this example includes 140 families with data from three family members (as three multivariate outcomes): a mother, a father, and an adult child. The example outcome is a scale mean (range from 1–4) of attitudes about gender roles in marriage, in which higher scores indicate more conservative attitudes. The example predictors are the gender of the adult child (0=girl, 1=boy) and the years of education for each family member (centered such that 0=12 years). In all models, we will use an unstructured **R** matrix (in which the residual variances and covariances are estimated separately for each outcome), although compound symmetry heterogeneous (with equal correlation across outcomes, but a separate variance for each outcome) or compound symmetry (equal covariance and equal variance across all outcomes) would be more parsimonious alternatives (if they fit not worse than unstructured via a likelihood ratio test).

We will predict all three family member outcomes simultaneously using two distinct analysis frameworks. In Part 1 we will estimate multivariate general linear models within univariate software (i.e., with an identity link and conditional multivariate normal distributions) using residual maximum likelihood (REML), and we will (try to) test fixed effects using Satterthwaite denominator degrees of freedom. In Part 2, we will estimate the same models using path analysis, (a truly multivariate modeling framework in which multiple columns can be predicted at once), whose software requires us to switch to maximum likelihood and to test fixed effects without denominator degrees of freedom. I am using manual dummy codes to distinguish the three outcomes rather than treating them as factor variables (i.e., letting the program create contrasts to do so), given that the latter option is not as readily available for path analysis.

The marginal outcome distributions of the showed some positive skew (with an observed floor effect for the adult children), but a conditional normal distribution appears to be a reasonable choice among the readily-available options for multivariate models. This is evidenced in the final model by predicted outcomes that stayed within the outcome bounds without the use a link function to do so, and plausible homogeneity of variance across predicted outcomes. In Part 2, we will also invoke robust standard errors that protect against deviations from residual multivariate normality.

**Part 1 will require “reshaping” (i.e., stacking) our original data stored in wide (multivariate) format, in which one row holds all variables per family, with per-person versions in separate columns...**

	FamilyID: Family ID Number	KidBoy: Kid's Gender (0=girl, 1=boy)	KidEd12: Kid's Years of Education (0=12)	MomEd12: Mother's Years of Education (0=12)	DadEd12: Father's Years of Education (0=12)	KidMarital: Kid's Marital Gender Attitudes Mean (1-4)	MomMarital: Mom's Marital Gender Attitudes Mean (1-4)	DadMarital: Dad's Marital Gender Attitudes Mean (1-4)
1	3996	1	2	2	2	1	1.833333333	1
2	4425	1	3	0	0	1	1.333333333	2.5

...into this new format called long (stacked, univariate), with one row per person per family:

	FamilyID: Family ID Number	KidBoy: Kid's Gender (0=girl, 1=boy)	KidEd12: Kid's Years of Education (0=12)	MomEd12: Mother's Years of Education	DadEd12: Father's Years of Education (0=12)	DV: 1K,2M,3D	kid: Is Adult Child (0=no, 1=yes)	mom: Is Mother (0=no, 1=yes)	dad: Is Father (0=no, 1=yes)	marital: Marital Gender Attitudes Mean (1-4)
1	3996	1	2	2	2	1.Kid	1	0	0	1
2	3996	1	2	2	2	2.Mom	0	1	0	1.83333333
3	3996	1	2	2	2	3.Dad	0	0	1	1
4	4425	1	3	0	0	1.Kid	1	0	0	1
5	4425	1	3	0	0	2.Mom	0	1	0	1.33333333
6	4425	1	3	0	0	3.Dad	0	0	1	2.5

**Part 2 will use the original wide-format data for path analysis instead.**

**STATA Syntax for Importing and Stacking Wide Data into Long (to get one row per person per family):**

```
// Paste in the folder address where datafile is saved between quotes
cd "C:\Dropbox\25_PSQF6270\PSQF6270_Example5"

// Import Example 5a wide-format Stata data
use "Example5Wide.dta", clear

// Rename variables with numeric suffix to use with reshape (old) (new)
rename (kidmarital mommarital dadmarital) (marital1 marital2 marital3)

// Stack data: list multivariate variables first, i (higher index) j (repeated)
reshape long marital, i(familyid) j(DVnum)

// Create per-outcome dummy codes
gen kid=0
gen mom=0
gen dad=0
recode kid (0=1) if DVnum==1
recode mom (0=1) if DVnum==2
recode dad (0=1) if DVnum==3

// Label new variables
label variable DVnum "DVnum: 1K,2M,3D"
label variable kid "kid: Is Adult Child (0=no, 1=yes)"
label variable mom "mom: Is Mother (0=no, 1=yes)"
label variable dad "dad: Is Father (0=no, 1=yes)"
label variable marital "marital: Marital Gender Attitudes Mean (1-4)"

// Remove missing predictors or row-specific outcome (will happen anyway)
egen nummiss = rowmiss(kidboy kided12 momed12 daded12 marital)
drop if nummiss>0
```

**R Syntax for Importing and Stacking Wide-Format Data into Long-Format (to get one row per person per family), after loading packages *haven*, *TeachingDemos*, *psych*, *multcomp*, *prediction*, *nlme*, and *lavaan*, as shown online:**

```
# Set working directory (to import and export files to)
# Paste in the folder address where dataset is saved in quotes
setwd("C:/Dropbox/25_PSQF6270/PSQF6270_Example5")

# Import Example 5 wide-format SAS data
Example5_wide = read_sas(data_file="Example5Wide.sas7bdat")
# Convert to data frame without labels to use for analysis
Example5_wide = as.data.frame(Example5_wide)

# Stack into long format (one row per outcome per family)
Example5 = reshape(Example5_wide, direction="long", idvar="FamilyID",
  varying=c("KidMarital", "MomMarital", "DadMarital"),
  v.names="marital", timevar="DVnum", times=c(1,2,3))

# Create per-person dummy codes
Example5$kid=0
Example5$mom=0
Example5$dad=0
Example5$kid[which(Example5$DVnum==1)]=1
Example5$mom[which(Example5$DVnum==2)]=1
Example5$dad[which(Example5$DVnum==3)]=1

# Remove missing predictors or row-specific outcome (will happen anyway)
Example5 = Example5[complete.cases(Example5[,
  c("KidBoy", "KidEd12", "MomEd12", "DadEd12", "marital")]),]
```

## Part 1: Multivariate General Linear Models via Univariate Software

### Model 0a: Empty Means, Unstructured Variance Model Predicting Marital Conservative Gender Attitudes

*General Intercept Version:*  $\widehat{Marital}_{fi} = \beta_{00} + \beta_{01}(Kid_{fi}) + \beta_{02}(Mom_{fi})$

#### STATA Syntax and Partial Output for Model 0a:

```
display "STATA Empty Means, Unstructured Variance Models for Marital Attitudes"
display "STATA Model 0a: General Intercept (Dad=Ref DV) using 2 Dummy Codes"
mixed marital c.kid c.mom,          /// Fixed intercept will be for dad (as omitted DV)
|| familyid: , noconstant          /// This NOCONSTANT removes default family random intercept
nolog reml residuals(unstructured,t(DVnum))          /// Unstructured R matrix by DV
difficult dfmethod(satterthwaite) dftable(pvalue)    /// Use Satterthwaite denominator DF
display "-2LL= " e(ll)*-2          /// Print -2LL for model
lincom _cons*1 + c.kid*1, small    /// Kid Intercept (Dad + diff)
lincom _cons*1 + c.mom*1, small    /// Mom Intercept (Dad + diff)
lincom c.kid*-1 + c.mom*1, small    /// Kid vs. Mom: Intercept Diff
```

```
Mixed-effects REML regression          Number of obs      =      420
Group variable: familyid              Number of groups   =      140
                                      Obs per group:
                                      min =          3
                                      avg =         3.0
                                      max =          3
DF method: Satterthwaite              DF:
                                      min =      139.00
                                      avg =      139.00
                                      max =      139.00
                                      F(2,   139.00) =      16.19 → Multiv Wald test with DDF
Log restricted-likelihood = -353.47735  Prob > F          =      0.0000
```

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]
familyid: (empty)   (No random effect variances in this model)			
Residual: Unstructured			
var(e1)	.3311924	.0397272	.2618044 .4189707
var(e2)	.3230136	.0387461	.2553391 .4086242
var(e3)	.3195886	.0383353	.2526318 .4042916
cov(e1,e2)	.041334	.027963	-.0134724 .0961405
cov(e1,e3)	.0824049	.0284663	.026612 .1381978
cov(e2,e3)	.0937145	.0283876	.0380758 .1493531

```
estat wcorrelation, covariance          /// R matrix of variances and covariances across outcomes
estat wcorrelation                     /// RCORR matrix of correlations across outcomes
```

Covariances for familyid = 3996:				Correlations:			
DV	1	2	3	DV	1	2	3
1	0.331			1	1.000		
2	0.041	0.323		2	0.126	1.000	
3	0.082	0.094	0.320	3	0.253	0.292	1.000

R and RCORR from  
estat wcorrelation

SDs for R also printed  
(not shown here)

#### R Syntax and Partial Output for Model 0a:

```
print("R Empty Means, Unstructured Variance Models for Marital Attitudes")
print("R Model 0a: General Intercept (Dad=Ref DV) using 2 Dummy Codes")
Model0a = gls(data=Example5, method="REML",
              model=marital~1+kid+mom,          # Fixed intercept will be for dad (as omitted)
              correlation=corSymm(form=~DVnum|FamilyID), # Unstructured correlations
              weights=varIdent(form=~1|DVnum))    # Separate variance by DV
print("Print -2LL and Results"); -2*logLik(Model0a); summary(Model0a)
```

'log Lik.' 706.95471 (df=9) → -2LL for model

```
Correlation Structure: General
1      2
2 0.126
3 0.253 0.292
```

Inside of RCORR  
(given in full below)

```
Variance function:
  Structure: Different standard deviations per stratum
  Formula: ~1 | DV
  Parameter estimates:
    1.Kid      2.Mom      3.Dad
1.00000000 0.98757870 0.98232045
```

Weird multiplication factors to compute  
SD relative to first DV → ignore this

```
Residual standard error: 0.57549394
Degrees of freedom: 420 total; 417 residual
```

Naïve denominator DF given

```
print("Show R and RCORR matrices for first family in the data")
getVarCov(Model0a, individual="3996"); # R matrix = variances and covariances across outcomes
corMatrix(Model0a$modelStruct$corStruct)[[3]] # 3=rows/columns of R here, RCORR = correlations
```

```
Marginal variance covariance matrix
      [,1]      [,2]      [,3]
[1,] 0.331190 0.041336 0.082407
[2,] 0.041336 0.323020 0.093715
[3,] 0.082407 0.093715 0.319590
```

Actual R matrix!

```
> corMatrix(Model0a$modelStruct$corStruct)[[3]]
      [,1]      [,2]      [,3]
[1,] 1.00000000 0.12637845 0.25329512
[2,] 0.12637845 1.00000000 0.29167759
[3,] 0.25329512 0.29167759 1.00000000
```

