Distinguishing Genetic Effects from Common Environment in Predicting Negative Affect

Negative affect was measured for measured for 111 monozygotic (MZ) and 99 dizygotic (DZ) twin pairs. Below we conduct "ACE" modeling via multilevel models with heterogeneous variances, in which the goal is to differentiate proportions of variance due to heritability (A or H^2), common environment (C^2), and unshared or unique environment (E^2). We will use maximum likelihood (ML) estimation in order to provide comparability across results from SAS MIXED and NLMIXED (and any structural equation models).

SAS Data Manipulation:

```
* To use FitTest macro;
* FitFewer = Name of ODS InfoCrit table for nested model;
* FitMore  = Name of ODS InfoCrit table for comparison model;
%MACRO FitTest(FitFewer=,FitMore=);
DATA &FitFewer.; LENGTH Name $30.; SET &FitFewer.; Name="&FitFewer."; RUN;
DATA &FitMore.;  LENGTH Name $30.; SET &FitMore.;  Name="&FitMore.";  RUN;
DATA FitCompare; LENGTH Name $30.; SET &FitFewer. &FitMore.; RUN;
DATA FitCompare; SET FitCompare; DevDiff=Lag1(Neg2LogLike)-Neg2LogLike;
     DFdiff=Parms-LAG1(Parms); Pvalue=1-PROBCHI(DevDiff,DFdiff);
     DROP AICC HQIC CAIC; RUN;
TITLE9 "Likelihood Ratio Test for &FitFewer. vs. &FitMore.";
PROC PRINT NOOBS DATA=FitCompare; RUN; TITLE9;
%MEND FitTest;

* Location for original SAS files for these models;
LIBNAME import SPSS 'C:\Dropbox\_Archive\Example Data\Dave Twins\DaveTwin.por';

* Saving file in work library;
DATA work.TwinData; SET import.DaveTwin;
       PairID=FamID2; * Renaming family ID and coding for zygosity;
       IF zyg=1 THEN DO; Zygosity="MZ"; MZ=1; DZ=0; END;
       IF zyg=7 THEN DO; Zygosity="DZ"; MZ=0; DZ=1; END;
       * Stacking into 1 row per person (2 per family);
       Twin=1; NegAffect=Day1; OUTPUT;
       Twin=2; NegAffect=Day2; OUTPUT;
       KEEP PairID Twin Zygosity NegAffect;
RUN;
PROC SORT DATA=work.TwinData; BY PairID Twin Zygosity; RUN;
```

SAS Mixed Models and Output:

```
TITLE1 "No Random Intercept, Homogeneous Variance by MZ-DZ";
TITLE2 "Baseline model -- assumes no common variance of any kind";
PROC MIXED DATA=work.TwinData COVTEST NOCLPRINT IC NAMELEN=100 METHOD=ML;
       CLASS PairID Twin Zygosity;
       MODEL NegAffect = Zygosity / NOINT CL SOLUTION DDFM=Satterthwaite;
       REPEATED Twin / TYPE=VC SUBJECT=PairID;  * No covariance between twins;
       ODS OUTPUT InfoCrit=FitIndependence; RUN;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr > Z
Twin	PairID	1.5684	0.1082	14.49	<.0001

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
1380.9	**3**	1386.9	1387.0	1391.0	1397.0	1400.0

Solution for Fixed Effects → Means per Zygosity Directly via NOINT

Effect	Zygosity	Estimate	Standard Error	DF	t Value	Pr > |t|	Lower	Upper
Zygosity	DZ	6.9627	0.08900	420	78.23	<.0001	6.7877	7.1376
Zygosity	MZ	6.8947	0.08405	420	82.03	<.0001	6.7294	7.0599

```
TITLE1 "Add Random Intercept, Homogeneous Variances by MZ-DZ";
TITLE2 "Tests separation of shared (common) and unshared (environment) variance";
PROC MIXED DATA=work.TwinData COVTEST NOCLPRINT IC NAMELEN=100 METHOD=ML;
       CLASS PairID Twin Zygosity;
       MODEL NegAffect = Zygosity / NOINT CL SOLUTION DDFM=Satterthwaite;
       RANDOM INTERCEPT / VCORR TYPE=UN SUBJECT=PairID;      * Level 2;
       REPEATED Twin    / TYPE=VC SUBJECT=PairID;           * Level 1;
       ODS OUTPUT InfoCrit=FitSharedVar; RUN;