Actual RCORR matrix!

```
print("DF=2 Intercept Diff -- Get error that it used Chi-Square instead of F")
F0a = glht(model=Model0a, linfct=rbind(c(0,1,0),c(0,0,1)), df=139)
```

```
Global Test:
  Chisq DF      Pr(>Chisq)
1 32.376  2 0.00000009324
```

R told me it wouldn't compute the F test...  
except it secretly did! So below I just asked for it

```
SaveF0a = summary(F0a, test=Ftest()); SaveF0a # Joint F-test
print("Get and show hidden results for F, dfnum, dfden, and p-value")
SaveF0a$test$fstat; SaveF0a$test$df; SaveF0a$df
[1,] 16.18809      [1,] 2      [1,] 139
```

```
pf(SaveF0a$test$fstat, df1=SaveF0a$test$df, df2=SaveF0a$df, lower.tail=FALSE)
[1,] 0.00000047859907
```

```
# model=marital~1+kid+mom
print("Missing Intercepts and Difference -- Had to give it correct Denominator DF")
summary(glht(model=Model0a, df=139, linfct=rbind(
  "Kid Intercept (Dad+Diff)" = c(1,1,0), # in order of fixed effects
  "Mom Intercept (Dad+Diff)" = c(1,0,1),
  "Kid vs. Mom: Intercept Diff" = c(0,-1,1))), test=adjusted("none"))
```

**Model 0a:**  $\widehat{Marital}_{fi} = \beta_{00} + \beta_{01}(Kid_{fi}) + \beta_{02}(Mom_{fi})$

### Model-Estimated Fixed Effects using General Intercept Version Model 0a (from SAS):

Solution for Fixed Effects						
Effect	Estimate	Standard Error	DF	t Value	Pr >  t	
Intercept	1.9560	0.04778	139	40.94	<.0001	Dad intercept beta00
kid	-0.3264	0.05892	139	-5.54	<.0001	Kid intercept diff beta01
mom	-0.05619	0.05702	139	-0.99	0.3261	Mom intercept diff beta02

### Requested Linear Combination Estimates using General Intercept Version Model 0a (from SAS):

Estimates						
Label	Estimate	Standard Error	DF	t Value	Pr >  t	
Kid Intercept (Dad+diff)	1.6295	0.04864	139	33.50	<.0001	beta00 + beta01
Mom Intercept (Dad+diff)	1.8998	0.04803	139	39.55	<.0001	beta00 + beta02
Kid vs. Mom: Intercept Diff	0.2702	0.06389	139	4.23	<.0001	beta02 - beta01

**Model 0b: Empty Means, Unstructured Variance Model for Marital Conservative Gender Attitudes**

*DV-Specific Intercept Version:*  $\widehat{Marital}_{fi} = \beta_{00}(Dad_{fi}) + \beta_{01}(Kid_{fi}) + \beta_{02}(Mom_{fi})$

**STATA Syntax for Model 0b:**

```
display "STATA Model 0b: DV-Specific Intercepts using All 3 Dummy Codes"
mixed marital c.kid c.mom c.dad, noconstant /// This NOCONSTANT removes general fixed intercept
|| familyid: , noconstant /// This NOCONSTANT removes family random intercept
nolog reml residuals(unstructured,t(DVnum)) /// Unstructured R matrix by DV
difficult dfmethod(satterthwaite) dftable(pvalue) // Use Satterthwaite denominator DF
display "-2LL= " e(11)*-2 // Print -2LL for model
estat wcorrelation, covariance // R matrix
estat wcorrelation // RCORR matrix
test (c.kid=c.mom) (c.kid=c.dad), small // DF=2 Intercept Diff (small = use denominator DF)
lincom c.kid*-1 + c.mom*1, small // Kid vs. Mom: Intercept Diff
lincom c.kid*-1 + c.dad*1, small // Kid vs. Dad: Intercept Diff
lincom c.mom*-1 + c.dad*1, small // Mom vs. Dad: Intercept Diff
```

**R Syntax for Model 0b:**

```
print("R Model 0b: DV-Specific Intercepts using All 3 Dummy Codes")
Model0b = gls(data=Example5, method="REML",
              model=marital~0+kid+mom+dad, # 0 removes fixed intercept
              correlation=corSymm(form=~DVnum|FamilyID), # Unstructured correlations
              weights=varIdent(form=~1|DVnum)) # Separate variance by DV
print("Print -2LL and Results"); -2*logLik(Model0b); summary(Model0b)
print("Show R and RCORR matrices for first family in the data")
getVarCov(Model0b, individual="3996") # R matrix = variances and covariances across outcomes
corMatrix(Model0b$modelStruct$corStruct)[[3]] # 3=rows/columns of R here, RCORR = correlations

print("DF=2 Intercept Diff -- Get error that it used Chi-Square instead of F")
F0c = glht(model=Model0b, linfct=rbind(c(-1,1,0),c(0,-1,1)), df=139)
SaveF0c = summary(F0c, test=Ftest()); SaveF0a # Joint F-test
print("Get and show hidden results for F, dfnum, dfden, and p-value")
SaveF0c$test$fstat; SaveF0c$test$df; SaveF0c$df
pf(SaveF0c$test$fstat,df1=SaveF0c$test$df,df2=SaveF0c$df,lower.tail=FALSE)

print("Pairwise Intercept Diffs -- Had to give it correct Denominator DF")
summary(glht(model=Model0b, df=139, linfct=rbind(
  "Kid vs. Mom: Intercept Diff" = c(-1,1,0), # in order of fixed effects
  "Kid vs. Dad: Intercept Diff" = c(-1,0,1),
  "Mom vs. Dad: Intercept Diff" = c(0,-1,1))), test=adjusted("none"))
```

**Model 0b:**  $\widehat{Marital}_{fi} = \beta_{00}(Dad_{fi}) + \beta_{01}(Kid_{fi}) + \beta_{02}(Mom_{fi})$

**Model-Estimated Fixed Effects using DV-Specific Intercept Version from SAS:**

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr >  t	
kid	1.6295	0.04864	139	33.50	<.0001	Kid intercept beta01
mom	1.8998	0.04803	139	39.55	<.0001	Mom intercept beta02
dad	1.9560	0.04778	139	40.94	<.0001	Dad intercept beta00

**Requested Linear Combination Estimates using DV-Specific Intercept Version from SAS:**

Label	Estimates					
	Estimate	Standard Error	DF	t Value	Pr >  t	
Kid vs. Mom: Intercept Diff	0.2702	0.06389	139	4.23	<.0001	beta02 - beta01
Kid vs. Dad: Intercept Diff	0.3264	0.05892	139	5.54	<.0001	beta00 - beta01
Mom vs. Dad: Intercept Diff	0.05619	0.05702	139	0.99	0.3261	beta00 - beta02

To avoid confusion, we will proceed using Model 0b: DV-specific intercepts implemented via three dummy codes. This approach also aligns most directly with path model variants of these models (Part 2).

**Model 1: DV-Specific Intercepts: To what extent does the kid's gender predict each persons' attitude?**

$$\widehat{Marital}_{fi} = \beta_{00}(Dad_{fi}) + \beta_{01}(Kid_{fi}) + \beta_{02}(Mom_{fi}) \\ + \beta_{10}(Dad_{fi})(KidBoy_f) + \beta_{11}(Kid_{fi})(KidBoy_f) + \beta_{12}(Mom_{fi})(KidBoy_f)$$

**STATA Syntax for Model 1:**

```
display "STATA Model 1: DV-Specific Intercepts -- Add Kid Gender"
mixed marital c.kid c.mom c.dad c.kid#c.kidboy c.mom#c.kidboy c.dad#c.kidboy, noconstant ///
|| familyid: , noconstant /// This NOCONSTANT removes family random intercept
nolog reml residuals(unstructured,t(DVnum)) /// Unstructured R matrix by DV
difficult dfmethod(satterthwaite) dftable(pvalue) // Use Satterthwaite denominator DF
display "-2LL= " e(11)*-2 // Print -2LL for model
estat wcorrelation, covariance // R matrix
estat wcorrelation // RCORR matrix

// DF=2 Diff in Kidboy Slope
test (c.kid#c.kidboy=c.mom#c.kidboy) (c.kid#c.kidboy=c.dad#c.kidboy), small
lincom c.kid#c.kidboy*-1 + c.mom#c.kidboy*1, small // Kid vs. Mom: Kidboy Slope Diff
lincom c.kid#c.kidboy*-1 + c.dad#c.kidboy*1, small // Kid vs. Dad: Kidboy Slope Diff
lincom c.mom#c.kidboy*-1 + c.dad#c.kidboy*1, small // Mom vs. Dad: Kidboy Slope Diff
lincom 0.5*(c.mom#c.kidboy*1 + c.dad#c.kidboy*1), small // Parent: Kidboy Slope
// Mom vs. Dad: Kidboy Slope Diff
lincom 0.5*(c.kid#c.kidboy*-2 + c.mom#c.kidboy*1 + c.dad#c.kidboy*1), small
```

**R Syntax for Model 1:**

```
print("R Model 1: DV-Specific Intercepts -- Add Kid Gender")
Modell = gls(data=Example5, method="REML",
             model=marital~0+kid+mom+dad+kid:KidBoy+mom:KidBoy+dad:KidBoy,
             correlation=corSymm(form=~DVnum|FamilyID), # Unstructured correlations
             weights=varIdent(form=~1|DVnum)) # Separate variance by DV
print("Print -2LL and Results"); -2*logLik(Modell); summary(Modell)

print("Show R and RCORR matrices for first family in the data")
getVarCov(Modell, individual="3996"); corMatrix(Modell$modelStruct$corStruct)[[3]]

print("DF=2 Diff in KidBoy Slope -- Get error that it used Chi-Square instead of F")
F1 = glht(model=Modell, linfct=rbind(c(0,0,0,-1,1,0),c(0,0,0,-1,0,1)), df=138)
SaveF1 = summary(F1, test=Ftest()); SaveF0a # Joint F-test
print("Get and show hidden results for F, dfnum, dfden, and p-value")
SaveF1$test$fstat; SaveF1$test$df; SaveF1$df
pf(SaveF1$test$fstat,df1=SaveF1$test$df,df2=SaveF1$df,lower.tail=FALSE)

print("KidBoy Slope Diffs -- Had to give it correct Denominator DF")
summary(glht(model=Modell, df=138, linfct=rbind(
  "Kid vs. Mom: KidBoy Slope Diff" = c(0,0,0,-1, 1, 0), # in order of fixed effects
  "Kid vs. Dad: KidBoy Slope Diff" = c(0,0,0,-1, 0, 1),
  "Mom vs. Dad: KidBoy Slope Diff" = c(0,0,0, 0, -1, 1),
  "Parent KidBoy Effect" = c(0,0,0, 0,1/2,1/2),
  "Kids vs. Parent KidBoy Effect Diff" = c(0,0,0,-1,1/2,1/2))), test=adjusted("none"))
```

**Partial SAS Output for Model 1: DV-Specific Intercepts adding Kid's Gender as Predictor for Each Attitude**

Estimated R Matrix for FAMILYID 3996				Estimated R Correlation Matrix for FAMILYID 3996			
Row	Col1	Col2	Col3	Row	Col1	Col2	Col3
1	0.3136	0.03725	0.07733	1	1.0000	0.1168	0.2440
2	0.03725	0.3244	0.09315	2	0.1168	1.0000	0.2890
3	0.07733	0.09315	0.3203	3	0.2440	0.2890	1.0000

Contrasts				
Label	Num DF	Den DF	F Value	Pr > F
DF=2 Diff in KidBoy Slope?	2	138	1.90	0.1529



## Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr >  t	
kid	1.4950	0.06554	138	22.81	<.0001	Kid intercept beta01
mom	1.8703	0.06666	138	28.06	<.0001	Mom intercept beta02
dad	1.9178	0.06624	138	28.95	<.0001	Dad intercept beta00
<b>kid*KidBoy</b>	<b>0.2811</b>	<b>0.09474</b>	<b>138</b>	<b>2.97</b>	<b>0.0035</b>	<b>girl vs boy for Kid beta11</b>
mom*KidBoy	0.06152	0.09636	138	0.64	0.5242	girl vs boy for Mom beta12
dad*KidBoy	0.07970	0.09575	138	0.83	0.4066	girl vs boy for Dad beta10

## Estimates

Label	Estimate	Standard Error	DF	t Value	Pr >  t	b = beta
Kid vs. Mom: KidBoy Slope Diff	-0.2196	0.1270	138	-1.73	0.0860	b12 - b11
Kid vs. Dad: KidBoy Slope Diff	-0.2014	0.1171	138	-1.72	0.0877	b10 - b11
Mom vs. Dad: KidBoy Slope Diff	0.01818	0.1145	138	0.16	0.8741	b10 - b12
Parent KidBoy Slope	0.07061	0.07711	138	0.92	0.3614	0.5*(b10+b12)
Kid vs. Parents: KidBoy Slope Diff	-0.2105	0.1079	138	-1.95	0.0531	0.5*(b10+b12) - b11

It looks like we need to control for the effect of kid gender only for the kid (which makes sense, since we don't know about the gender of any siblings). Next, we'll test the effects of each person's education on their own attitude, followed by the incremental effect of dad's education on kid and mom attitudes after controlling for own education.

**Model 2: DV-Specific Intercepts: To what extent does one's own education predict one's own attitude?**

$$\widehat{Marital}_{fi} = \beta_{00}(Dad_{fi}) + \beta_{01}(Kid_{fi}) + \beta_{02}(Mom_{fi}) + \beta_{11}(Kid_{fi})(KidBoy_f) \\ + \beta_{20}(Dad_{fi})(DadEd_f - 12) + \beta_{31}(Kid_{fi})(KidEd_f - 12) + \beta_{42}(Mom_{fi})(MomEd_f - 12)$$

**STATA Syntax for Model 2:**

```
display "STATA Model 2: DV-Specific Intercepts -- KidBoy on Kid Only, Add Own Education"
mixed marital c.kid c.mom c.dad c.kid#c.kidboy ///
      c.kid#c.kided12 c.mom#c.momed12 c.dad#c.daded12, noconstant ///
|| familyid: , noconstant /// This NOCONSTANT removes family random intercept
nolog reml residuals(unstructured,t(DVnum)) /// Unstructured R matrix by DV
difficult dfmethod(satterthwaite) dftable(pvalue) // Use Satterthwaite denominator DF
display "-2LL= " e(11)*-2 // Print -2LL for model
estat wcorrelation, covariance // R matrix
estat wcorrelation // RCORR matrix
```

**R Syntax for Model 2:**

```
print("R Model 2: DV-Specific Intercepts -- KidBoy on Kid Only, Add Own Educ")
Model2 = gls(data=Example5, method="REML",
      model=marital~0+kid+mom+dad +kid:KidBoy +kid:KidEd12+mom:MomEd12+dad:DadEd12,
      correlation=corSymm(form=~DVnum|FamilyID), # Unstructured correlations
      weights=varIdent(form=~1|DVnum)) # Separate variance by DV
print("Print -2LL and Results"); -2*logLik(Model2); summary(Model2)

print("Show R and RCORR matrices for first family in the data")
getVarCov(Model2, individual="3996"); corMatrix(Model2$modelStruct$corStruct)[[3]]
```

**Partial SAS Output for Model 2:**

Estimated R Matrix for FAMILYID 3996				Estimated R Correlation Matrix for FAMILYID 3996			
Row	Col1	Col2	Col3	Row	Col1	Col2	Col3
1	0.3156	0.03837	0.07669	1	1.0000	0.1207	0.2501
2	0.03837	0.3205	0.08441	2	0.1207	1.0000	0.2732
3	0.07669	0.08441	0.2979	3	0.2501	0.2732	1.0000

Solution for Fixed Effects						
Effect	Estimate	Standard Error	DF	t Value	Pr >  t	
kid	1.5117	0.09814	141	15.40	<.0001	Kid intercept beta01
mom	1.9359	0.05976	142	32.39	<.0001	Mom intercept beta02
dad	2.0700	0.05663	145	36.55	<.0001	Dad intercept beta00
kid*KidBoy	0.2641	0.09204	137	2.87	0.0048	girl vs boy for Kid beta11
kid*KidEd12	-0.00280	0.02344	138	-0.12	0.9052	Kid Ed for kid beta31
mom*MomEd12	-0.01725	0.01711	142	-1.01	0.3150	Mom Ed for mom beta42
dad*DadEd12	-0.05447	0.01570	143	-3.47	0.0007	Dad Ed for dad beta20

### Model 3: DV-Specific Intercepts: To what extent does dad's education also predict kid and mom attitudes?

$$\begin{aligned} \widehat{Marital}_{fi} = & \beta_{00}(Dad_{fi}) + \beta_{01}(Kid_{fi}) + \beta_{02}(Mom_{fi}) + \beta_{11}(Kid_{fi})(KidBoy_f) \\ & + \beta_{20}(Dad_{fi})(DadEd_f - 12) + \beta_{31}(Kid_{fi})(KidEd_f - 12) + \beta_{42}(Mom_{fi})(MomEd_f - 12) \\ & + \beta_{21}(Kid_{fi})(DadEd_f - 12) + \beta_{22}(Mom_{fi})(DadEd_f - 12) \end{aligned}$$

### STATA Syntax for Model 3:

```
display "STATA Model 3: DV-Specific Intercepts -- Add Dad Educ (Control for Own Educ)"
mixed marital c.kid c.mom c.dad c.kid#c.kidboy c.kid#c.kided12 c.mom#c.momed12 ///
  c.dad#c.daded12 c.kid#c.daded12 c.mom#c.daded12, noconstant ///
  || familyid: , noconstant /// This NOCONSTANT removes family random intercept
  nolog reml residuals(unstructured,t(DVnum)) /// Unstructured R matrix by DV
  difficult dfmethod(satterthwaite) dftable(pvalue) // Use Satterthwaite denominator DF
display "-2LL= " e(l1)*-2 // Print -2LL for model
estat wcorrelation, covariance // R matrix
estat wcorrelation // RCORR matrix
lincom c.kid#c.daded12*-1 + c.mom#c.daded12*1, small // Kid vs. Mom: DadEd12 Slope Diff
lincom c.kid#c.daded12*-1 + c.dad#c.daded12*1, small // Kid vs. Dad: DadEd12 Slope Diff
lincom c.mom#c.daded12*-1 + c.dad#c.daded12*1, small // Mom vs. Dad: DadEd12 Slope Diff
predict Model3pred, xb // Save yhat from fixed effects
predict Model3res, rstandard // Save "standardized" residuals from fixed effects
hist Model3res // Histogram of residuals (for normality)
graph export "STATA Model 3 Residual Histogram.png", replace
twoway (scatter Model3res Model3pred) // Scatterplot by predicted (for constant variance)
graph export "STATA Model 3 Residual Scatterplot.png", replace
```

### R Syntax for Model 3:

```
print("R Model 3: DV-Specific Intercepts -- Add Dad Educ (Control for Own Educ)")
Model3 = gls(data=Example5, method="REML",
  model=marital~0+kid+mom+dad+ kid:KidBoy +kid:KidEd12+mom:MomEd12+dad:DadEd12
  +kid:DadEd12+mom:DadEd12,
  correlation=corSymm(form=~DVnum|FamilyID), # Unstructured correlations
  weights=varIdent(form=~1|DVnum)) # Separate variance by DV
print("Print -2LL and Results"); -2*logLik(Model3); summary(Model3)
print("Show R and RCORR matrices for first family in the data")
getVarCov(Model3, individual="3996"); corMatrix(Model3$modelStruct$corStruct)[[3]]

print("DadEd Slope Diffs -- Had to give it correct Denominator DF")
summary(glht(model=Model3, df=136, linfct=rbind(
  "Kid vs. Mom: DadEd12 Slope Diff" = c(0,0,0,0,0,0,0,-1,1), # in order of fixed effects
  "Kid vs. Dad: DadEd12 Slope Diff" = c(0,0,0,0,0,0,0,1,-1,0),
  "Mom vs. Dad: DadEd12 Slope Diff" = c(0,0,0,0,0,0,0,1,0,-1))) , test=adjusted("none"))

print("Save yhat from fixed effects and Pearson residuals")
Example5$Model3pred = predict(Model3, type="response")
Example5$Model3res = residuals(Model3, type="pearson")
print("Histogram of Residuals for normality")
hist(x=Example5$Model3res, freq=FALSE, ylab="Density",xlab="Model 3 Residuals")
print("Scatterplot of residuals by predicted for constant variance")
plot(x=Example5$Model3res, y=Example5$Model3pred,
  ylab="Residual",xlab="Model 3 Predicted Outcome")
```



**Partial SAS Output for Model 3:**

$$\begin{aligned} \widehat{Marital}_{fi} = & \beta_{00}(Dad_{fi}) + \beta_{01}(Kid_{fi}) + \beta_{02}(Mom_{fi}) + \beta_{11}(Kid_{fi})(KidBoy_f) \\ & + \beta_{20}(Dad_{fi})(DadEd_f - 12) + \beta_{31}(Kid_{fi})(KidEd_f - 12) + \beta_{42}(Mom_{fi})(MomEd_f - 12) \\ & + \beta_{21}(Kid_{fi})(DadEd_f - 12) + \beta_{22}(Mom_{fi})(DadEd_f - 12) \end{aligned}$$

Estimated R Matrix for FAMILYID 3996				Estimated R Correlation Matrix for FAMILYID 3996			
Row	Col1	Col2	Col3	Row	Col1	Col2	Col3
1	0.3179	0.03856	0.07720	1	1.0000	0.1204	0.2508
2	0.03856	0.3229	0.08514	2	0.1204	1.0000	0.2744
3	0.07720	0.08514	0.2982	3	0.2508	0.2744	1.0000