* Test of random intercept variance;
%FitTest(FitFewer=FitIndependence, FitMore=FitSharedVar);
```

Estimated V Correlation

Matrix for PairID 5000

Row	Col1	Col2
1	1.0000	0.2525
2	**0.2480**	1.0000

ICC (shown in VCORR):
0.3984 / (0.3890 + 1.1795) = .2480

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr > Z
UN(1,1)	PairID	0.3890	0.1115	3.49	0.0002
Twin	PairID	1.1795	0.1151	10.25	<.0001

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
1	**13.33**	0.0003

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
1367.6	**4**	1375.6	1375.7	1381.0	1389.0	1393.0

Solution for Fixed Effects → Means per Zygosity Directly via NOINT

Effect	Zygosity	Estimate	Standard Error	DF	t Value	Pr > |t|	Lower	Upper
Zygosity	DZ	6.9627	0.09943	210	70.03	<.0001	6.7667	7.1587
Zygosity	MZ	6.8947	0.09390	210	73.43	<.0001	6.7096	7.0798

Likelihood Ratio Test for FitIndependence vs. FitSharedVar

Name	Neg2Log Like	Parms	AIC	BIC	DevDiff	DFdiff	Pvalue
FitIndependence	1380.9	3	1386.9	1397.0	.	.	.
FitSharedVar	1367.6	4	1375.6	1389.0	13.3311	1	.000261043

```
TITLE1 "Keep Random Intercept, Add Heterogeneous Variances by MZ-DZ via GROUP=";
TITLE2 "Tests separation of shared variance into common versus genetic";
PROC MIXED DATA=work.TwinData COVTEST NOCLPRINT IC NAMELEN=100 METHOD=ML;
       CLASS PairID Twin Zygosity;
       MODEL NegAffect = Zygosity / NOINT CL SOLUTION DDFM=Satterthwaite;
* Printing VCORR for 1st and 4th pairs, to get separate ICCs for MZ and DZ;
       RANDOM INTERCEPT / VCORR=1,4 TYPE=UN SUBJECT=PairID GROUP=Zygosity;   * Level 2;
       REPEATED Twin    /           TYPE=VC SUBJECT=PairID GROUP=Zygosity;   * Level 1;
       ODS OUTPUT InfoCrit=FitGeneticVar; RUN;

* Test of 2 different variances by zygosity;
%FitTest(FitFewer=FitSharedVar, FitMore=FitGeneticVar);
```

Estimated V Correlation

Matrix for PairID 5000 (MZ)

Row	Col1	Col2
1	1.0000	0.3058
2	0.3017	1.0000

MZ ICC (shown in VCORR):
0.4088 / (0.4088 + 0.9462) = .3017

Estimated V Correlation
Matrix for PairID 5013 (DZ)

Row	Col1	Col2
1	1.0000	0.2077
2	0.2029	1.0000

DZ ICC (shown in VCORR):
0.3668 / (0.3668 + 1.4410) = .2029

Covariance Parameter Estimates

Cov Parm	Subject	Group	Estimate	Standard Error	Z Value	Pr > Z
UN(1,1)	PairID	Zygosity DZ	0.3668	0.1854	1.98	0.0239
UN(1,1)	PairID	Zygosity MZ	0.4088	0.1343	3.04	0.0012
Twin	PairID	Zygosity DZ	1.4410	0.2048	7.04	<.0001
Twin	PairID	Zygosity MZ	0.9462	0.1270	7.45	<.0001

Information Criteria

Neg2LogLike	Parms	AIC	AICC	HQIC	BIC	CAIC
1361.8	**6**	1373.8	1374.0	1381.9	1393.9	1399.9

Solution for Fixed Effects → Means per Zygosity Directly via NOINT

Effect	Zygosity	Estimate	Standard Error	DF	t Value	Pr > |t|	Lower	Upper
Zygosity	DZ	6.9627	0.1048	99	66.44	<.0001	6.7547	7.1706
Zygosity	MZ	6.8947	0.08913	111	77.35	<.0001	6.7180	7.0713

Likelihood Ratio Test for FitSharedVar vs. FitGeneticVar

Name	Neg2Log Like	Parms	AIC	BIC	DevDiff	DFdiff	Pvalue
FitSharedVar	1367.6	4	1375.6	1389.0	.	.	.
FitGeneticVar	1361.8	6	1373.8	1393.9	5.78202	2	0.055520

Heritability (A or H^2), or the contribution of genetics, can be found as twice the difference of the intraclass correlation between MZ and DZ twins. **Common environment** (C^2) can be found as the difference between the intraclass correlation for MZ twins and the heritability estimate (usually constrained to be ≥ 0), and the **unique environment** (E^2) can be found as the remainder (i.e., 1 – [heritability + common environment]).

Heritability (H^2) = 2*($ICC_{MZ} - ICC_{DZ}$) = 2*(0.3017 −0 .2029) = 0.1976
Common Environment (C^2) = $ICC_{MZ} - H^2$ = 0.3017 – 0.1976 = 0.1041
Unique Environment (E^2) = 1 – ($H^2 + C^2$) = 1 – (0.1976 + 0.1041) = 0.6983

In addition to point estimates for these variance sources, it is often required to provide standard errors and confidence intervals for them as well. I don't know of a way to do so in MIXED, but once can use ESTIMATE statements to create linear combinations of variances in NLMIXED, as shown below.