Solution for Fixed Effects						
Effect	Estimate	Standard Error	DF	t Value	Pr >  t	
kid	1.5123	0.1003	140	15.08	<.0001	Kid intercept beta01
mom	1.9373	0.06305	138	30.73	<.0001	Mom intercept beta02
dad	2.0707	0.05769	138	35.89	<.0001	Dad intercept beta00
kid*KidBoy	0.2639	0.09258	136	2.85	0.0050	girl vs boy for Kid beta11
kid*KidEd12	-0.00264	0.02458	136	-0.11	0.9147	Kid Ed for kid beta31
mom*MomEd12	-0.01624	0.02068	137	-0.79	0.4338	Mom Ed for mom beta42
dad*DadEd12	-0.05484	0.01654	138	-3.32	0.0012	Dad Ed for dad beta20
kid*DadEd12	-0.00048	0.01791	138	-0.03	0.9787	Dad Ed for kid beta21
mom*DadEd12	-0.00169	0.02069	141	-0.08	0.9349	Dad Ed for mom beta22

Estimates							
Label	Estimate	Standard Error	DF	t Value	Pr >  t	b = beta	
Kid vs. Mom: DadEd12 Slope Diff	-0.00121	0.02601	162	-0.05	0.9629	b22 - b21	
Kid vs. Dad: DadEd12 Slope Diff	-0.05436	0.02127	143	-2.56	0.0117	b20 - b21	
Mom vs. Dad: DadEd12 Slope Diff	-0.05314	0.02335	154	-2.28	0.0242	b20 - b22	

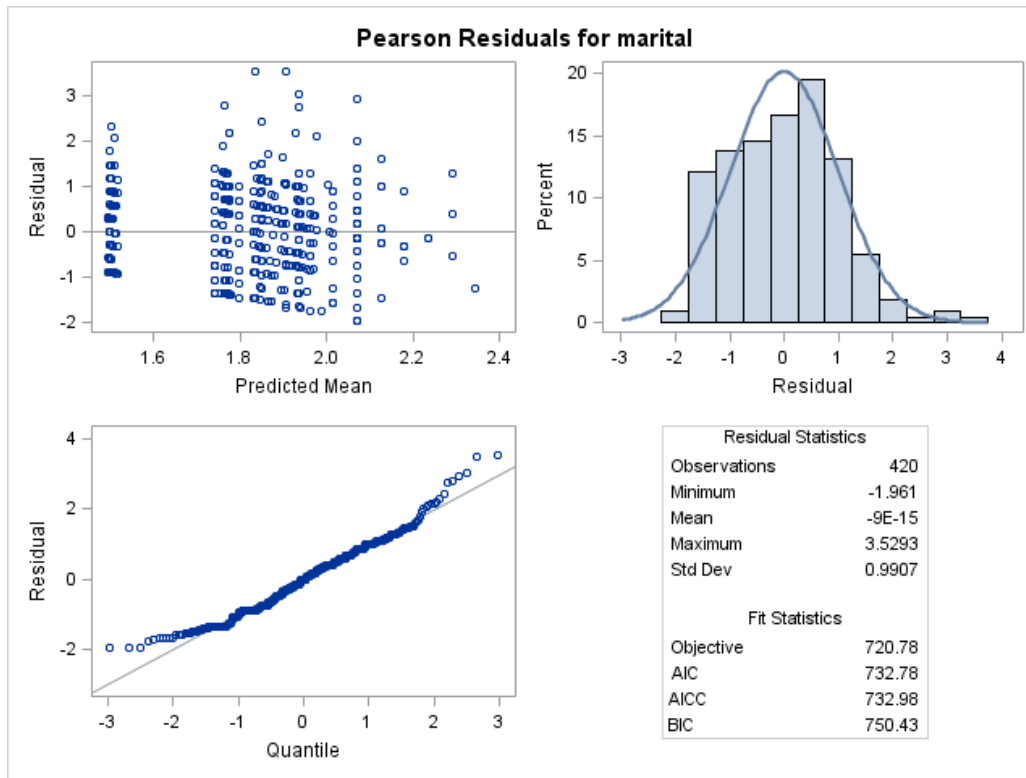
Moral of the story? Multivariate models can be estimated in univariate software to capture the relationships between person-specific predictors and person-specific outcomes (such as in “actor–partner” models for dyadic data as well).

**Example results section [using SAS] for Part 1 Models 0–3:**

The extent to which gender and education predicted marital attitudes was examined in 140 families, in which responses were collected from adult children, their mothers, and their fathers. Higher outcomes indicated more conservative marital attitudes (i.e., gender-traditional attitudes measured as the mean across items on a scale of 1 to 4). Given that the outcomes were correlated within families, multivariate general linear models (i.e., with conditionally multivariate normal residuals) were used to predict all three outcomes for each family simultaneously. All models were estimated using residual maximum likelihood and Satterthwaite denominator degrees of freedom. All models allowed separate means and residual variances across the three outcomes for the three types of family members, as well as covariances among the residuals from the same family. ESTIMATE statements were used to estimate simple slopes and simple slope differences as linear combinations of the model fixed effects. Prior to adding predictors, an empty means model (i.e., an unconditional model with no predictors) revealed significant differences in marital attitudes across type of family member,  $F(2, 139) = 16.19$ ,  $p < .001$ . Although mean attitudes were similar across mothers and fathers (1.90 and 1.96, respectively,  $p = .27$ ), the mean attitudes of children (1.63) were significantly less conservative on average than those of their parents ( $p < .001$  for both comparisons).

To begin, we examined the extent to which the gender of the adult child (coded 0=woman, 1=man) who was surveyed was related to the marital attitudes of each type of family member. Although the attitudes of adult male children were significantly more conservative than those of adult female children (diff = 0.28,  $p = .004$ ), there were no significant effects of the gender of the adult child for the marital attitudes of their mothers or fathers. Thus, we retained a predictor for the gender of the adult child only for the adult child’s outcome. We then examined the extent to which the education (centered at 12 years) of each type of family member predicted their own attitudes, which was significant only for the father: for every additional year of father’s education, his own attitudes were expected to be less conservative by 0.05 ( $p < .001$ ). Next, we examined whether father’s education incrementally predicted the marital attitudes of the mother or adult child after controlling for their own education, but neither effect was significant (and the effect of father’s education on his own attitudes was significantly larger).

**But how do we know if Model 3 is sufficient??** One aspect concerns the fit of the conditional distribution—in absence of Pearson  $\chi^2/DF$  for normal residuals, we can examine residual plots, such as shown for SAS below:



These plots suggest some deviation from normality of the residuals, although the assumption of constant variance looks not terribly unreasonable.

Unfortunately, multivariate options for generalized linear models do not include beta-binomial alternatives that might have been useful here (given that the outcomes are bounded by 1 and 4). Also, given that all predicted outcomes stayed in bounds, it appears we don't necessarily need a link function.

Instead, we can see how the results differ using "robust" standard errors...so stay tuned for Part 2!

The other issue whether all relationships among the predictors and outcomes have been captured adequately by the model... for a more efficient way to answer that question, **stay tuned for Part 2 using path analysis!**

## Part 2: Multivariate General Linear Models via Path Analysis Software

In Part 2, we begin by estimating Model 3 using path analysis in Mplus, STATA SEM, and R LAVAAN, which each require us to switch to maximum likelihood and test fixed effects without denominator degrees of freedom. For Model 4, we will also invoke “robust” standard errors (that correct for deviations from multivariate non-normality).

### STATA Syntax to prepare wide-format data file in .csv format for Mplus:

```
// Import Example 5 wide-format STATA data
use "Example5Wide.dta", clear

// Example of how to export a .csv file for use in Mplus
// Replace all missing values with -999 for Mplus
mvencode _all, mv(-999)

// export delimited below: using lists the path and name of the new .csv file
// replace means it will be replaced if a file already exists with that name
// delimiter indicates a comma-delimited file
// nolabel will save actual data (numbers) instead of any value labels included
// novarnames tells it not to write the names to the top of the .csv file
export delimited using "Example5Wide_STATA.csv", ///
    delimiter(",") replace nolabel novarnames
```

### R Syntax to prepare wide-format data file in .csv format for Mplus:

```
# Example of how to export a .csv file for use in Mplus
# Copy data, replace all missing values with -999 for Mplus
Example5_Mplus = Example5_wide
Example5_Mplus[is.na(Example5_Mplus)] <- -999

# Write to .csv file without column names
write.table(x=Example5_Mplus, col.names=FALSE, row.names=FALSE, sep="," ,
    file="Example5Wide_R.csv")
```

---

### Model 3: DV-Specific Intercepts: To what extent does dad's education also predict kid and mom attitudes?

$$\begin{aligned} \widehat{Marital}_{fi} = & \beta_{00}(Dad_{fi}) + \beta_{01}(Kid_{fi}) + \beta_{02}(Mom_{fi}) + \beta_{11}(Kid_{fi})(KidBoy_f) \\ & + \beta_{20}(Dad_{fi})(DadEd_f - 12) + \beta_{31}(Kid_{fi})(KidEd_f - 12) + \beta_{42}(Mom_{fi})(MomEd_f - 12) \\ & + \beta_{21}(Kid_{fi})(DadEd_f - 12) + \beta_{22}(Mom_{fi})(DadEd_f - 12) \end{aligned}$$

### STATA Syntax and Output for Previous Model 3 as a Path Model (estimated with ML; regular SEs):

```
// Import Example 5 wide STATA data
use "Example5Wide.dta", clear

* /// means continue the command + comment
* // means comment only

display "STATA Model 3: Own Education + Dad Education a Predictor of Each Attitude"
display "Using SEM to create path analysis model estimated with ML on wide-format data"
sem
    (kidmarit mommarit dadmarit <- _cons)          /// All intercepts estimated (by default)
    (kidmarit <- kidboy kided12)                  /// Regressions: y outcomes ON x predictors
    (mommarit <- momed12)                          ///
    (kidmarit mommarit dadmarit <- daded12),       ///
    var(e.kidmarit e.mommarit e.dadmarit)         /// All residual variances estimated (by default)
    covariance(e.kidmarit*e.mommarit              /// All pairwise residual covariances (not default)
               e.mommarit*e.dadmarit              ///
               e.kidmarit*e.dadmarit)              ///
    method(mlmv)                                  /// Full-information ML
    lincom _b[mommarital:daded12] - _b[kidmarital:daded12] // Kid v. Mom: Dad Educ Effect Diff
    lincom _b[dadmarital:daded12] - _b[kidmarital:daded12] // Kid v. Dad: Dad Educ Effect Diff
    lincom _b[dadmarital:daded12] - _b[mommarital:daded12] // Mom v. Dad: Dad Educ Effect Diff
    sem, coeflegend                                /// Print parameter labels, too (to use in lincom)
```

```

sem, standardized // Print fully standardized solution, too
estat gof, stats(all) // Print fit statistics
display "LL for H1 Model= " e(critvalue_s)
display "# of parameters= " e(df_m)
display "-2LL= " e(ll)*-2 // Print -2LL for model
estat eqgof // Print R2 per variable
estat residuals // Print how far off each predicted covariance is
estat mindices, minchi2(3.84) showpclass(all) // Print cheat codes to improve model fit p<.05

```

```

Structural equation model      Number of obs      =      140
Estimation method      =      mlmv
Log likelihood      =      -1374.4822 → Does NOT match Mplus because all predictors are in the likelihood,
                                not just the outcomes, but rest of the fit tests do match

```

UNSTANDARDIZED SOLUTION		OIM						IN MIXED
		Coef.	Std. Err.	z	P> z	[95% Conf. Interval]		
Structural								
kidmarital <-		These unstandardized <- paths are the fixed slopes in MIXED.						
	kidboy	.2638938	.0914365	2.89	0.004	.0846816	.4431059	B11
	kided12	-.002641	.0242338	-0.11	0.913	-.0501385	.0448565	B31
	daded12	-.0004795	.0176566	-0.03	0.978	-.0350857	.0341268	B21
	_cons	1.512271	.0989087	15.29	0.000	1.318414	1.706129	B01
mommarital <-								
	momed12	-.0162593	.0211854	-0.77	0.443	-.0577819	.0252634	B42
	daded12	-.0016793	.0206962	-0.08	0.935	-.0422431	.0388845	B22
	_cons	1.937305	.062596	30.95	0.000	1.814619	2.059991	B02
dadmarital <-								
	daded12	-.0548368	.016422	-3.34	0.001	-.0870233	-.0226502	B20
	_cons	2.070718	.0572756	36.15	0.000	1.95846	2.182976	B00

Below are the residual variances and covariances from the **R** matrix in MIXED.

var(e.kidmarital)	.3091381	.0369567			.2445646	.3907613	UN(1,1)
var(e.mommarital)	.3161529	.0379111			.2499347	.3999152	UN(2,2)
var(e.dadmarital)	.2938981	.0351275			.2325192	.3714795	UN(3,3)
cov(e.kidmarital,e.mommarital)	.0380059	.0266924	1.42	0.154	-.0143102	.090322	UN(2,1)
cov(e.kidmarital,e.dadmarital)	.0761007	.0263037	2.89	0.004	.0245463	.1276551	UN(3,1)
cov(e.mommarital,e.dadmarital)	.0839167	.0273732	3.07	0.002	.0302662	.1375671	UN(3,2)

LR test of model vs. saturated: chi2(6) = 10.93, Prob > chi2 = 0.0906

```

. lincom _b[mommarital:daded12] - _b[kidmarital:daded12] // Kid v. Mom: Dad Educ Effect Diff
( 1) - [kidmarital]daded12 + [mommarital]daded12 = 0

```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
(1)	-.0011998	.0258607	-0.05	0.963	-.0518858 .0494862	B22 - B21

```

. lincom _b[dadmarital:daded12] - _b[kidmarital:daded12] // Kid v. Dad: Dad Educ Effect Diff
( 1) - [kidmarital]daded12 + [dadmarital]daded12 = 0

```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
(1)	-.0543573	.0210185	-2.59	0.010	-.0955527 -.0131618	B20 - B21

```

. lincom _b[dadmarital:daded12] - _b[mommarital:daded12] // Mom v. Dad: Dad Educ Effect Diff
( 1) - [mommarital]daded12 + [dadmarital]daded12 = 0

```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
(1)	-.0531575	.023324	-2.28	0.023	-.0988717 -.0074432	B20 - B22

```
. sem, coeflegend // Print parameter labels, too (to use in lincom)
```

		Coef.	Legend
Structural			
kidmarital <-			
	kidboy	.2638938	_b[kidmarital:kidboy]
	kided12	-.002641	_b[kidmarital:kided12]
	daded12	-.0004795	_b[kidmarital:daded12]
	_cons	1.512271	_b[kidmarital:_cons]
mommarital <-			
	momed12	-.0162593	_b[mommarital:momed12]
	daded12	-.0016793	_b[mommarital:daded12]
	_cons	1.937305	_b[mommarital:_cons]
dadmarital <-			
	daded12	-.0548368	_b[dadmarital:daded12]
	_cons	2.070718	_b[dadmarital:_cons]
	var(e.kidmarital)	.3091381	_b[var(e.kidmarital):_cons]
	var(e.mommarital)	.3161529	_b[var(e.mommarital):_cons]
	var(e.dadmarital)	.2938981	_b[var(e.dadmarital):_cons]
	cov(e.kidmarital,e.mommarital)	.0380059	_b[cov(e.kidmarital,e.mommarital):_cons]
	cov(e.kidmarital,e.dadmarital)	.0761007	_b[cov(e.kidmarital,e.dadmarital):_cons]
	cov(e.mommarital,e.dadmarital)	.0839167	_b[cov(e.mommarital,e.dadmarital):_cons]

This table from `sem, coeflegend` provides the parameter names for the LINCOS statements above.

```
LR test of model vs. saturated: chi2(6) = 10.93, Prob > chi2 = 0.0906
```

```
. sem, standardized // Print fully standardized solution, too
```

Standardized Solution: All variables M=0, SD=1		OIM											
		Coef.	Std. Err.	z	P> z	[95% Conf. Interval]							
Structural		These standardized <- paths are standardized regression coefficients.											
kidmarital <-													
kidboy								.2306503	.0770785	2.99	0.003	.0795792	.3817214
kided12								-.0090898	.0834042	-0.11	0.913	-.1725591	.1543794
daded12								-.0023406	.0861915	-0.03	0.978	-.1712728	.1665916
	_cons		2.645964	.2466043	10.73	0.000	2.162628	3.129299					
mommarital <-													
	momed12		-.0782303	.1019932	-0.77	0.443	-.2781333	.1216728					
	daded12		-.0083038	.1023122	-0.08	0.935	-.208832	.1922244					
	_cons		3.433518	.2280257	15.06	0.000	2.986595	3.88044					
dadmarital <-													
	daded12		-.2716069	.0768234	-3.54	0.000	-.422178	-.1210357					
	_cons		3.676054	.2200636	16.70	0.000	3.244738	4.107371					
var(e.kidmarital)			.9463698	.0355442			.8792068	1.018663					
var(e.mommarital)			.9930701	.0139521			.9660976	1.020796					
var(e.dadmarital)			.9262297	.0417315			.8479448	1.011742					
		These standardized covariances are residual correlations (in RCORR).											
cov(e.kidmarital,e.mommarital)			.12157	.0835428	1.46	0.146	-.0421709	.2853109					
cov(e.kidmarital,e.dadmarital)			.2524724	.0792209	3.19	0.001	.0972022	.4077426					
cov(e.mommarital,e.dadmarital)			.2752969	.0801933	3.43	0.001	.1181209	.432473					

These standardized covariances are residual correlations (in RCORR).

```
. estat gof, stats(all) // Print fit statistics
```

Fit statistic	Value	Description (from STATA!)	Notes from Lesa:
Likelihood ratio			
chi2_ms(6)	10.929	model vs. saturated	-This is -2ΔLL for our H0-H1
p > chi2	0.091		Test of exact fit: NS is good!
chi2_bs(15)	52.998	baseline vs. saturated	-This is -2ΔLL for H0-H1 if
p > chi2	0.000		H0 had no paths at all



Population error			
RMSEA	0.077	Root mean squared error of approximation	Should be < .08 or so
90% CI, lower bound	0.000		
upper bound	0.148		
pclose	0.229	Probability RMSEA <= 0.05	Test of exact fit: NS is good!
-----			
Information criteria			
AIC	2778.964	Akaike's information criterion	Does not match Mplus
BIC	2823.089	Bayesian information criterion	Does not match Mplus
-----			
Baseline comparison			
CFI	0.870	Comparative fit index	Should be > .9 or so
TLI	0.676	Tucker-Lewis index	Should be > .9 or so
-----			
Size of residuals			
SRMR	0.039	Standardized root mean squared residual	Should be < .05 or so
CD	0.132	Coefficient of determination	Like an overall R2 across DVs

```
. estat eggof // Print R2 per variable
Equation-level goodness of fit
```

		Variance				
depvars	fitted	predicted	residual	R-squared	mc	mc2
-----						
observed						
kidmarital	.3266568	.0175187	.3091381	.0536302	.231582	.0536302
mommarital	.3183591	.0022062	.3161529	.0069299	.0832462	.0069299
dadmarital	.3173058	.0234077	.2938981	.0737703	.2716069	.0737703
-----						
overall				.1323532		
-----						

mc = correlation between depvar and its prediction  
mc2 = mc^2 is the Bentler-Raykov squared multiple correlation coefficient

```
. estat residuals // Print how far off each predicted covariance is
Residuals of observed variables
```

Mean residuals	kidmari~1	mommari~1	dadmari~1	kidboy	kided12	momed12	daded12
-----							
raw	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-----							

Above: the means are recovered perfectly because each outcome has its own intercept (and predictor means are not part of the model). Below: the bolded covariances indicate the biggest sources of misfit—it looks like momed12 needs to predict each outcome!

Covariance residuals	kidmari~1	mommari~1	dadmari~1	kidboy	kided12	momed12	daded12
-----							
kidmarital	0.002						
mommarital	0.003	0.002					
dadmarital	0.004	0.005	0.000				
kidboy	0.004	0.015	0.014	0.000			
kided12	-0.001	0.016	-0.008	0.000	0.000		
momed12	<b>0.068</b>	<b>-0.072</b>	<b>-0.280</b>	0.000	0.000	0.000	
daded12	-0.000	-0.000	-0.000	0.000	0.000	0.000	0.000
-----							

```
. estat mindices, minchi2(3.84) showpclass(all) // Print cheat codes to improve model fit at p<.05
Modification indices
```

	MI	df	P>MI	EPC	Standard EPC	
-----						
Structural						
dadmarital <-						
mommarital	9.061	1	0.00	3.68633	3.692443	This is already in the model as a cov
momed12	9.061	1	0.00	-.0599371	-.2888608	<b>This is MomEd → DadMarit</b>
-----						