```
TITLE1 "Last Model in MIXED Repeated in NLMIXED";
TITLE2 "Used to obtain SEs and CIs for ACE estimates";
PROC NLMIXED DATA=work.TwinData METHOD=GAUSS TECH=NEWRAP GCONV=1e-12;
* Must define all parameters to be estimated and provide start values;
PARMS fintDZ=6.9 fintMZ=6.9 vU0intDZ=0.37 vU0intMZ=0.41 varEresDZ=1.44 varEresMZ=0.95;
* Setting up composite equation WITHOUT level-1 residual;
      y = DZ*fintDZ + DZ*U0i + MZ*fintMZ + MZ*U0i;
* Model for heterogeneous variance at level-2 and level-1;
      vU0int  = DZ*vU0intDZ  + MZ*vU0intMZ;
      varEres = DZ*varEresDZ + MZ*varEresMZ;
* Computing ICCs and ACE model parameters;
      ICCMZ = vU0intMZ/(vU0intMZ+varEresMZ);
      ICCDZ = vU0intDZ/(vU0intDZ+varEresDZ);
      H2 = 2*(ICCMZ-ICCDZ);
      C2 = ICCMZ-H2;
* Telling it which DV, defining level-1 residual;
* NegAffect is normally distributed with a mean of "y" and a variance of "VarEres";
      MODEL NegAffect ~ NORMAL(y,varEres);
```

```
* Defining random effects: normally distributed with given means and variances;
      RANDOM U0i ~ NORMAL([0],[vU0int]) SUBJECT=PairID;
* Labeling parameters;
      ESTIMATE "Mean for MZ"                  fintMZ;
      ESTIMATE "Mean for DZ"                  fintDZ;
      ESTIMATE "ICC for MZ"                   ICCMZ;
      ESTIMATE "ICC for DZ"                   ICCDZ;
      ESTIMATE "Heritability A or H2"  H2;
      ESTIMATE "Common Environment C2" C2;
      ESTIMATE "Unique Environment E2" 1-(H2+C2);
RUN; TITLE1; TITLE2;
```

NOTE: ABSGCONV convergence criterion satisfied.

Fit Statistics

-2 Log Likelihood	**1361.8**
AIC (smaller is better)	1373.8
AICC (smaller is better)	1374.0
BIC (smaller is better)	1393.9

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Lower	Upper	Gradient
fintDZ	6.9627	0.1048	209	66.44	<.0001	6.7561	7.1693	3.26E-8
fintMZ	6.8947	0.08913	209	77.35	<.0001	6.7189	7.0704	8.22E-12
vU0intDZ	0.3668	0.1854	209	1.98	0.0492	0.001288	0.7322	-4.72E-8
vU0intMZ	0.4088	0.1343	209	3.04	0.0026	0.1440	0.6736	7.22E-12
varEresDZ	1.4410	0.2048	209	7.04	<.0001	1.0372	1.8448	-2.41E-8
varEresMZ	0.9462	0.1270	209	7.45	<.0001	0.6958	1.1966	5.19E-12

Additional Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Lower	Upper
Mean for MZ	6.8947	0.08913	209	77.35	<.0001	6.7189	7.0704
Mean for DZ	6.9627	0.1048	209	66.44	<.0001	6.7561	7.1693
ICC for MZ	0.3017	0.08628	209	3.50	0.0006	0.1316	0.4718
ICC for DZ	0.2029	0.09637	209	2.11	0.0365	0.01290	0.3929
Heritability A or H2	**0.1976**	**0.2587**	209	0.76	0.4457	**-0.3123**	**0.7076**
Common Environment C2	**0.1041**	**0.2112**	209	0.49	0.6227	**-0.3122**	**0.5203**
Unique Environment E2	**0.6983**	**0.08628**	209	8.09	<.0001	0.5282	0.8684

Thus, it appears the distinction between heritability and common environment is not reliable. Below is a comparison with previous results from MIXED—everything is exactly the same.

Solution for Fixed Effects → Means per Zygosity Directly via NOINT

Effect	Zygosity	Estimate	Standard Error	DF	t Value	Pr > |t|	Lower	Upper
Zygosity	DZ	6.9627	0.1048	99	66.44	<.0001	6.7547	7.1706
Zygosity	MZ	6.8947	0.08913	111	77.35	<.0001	6.7180	7.0713

Covariance Parameter Estimates

Cov Parm	Subject	Group	Estimate	Standard Error	Z Value	Pr > Z
UN(1,1)	PairID	Zygosity DZ	0.3668	0.1854	1.98	0.0239
UN(1,1)	PairID	Zygosity MZ	0.4088	0.1343	3.04	0.0012
Twin	PairID	Zygosity DZ	1.4410	0.2048	7.04	<.0001
Twin	PairID	Zygosity MZ	0.9462	0.1270	7.45	<.0001

Heritability (H^2) = 2*($ICC_{MZ} - ICC_{DZ}$) = 2*(0.3017 −0 .2029) = 0.1976
Common Environment (C^2) = $ICC_{MZ} - H^2$ = 0.3017 – 0.1976 = 0.1041
Unique Environment (E^2) = 1 – (H^2 + C^2) = 1 – (0.1976 + 0.1041) = 0.6983