EPC = expected parameter change

**R Syntax for Previous Model 3 as a Path Model (estimated with ML; regular SEs):**

```

print("R Model 3: Own Education + Dad Education a Predictor of Each Attitude")
# Create model syntax as separate text object
Syntax3 = "
# Residual variances estimated separately (by default)
  KidMarital ~~ KidMarital; MomMarital ~~ MomMarital; DadMarital ~~ DadMarital
# All possible pairwise residual covariances (not estimated by default)
  KidMarital ~~ MomMarital + DadMarital; MomMarital ~~ DadMarital
# All intercepts estimated separately (by default)
  KidMarital ~ 1; MomMarital ~ 1; DadMarital ~ 1
# Regressions: y outcomes ON x predictors (label to do math on later)
  KidMarital ~ KidBoy + KidEd12
  MomMarital ~ MomEd12
  KidMarital ~ (DadEd2K)*DadEd12
  MomMarital ~ (DadEd2M)*DadEd12
  DadMarital ~ (DadEd2D)*DadEd12
# Getting differences in effect of DadEd for each person
  KvMDadEd := DadEd2M - DadEd2K # Kid v. Mom: Dad Educ Effect Diff
  KvDDadEd := DadEd2D - DadEd2K # Kid v. Dad: Dad Educ Effect Diff
  MvDDadEd := DadEd2D - DadEd2M # Mom v. Dad: Dad Educ Effect Diff
"
print("lavaan path analysis model estimated with ML on wide-format data")
PathModel3 = lavaan(data=Example5_wide, model=Syntax3, estimator="MLR", mimic="mplus")
summary(PathModel3, fit.measures=TRUE, rsquare=TRUE, standardized=TRUE, ci=TRUE)

print("Request sorted modification indices for p<.05 to troubleshoot local misfit")
modindices(object=PathModel3, sort=TRUE, minimum.value=3.84)
print("Request residual covariance matrix = leftover from observed minus predicted")
resid(object=PathModel3, type="raw") # also type="cor" for correlation matrix

```

**Mplus Syntax and Output for Previous Model 3 as a Path Model (estimated with ML; regular SEs):**

```

!!!!!!!!!!!!!!!!!!!!!!!!!!!! Mplus Lines Cannot Exceed 90 Characters !!!!!!!!!!!!!!!!!!!!!!!
TITLE: Example 5 Model 3: Own Education + Dad Education a Predictor of Each Attitude

DATA:      FILE = Example5Wide.csv;      ! Can just list file name if input is in same folder
          FORMAT = free;                  ! FREE (default) or FIXED format
          TYPE = individual;              ! Individual (default) or matrix data as input
VARIABLE:
! List of ALL variables in original wide data file, in order;
! Mplus names must use 8 characters or fewer (so rename as needed);
  NAMES = FamilyID KidBoy KidEd12 MomEd12 DadEd12 KidMarit MomMarit DadMarit;
! List of ALL variables used in model;
  USEVARIABLES = KidBoy KidEd12 MomEd12 DadEd12 KidMarit MomMarit DadMarit;
! Missing data codes (here, -999);
  MISSING = ALL (-999);

ANALYSIS:   TYPE = GENERAL;                ! Used for path models
           ESTIMATOR = ML;                 ! Full-information regular maximum likelihood

OUTPUT:     CINTERVAL;                     ! Print confidence intervals
           STDYX;                          ! Print fully standardized solution, too
           RESIDUAL;                       ! Print how far off each predicted covariance is
           MODINDICES (3.84);              ! Print cheat codes to improve our model fit at p<.05

MODEL: ! * Indicates estimated parameter (all listed below for clarity)

! All residual variances estimated separately (by default)
  KidMarit* MomMarit* DadMarit*;

! All possible pairwise residual covariances (not estimated by default)
  KidMarit MomMarit DadMarit WITH KidMarit* MomMarit* DadMarit*;

! All intercepts estimated separately (by default)
  [KidMarit* MomMarit* DadMarit*];

```

```
! Regressions: y outcomes ON x predictors (label to do math on later)
KidMarit ON KidBoy* KidEd12*;
MomMarit ON MomEd12*;
KidMarit MomMarit DadMarit ON DadEd12* (DadEd2K DadEd2M DadEd2D);
```

```
! Getting differences in effect of DadEd for each person
```

```
MODEL CONSTRAINT:
```

```
NEW (KvMDadEd KvDDadEd MvDDadEd); ! List names of linear combinations here
```

```
KvMDadEd = DadEd2M - DadEd2K; ! Kid v. Mom: Dad Educ Effect Diff
```

```
KvDDadEd = DadEd2D - DadEd2K; ! Kid v. Dad: Dad Educ Effect Diff
```

```
MvDDadEd = DadEd2D - DadEd2M; ! Mom v. Dad: Dad Educ Effect Diff
```

```
THE MODEL ESTIMATION TERMINATED NORMALLY
```

```
MODEL FIT INFORMATION
```

Number of Free Parameters	15	Notes from Lesa:
Loglikelihood		
H0 Value	-337.106	For OUR model: Larger is better
H1 Value	-331.641	For model with all possible paths estimated
Information Criteria		
Akaike (AIC)	704.211	For our model: Smaller is better
Bayesian (BIC)	748.336	
Sample-Size Adjusted BIC	700.878	
(n* = (n + 2) / 24)		
Chi-Square Test of Model Fit		
Value	10.929	This is -2ALL for our H0-H1
Degrees of Freedom	6	This is counting the covariances between X's and Y's too
P-Value	0.0906	Test of exact fit: Nonsignificant is good!
RMSEA (Root Mean Square Error Of Approximation)		
Estimate	0.077	Should be < .08 or so
90 Percent C.I.	0.000 0.148	
Probability RMSEA <= .05	0.229	Test of close fit: Nonsignificant is good!
CFI/TLI		
CFI	0.870	Should be > .9 or so
TLI	0.676	Should be > .9 or so
Chi-Square Test of Model Fit for the Baseline Model		
Value	52.998	This is -2ALL for H0-H1 if H0 had no paths at all
Degrees of Freedom	15	
P-Value	0.0000	
SRMR (Standardized Root Mean Square Residual)		
Value	0.046	Should be < .05 or so

```
MODEL RESULTS (UNSTANDARDIZED SOLUTION; Mplus reorders them to list paths first)
```

	Estimate	S.E.	Est./S.E.	P-Value	Two-Tailed	IN MIXED
KIDMARIT ON						
KIDBOY	0.264	0.091	2.886	0.004	B11	
KIDED12	-0.003	0.024	-0.109	0.913	B31	
DADED12	0.000	0.018	-0.027	0.978	B21	
MOMMARIT ON						
MOMED12	-0.016	0.021	-0.767	0.443	B42	
DADED12	-0.002	0.021	-0.081	0.935	B22	
DADMARIT ON						
DADED12	-0.055	0.016	-3.339	0.001	B20	
KIDMARIT WITH						
MOMMARIT	0.038	0.027	1.424	0.154	UN(2,1)	
DADMARIT	0.076	0.026	2.893	0.004	UN(3,1)	
MOMMARIT WITH						
DADMARIT	0.084	0.027	3.066	0.002	UN(3,2)	
Intercepts						
KIDMARIT	1.512	0.099	15.290	0.000	B01	
MOMMARIT	1.937	0.063	30.949	0.000	B02	
DADMARIT	2.071	0.057	36.154	0.000	B00	
Residual Variances						
KIDMARIT	0.309	0.037	8.365	0.000	UN(1,1)	
MOMMARIT	0.316	0.038	8.339	0.000	UN(2,2)	
DADMARIT	0.294	0.035	8.367	0.000	UN(3,3)	
New/Additional Parameters (FROM MODEL CONSTRAINT, like ESTIMATE or LINCOM)						
KVMDADED	-0.001	0.026	-0.046	0.963	B22 - B21	
KVDDADED	-0.054	0.021	-2.586	0.010	B20 - B21	
MVDDADED	-0.053	0.023	-2.279	0.023	B20 - B22	

These unstandardized ON paths are the fixed slopes from MIXED.

These unstandardized WITH covariances are residual covariances (in R).

Note that because we are using ML, the residual variances are smaller than in MIXED (that used REML instead to avoid this downward bias).

**STANDARDIZED MODEL RESULTS - ALL VARIABLES HAVE MEAN=0, SD=1**

STDYX Standardization

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
KIDMARIT ON				
KIDBOY	0.231	0.078	2.950	0.003
KIDED12	-0.009	0.083	-0.109	0.913
DADED12	-0.002	0.086	-0.027	0.978
MOMMARIT ON				
MOMED12	-0.078	0.102	-0.766	0.444
DADED12	-0.008	0.102	-0.081	0.935
DADMARIT ON				
DADED12	-0.272	0.078	-3.470	0.001
KIDMARIT WITH				
MOMMARIT	0.122	0.084	1.455	0.146
DADMARIT	0.252	0.079	3.187	0.001
MOMMARIT WITH				
DADMARIT	0.275	0.080	3.433	0.001
Intercepts				
KIDMARIT	2.646	0.247	10.723	0.000
MOMMARIT	3.434	0.228	15.057	0.000
DADMARIT	3.676	0.221	16.659	0.000
Residual Variances				
KIDMARIT	0.946	0.036	26.246	0.000
MOMMARIT	0.993	0.014	71.055	0.000
DADMARIT	0.926	0.043	21.782	0.000
R-SQUARE				
Observed				
Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
KIDMARIT	0.054	0.036	1.487	0.137
MOMMARIT	0.007	0.014	0.496	0.620
DADMARIT	0.074	0.043	1.735	0.083

These standardized ON paths are standardized regression coefficients.

These standardized WITH covariances are residual correlations (in RCORR).

**ESTIMATED MODEL AND RESIDUALS (OBSERVED - ESTIMATED)****Residuals for Means**

KIDMARIT	MOMMARIT	DADMARIT	KIDBOY	KIDED12	MOMED12	DADED12
0.000	0.000	0.000	0.000	0.000	0.000	0.000

The means are recovered perfectly because each outcome has its own intercept (and predictor means are not part of the model).

**Residuals for Covariances**

	KIDMARIT	MOMMARIT	DADMARIT	KIDBOY	KIDED12	MOMED12	DADED12
KIDMARIT	0.002						
MOMMARIT	0.003	0.002					
DADMARIT	0.004	0.005	0.000				
KIDBOY	0.004	0.015	0.014	0.000			
KIDED12	-0.001	0.016	-0.008	0.000	0.000		
MOMED12	<b>0.068</b>	<b>-0.072</b>	<b>-0.280</b>	0.000	0.000	0.000	
DADED12	0.000	0.000	0.000	0.000	0.000	0.000	0.000

**After commenting out the MODEL CONSTRAINT code and running it again, we get these “helpful” suggestions for how to improve model fit:**

Minimum M.I. value for printing the modification index 3.840

	M.I.	E.P.C.	Std E.P.C.	StdYX E.P.C.
--	------	--------	------------	--------------

The bolded covariances above indicate the biggest sources of misfit—it looks like momed12 needs to predict each outcome!

**ON Statements**

DADMARIT ON MOMMARIT	9.062	3.687	3.687	3.693	This is already in the model as a cov
DADMARIT ON MOMED12	9.061	-0.060	-0.060	-0.289	This is MomEd → DadMarit

**WITH Statements**

MOMED12 WITH DADMARIT	9.336	-0.294	-0.294	-0.200	This is MomEd ↔ DadMarit
DADED12 WITH DADMARIT	8.134	0.491	0.491	0.324	This is already in the model as a path

**Model 4 in Univariate Software, DV-Specific Intercepts: To what extent does mom's education also predict kid and dad attitudes?** *Uses long-format data, ML, and robust standard errors for multivariate non-normality*

$$\begin{aligned} \widehat{Marital}_{fi} = & \beta_{00}(Dad_{fi}) + \beta_{01}(Kid_{fi}) + \beta_{02}(Mom_{fi}) + \beta_{11}(Kid_{fi})(KidBoy_f) \\ & + \beta_{20}(Dad_{fi})(DadEd_f - 12) + \beta_{31}(Kid_{fi})(KidEd_f - 12) + \beta_{42}(Mom_{fi})(MomEd_f - 12) \\ & + \beta_{21}(Kid_{fi})(DadEd_f - 12) + \beta_{22}(Mom_{fi})(DadEd_f - 12) \\ & + \beta_{41}(Kid_{fi})(MomEd_f - 12) + \beta_{40}(Dad_{fi})(MomEd_f - 12) \end{aligned}$$

```

display "STATA Model 4: DV-Specific Intercepts -- Add Mom Educ (Controlling for Own+Dad Educ)"
display "To match path model in Part 2, switch to ML estimation, robust SEs"
display "Satterthwaite DF not allowed with EMPIRICAL, so switch to residual (N-k)"
mixed marital c.kid c.mom c.dad c.kid#c.kidboy c.kid#c.kided12 c.mom#c.momed12 ///
    c.dad#c.daded12 c.kid#c.daded12 c.mom#c.daded12 ///
    c.kid#c.momed12 c.dad#c.momed12, noconstant ///
    || familyid: , noconstant /// This NOCONSTANT removes family random intercept
    nolog mle residuals(unstructured,t(DVnum)) /// Unstructured R matrix by DV
    difficult vce(robust) /// Use robust SEs, so no denominator DF allowed
display "-2LL=" e(ll)*-2 /// Print -2LL for model
estat wcorrelation, covariance /// R matrix
estat wcorrelation /// RCORR matrix
predict pred, xb // Add column pred of predicted outcomes to data
lincom c.kid#c.daded12*-1 + c.mom#c.daded12*1, small // Kid vs. Mom: DadEd12 Slope Diff
lincom c.kid#c.daded12*-1 + c.dad#c.daded12*1, small // Kid vs. Dad: DadEd12 Slope Diff
lincom c.mom#c.daded12*-1 + c.dad#c.daded12*1, small // Mom vs. Dad: DadEd12 Slope Diff
lincom c.kid#c.momed12*-1 + c.mom#c.momed12*1, small // Kid vs. Mom: MomEd12 Slope Diff
lincom c.kid#c.momed12*-1 + c.dad#c.momed12*1, small // Kid vs. Dad: MomEd12 Slope Diff
lincom c.mom#c.momed12*-1 + c.dad#c.momed12*1, small // Mom vs. Dad: MomEd12 Slope Diff
// Get correlation of actual and predicted outcomes to form R2
pwcrr marital pred if DV==1, sig
display "DV=1 Kid R2= " r(rho)^2 // Print R2 relative to empty model
pwcrr marital pred if DV==2, sig
display "DV=2 Mom R2= " r(rho)^2 // Print R2 relative to empty model
pwcrr marital pred if DV==3, sig
display "DV=3 Dad R2= " r(rho)^2 // Print R2 relative to empty model

print("R Model 4: DV-Specific Intercepts -- Add Mom Educ (Controlling for Own+Dad Educ)")
print("To match path model, switch to ML estimation, but robust SEs not directly available")
Model4 = gls(data=Example5, method="ML",
    model=marital~0+kid+mom+dad+ kid:KidBoy +kid:KidEd12+mom:MomEd12+dad:DadEd12
    +kid:DadEd12+mom:DadEd12 +kid:MomEd12+dad:MomEd12,
    correlation=corSymm(form=~DVnum|FamilyID), # Unstructured correlations
    weights=varIdent(form=~1|DVnum)) # Separate variance by DV
print("Print -2LL and Results"); -2*logLik(Model4); summary(Model4)
print("Show R and RCORR matrices for first family in the data")
getVarCov(Model4, individual="3996"); corMatrix(Model4$modelStruct$corStruct)[[3]]

print("DadEd Slope Diffs -- Had to give it correct Denominator DF")
summary(glht(model=Model4, df=135, linfct=rbind(
    "Kid vs. Mom: DadEd12 Slope Diff" = c(0,0,0,0,0,0,0,-1,1,0,0), # in order of fixed effects
    "Kid vs. Dad: DadEd12 Slope Diff" = c(0,0,0,0,0,0,1,-1,0,0,0),
    "Mom vs. Dad: DadEd12 Slope Diff" = c(0,0,0,0,0,0,1,0,-1,0,0),
    "Kid vs. Mom: MomEd12 Slope Diff" = c(0,0,0,0,0,1,0,0,0,-1,0),
    "Kid vs. Dad: MomEd12 Slope Diff" = c(0,0,0,0,0,0,0,0,0,-1,1),
    "Mom vs. Dad: MomEd12 Slope Diff" = c(0,0,0,0,0,-1,0,0,0,0,1))), test=adjusted("none"))

print("Save predicted marital attitudes and correlate with actual marital attitudes")
Example5$Pred = predict(Model4, type="response")

rPred1 = cor.test(x=Example5$Pred[which(Example5$DVnum==1)],
    y=Example5$marital[which(Example5$DVnum==1)], method="pearson")
print("R and R2 for DV=1 Kid"); rPred1$estimate; rPred1$estimate^2

rPred2 = cor.test(x=Example5$Pred[which(Example5$DVnum==2)],
    y=Example5$marital[which(Example5$DVnum==2)], method="pearson")
print("R and R2 for DV=2 Mom"); rPred2$estimate; rPred2$estimate^2

```



```
rPred3 = cor.test(x=Example5$Pred[which(Example5$DVnum==3)],
                  y=Example5$marital[which(Example5$DVnum==3)], method="pearson")
print("R and R2 for DV=3 Dad"); rPred3$estimate; rPred3$estimate^2
```

**Model 4 in Path Model Software, DV-Specific Intercepts:** To what extent does **mom's education also predict kid and dad attitudes**? *uses wide-format data, ML, and "robust" standard errors for multivariate non-normality*

$$\begin{aligned} \widehat{Marital}_{fi} = & \beta_{00}(Dad_{fi}) + \beta_{01}(Kid_{fi}) + \beta_{02}(Mom_{fi}) + \beta_{11}(Kid_{fi})(KidBoy_f) \\ & + \beta_{20}(Dad_{fi})(DadEd_f - 12) + \beta_{31}(Kid_{fi})(KidEd_f - 12) + \beta_{42}(Mom_{fi})(MomEd_f - 12) \\ & + \beta_{21}(Kid_{fi})(DadEd_f - 12) + \beta_{22}(Mom_{fi})(DadEd_f - 12) \\ & + \beta_{40}(Dad_{fi})(MomEd_f - 12) + \beta_{41}(Kid_{fi})(MomEd_f - 12) \end{aligned}$$

```

display "STATA Model 4: Own + Dad + Mom Education as a Predictor of Each Attitude"
display "Using SEM to create path analysis model estimated with ML on wide-format data"
sem
(kidmarit mommarit dadmarit <- _cons)          /// All intercepts estimated (by default)
(kidmarit <- kidboy kided12)                    /// Regressions: y outcomes ON x predictors
(kidmarit mommarit dadmarit <- daded12)         ///
(kidmarit mommarit dadmarit <- momed12),        /// New effects go here
var(e.kidmarit e.mommarit e.dadmarit)          /// All residual variances estimated (by default)
covariance(e.kidmarit*e.mommarit               /// All pairwise residual covariances (not default)
           e.mommarit*e.dadmarit               ///
           e.kidmarit*e.dadmarit)              ///
method(mlmv) vce(robust)                       // Full-information ML and robust SEs
lincom _b[mommarital:daded12] - _b[kidmarital:daded12] // Kid v. Mom: Dad Educ Effect Diff
lincom _b[dadmarital:daded12] - _b[kidmarital:daded12] // Kid v. Dad: Dad Educ Effect Diff
lincom _b[dadmarital:daded12] - _b[mommarital:daded12] // Mom v. Dad: Dad Educ Effect Diff
lincom _b[mommarital:momed12] - _b[kidmarital:momed12] // Kid v. Mom: Mom Educ Effect Diff
lincom _b[dadmarital:momed12] - _b[kidmarital:momed12] // Kid v. Dad: Mom Educ Effect Diff
lincom _b[dadmarital:momed12] - _b[mommarital:momed12] // Mom v. Dad: Mom Educ Effect Diff
sem, coeflegend                                // Print parameter labels, too (to use in lincom)
sem, standardized                             // Print fully standardized solution, too
estat gof, stats(all)                         // Print fit statistics
display "LL for H1 Model= " e(critvalue_s)
display "# of parameters= " e(df_m)
display "-2LL= " e(ll)*-2                     // Print -2LL for model
estat eggof                                   // Print R2 per variable
estat residuals                               // Print how far off each predicted covariance is
estat mindices, minchi2(3.84) showpclass(all) // Print cheat codes to improve model fit p<.05

```

!!!!!!!!!!!!!!!!!!!!!!!!!!!! Mplus Lines Cannot Exceed 90 Characters !!!!!!!!!!!!!!!!!!!!!!!!!!!!!

**TITLE:** Example 5 Model 4: Own Ed + Dad & Mom Ed a Predictor of Each Attitude

DATA, VARIABLE, and OUTPUT are the same as Model 3 except for **ANALYSIS: ESTIMATOR = MLR;**

MODEL: ! \* --> Estimated parameter (all listed below for clarity)

```
! All residual variances estimated separately (by default)
KidMarit* MomMarit* DadMarit*;
```

```
! All possible pairwise residual covariances (not estimated by default)
KidMarit MomMarit DadMarit WITH KidMarit* MomMarit* DadMarit*;
```

```
! All intercepts estimated separately (by default)
[KidMarit* MomMarit* DadMarit*];
```

```
! Regressions: y outcomes ON x predictors (label to do math on later)
KidMarit ON KidBoy* KidEd12*;
KidMarit MomMarit DadMarit ON DadEd12* (DadEd2K DadEd2M DadEd2D);
KidMarit MomMarit DadMarit ON MomEd12* (MomEd2K MomEd2M MomEd2D); ! New effects here
```

```

! Getting differences in effect of DadEd for each person
MODEL CONSTRAINT: ! List names of linear combinations here
NEW (KvMDadEd KvDDadEd MvDDadEd KvMMomEd KvDMomEd MvDMomEd);
KvMDadEd = DadEd2M - DadEd2K;      ! Kid v. Mom: Dad Educ Effect Diff
KvDDadEd = DadEd2D - DadEd2K;      ! Kid v. Dad: Dad Educ Effect Diff
MvDDadEd = DadEd2D - DadEd2M;      ! Mom v. Dad: Dad Educ Effect Diff
KvMMomEd = MomEd2M - MomEd2K;      ! Kid v. Mom: Mom Educ Effect Diff
KvDMomEd = MomEd2D - MomEd2K;      ! Kid v. Dad: Mom Educ Effect Diff
MvDMomEd = MomEd2D - MomEd2M;      ! Mom v. Dad: Mom Educ Effect Diff

print("R Model 4: Own + Dad + Mom Education a Predictor of Each Attitude")
# Create model syntax as separate text object
Syntax4 = "
# Residual variances estimated separately (by default)
KidMarital ~~ KidMarital; MomMarital ~~ MomMarital; DadMarital ~~ DadMarital
# All possible pairwise residual covariances (not estimated by default)
KidMarital ~~ MomMarital + DadMarital; MomMarital ~~ DadMarital
# All intercepts estimated separately (by default)
KidMarital ~ 1; MomMarital ~ 1; DadMarital ~ 1

# Regressions: y outcomes ON x predictors (label to do math on later)
KidMarital ~ KidBoy + KidEd12
KidMarital ~ (DadEd2K)*DadEd12
MomMarital ~ (DadEd2M)*DadEd12
DadMarital ~ (DadEd2D)*DadEd12
# New effects here
KidMarital ~ (MomEd2K)*MomEd12
MomMarital ~ (MomEd2M)*MomEd12
DadMarital ~ (MomEd2D)*MomEd12

# Getting differences in effect of DadEd for each person
KvMDadEd := DadEd2M - DadEd2K; # Kid v. Mom: Dad Educ Effect Diff
KvDDadEd := DadEd2D - DadEd2K; # Kid v. Dad: Dad Educ Effect Diff
MvDDadEd := DadEd2D - DadEd2M; # Mom v. Dad: Dad Educ Effect Diff
KvMMomEd := MomEd2M - MomEd2K; # Kid v. Mom: Mom Educ Effect Diff
KvDMomEd := MomEd2D - MomEd2K; # Kid v. Dad: Mom Educ Effect Diff
MvDMomEd := MomEd2D - MomEd2M; # Mom v. Dad: Mom Educ Effect Diff
"

print("lavaan path analysis model estimated with ML on wide-format data")
PathModel4 = lavaan(data=Example5_wide, model=Syntax4, estimator="MLR", mimic="mplus")
summary(PathModel4, fit.measures=TRUE, rsquare=TRUE, standardized=TRUE, ci=TRUE)
print("Request sorted modification indices for p<.05 to troubleshoot local misfit")
modindices(object=PathModel4, sort=TRUE, minimum.value=3.84)
print("Request residual correlation matrix =leftover from observed minus predicted")
resid(object=PathModel4, type="raw") # also type="cor" for correlation matrix

```

## R LAVAAN Output—shows both regular ML and “robust” ML fit statistics:

Estimator	ML		
Optimization method	NLMINB		
Number of model parameters	17		
Number of observations	140		
Number of missing patterns	1		

Model Test User Model:	Standard	Robust	
Test Statistic	1.034	1.026	This is -2ΔLL for our H0-H1
Degrees of freedom	4	4	
P-value (Chi-square)	0.905	0.906	
Scaling correction factor		1.007	
Yuan-Bentler correction (Mplus variant)			

Model Test Baseline Model:			
Test statistic	52.998	52.902	
Degrees of freedom	15	15	
P-value	0.000	0.000	
Scaling correction factor		1.002	

User Model versus Baseline Model:			
Comparative Fit Index (CFI)	1.000	1.000	Want close to 1
Tucker-Lewis Index (TLI)	1.293	1.294	
Robust Comparative Fit Index (CFI)		1.000	
Robust Tucker-Lewis Index (TLI)		1.296	
Loglikelihood and Information Criteria:			
Loglikelihood user model (H0)	-332.158	-332.158	For our model: Larger is better
Scaling correction factor for the MLR correction		1.007	1=multivariate normality (so not bad!)
Loglikelihood unrestricted model (H1)	-331.641	-331.641	For model with all paths estimated
Scaling correction factor for the MLR correction		1.007	
Akaike (AIC)	698.316	698.316	For our model: Smaller is better
Bayesian (BIC)	748.324	748.324	For our model: Smaller is better
Sample-size adjusted Bayesian (BIC)	694.538	694.538	For our model: Smaller is better
Root Mean Square Error of Approximation:			
RMSEA	0.000	0.000	Want close to 0
90 Percent confidence interval - lower	0.000	0.000	
90 Percent confidence interval - upper	0.052	0.051	
P-value RMSEA <= 0.05	0.947	0.948	Test of RMSEA <=.05
Robust RMSEA		0.000	
90 Percent confidence interval - lower		0.000	
90 Percent confidence interval - upper		0.052	
Standardized Root Mean Square Residual:			
SRMR	0.016	0.016	Want close to 0

### Parameter estimates, their SEs, and standardized estimates would be Table 1

#### Regressions: -- THESE ARE THE FIXED SLOPES FROM MIXED

	Estimate	Std.Err	z-value	P(> z )	ci.lower	ci.upper	Std.lv	Std.all=STDYX IN MPLUS
KidMarital ~								
KidBoy	0.258	0.093	2.786	0.005	0.076	0.439	0.258	0.225 B11
KidEd12	-0.011	0.024	-0.441	0.659	-0.058	0.037	-0.011	-0.037 B31
DadEd12 (DE2K)	-0.007	0.020	-0.367	0.714	-0.046	0.032	-0.007	-0.035 B21
MomMarital ~								
DadEd12 (DE2M)	0.006	0.020	0.316	0.752	-0.033	0.046	0.006	0.031 B22
DadMarital ~								
DadEd12 (DE2D)	-0.024	0.017	-1.388	0.165	-0.057	0.010	-0.024	-0.117 B20
KidMarital ~								
MomEd12 (ME2K)	0.015	0.022	0.681	0.496	-0.028	0.059	0.015	0.072 B41
MomMarital ~								
MomEd12 (ME2M)	-0.031	0.022	-1.412	0.158	-0.073	0.012	-0.031	-0.148 B42
DadMarital ~								
MomEd12 (ME2D)	-0.056	0.019	-2.974	0.003	-0.094	-0.019	-0.056	-0.272 B40

#### Covariances: -- THESE ARE RESIDUAL COVARIANCES FROM R MATRIX OFF-DIAGONALS

	Estimate	Std.Err	z-value	P(> z )	ci.lower	ci.upper	Std.lv	Std.all
.KidMarital ~~								
.MomMarital	0.039	0.028	1.400	0.162	-0.016	0.094	0.039	0.126 UN(1,2)
.DadMarital	0.080	0.024	3.296	0.001	0.033	0.128	0.080	0.274 UN(1,3)
.MomMarital ~~								
.DadMarital	0.080	0.020	4.011	0.000	0.041	0.119	0.080	0.270 UN(2,3)

#### Intercepts: -- THESE ARE THE FIXED INTERCEPTS FROM MIXED

	Estimate	Std.Err	z-value	P(> z )	ci.lower	ci.upper	Std.lv	Std.all
.KidMarital	1.522	0.101	15.125	0.000	1.325	1.719	1.522	2.664 B01
.MomMarital	1.951	0.063	30.825	0.000	1.827	2.075	1.951	3.445 B02
.DadMarital	2.123	0.060	35.574	0.000	2.006	2.240	2.123	3.769 B00

#### Variances: -- THESE ARE THE RESIDUAL VARIANCES FROM R MATRIX DIAGONAL

	Estimate	Std.Err	z-value	P(> z )	ci.lower	ci.upper	Std.lv	Std.all
.KidMarital	0.308	0.031	10.096	0.000	0.248	0.368	0.308	0.944 UN(1,1)
.MomMarital	0.315	0.045	7.081	0.000	0.228	0.402	0.315	0.983 UN(2,2)
.DadMarital	0.278	0.034	8.092	0.000	0.211	0.345	0.278	0.876 UN(3,3)

#### R-Square: -- THESE ARE CLOSE TO BUT NOT THE SAME AS WAS FOUND IN THE UNIVARIATE MODELS

	Estimate
KidMarital	0.056
MomMarital	0.017
DadMarital	0.124

**Defined Parameters: -- THESE ARE ESTIMATE/LINCOM/GLHT/MODEL CONSTRAINT LINEAR COMBINATIONS**

	Estimate	Std.Err	z-value	P(> z )	ci.lower	ci.upper	Std.lv	Std.all	
KvMDadEd	0.014	0.026	0.522	0.602	-0.038	0.065	0.014	0.067	B22 - B21
KvDDadEd	-0.016	0.022	-0.741	0.459	-0.059	0.027	-0.016	-0.081	B20 - B21
MvDDadEd	-0.030	0.025	-1.212	0.225	-0.078	0.018	-0.030	-0.148	B20 - B22
KvMMomEd	-0.046	0.029	-1.569	0.117	-0.103	0.011	-0.046	-0.219	B42 - B41
KvDMomEd	-0.072	0.027	-2.657	0.008	-0.124	-0.019	-0.072	-0.344	B40 - B41
MvDMomEd	-0.026	0.025	-1.041	0.298	-0.074	0.023	-0.026	-0.124	B40 - B42

**\$cov - THESE ARE THE DISCREPANCIES FOR OBSERVED MINUS PREDICTED COVARIANCES**

	KdMrtl	MmMrtl	DdMrtl	KidBoy	KdEd12	DdEd12	MmEd12
KidMarital	0.002						
MomMarital	0.004	0.000					
DadMarital	0.003	0.000	0.000				
KidBoy	0.005	<b>0.015</b>	<b>0.015</b>	0.000			
KidEd12	0.013	<b>0.029</b>	<b>0.043</b>	0.000	0.000		
DadEd12	0.000	0.000	0.000	0.000	0.000	0.000	
MomEd12	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Only the kid predictors on the mom and dad outcomes have leftover covariance, and no single added paths would help the model.

**\$mean**

	KidMarital	MomMarital	DadMarital	KidBoy	KidEd12	DadEd12	MomEd12
	0	0	0	0	0	0	0

**Example results section for Part 2 Models 3–4 [picking up from Part 1; using R LAVAAN output]:**

Next, we examined whether father's education incrementally predicted the marital attitudes of the mother or adult child after controlling for their own education, but neither effect was significant (and the effect of father's education on his own attitudes was significantly larger). The effect of father's education on his own attitudes remained significant, while the effect of education on their own attitudes for the adult child and mother remained nonsignificant).

Finally, we examined the incremental effects of mother's education on marital attitudes, and results from this final model are shown in Table 1. For every additional year of mother's education, father's attitudes were expected to be significantly less conservative by 0.056 ( $p = .003$ ). The effect of mother's education on the adult child attitudes was nonsignificant and significantly smaller than its effect on father's attitudes.

We re-estimated the final model as a path analysis in the R package lavaan (using robust maximum likelihood) in order to obtain indices of absolute model fit. The model had excellent fit,  $\chi^2(4) = 1.026$ ,  $p = .906$ , RMSEA = .00 [CI = .00–.051], CFI = 1.00, indicating that no further paths were needed. This final model is depicted in Figure 1 below.

Figure 1 (line types used to help visually distinguish the paths; standardized coefficients may also be added)

