

# Comparing the SAS<sup>®</sup> GLM and MIXED Procedures for Repeated Measures

Russ Wolfinger and Ming Chang, SAS Institute Inc., Cary, NC

## Abstract

Repeated measures analyses in the SAS GLM procedure involve the traditional univariate and multivariate approaches. The SAS MIXED procedure employs a more general covariance structure approach. This paper compares the two procedures and helps you understand their methodologies. A numerical example illustrates many of the key similarities and differences.

## Introduction

The analysis of repeated measures involves data which consist of multiple measurements on experimental units such as individuals, animals, or machines. These experimental units are called *subjects*. This paper focuses on *longitudinal data*, in which the repeated measurements on a subject occur over time (for example, a growth curve), and how they are handled by the REPEATED statements in the GLM procedure (SAS Institute Inc. 1989) and the MIXED procedure (SAS Institute Inc. 1992).

As an example of longitudinal data, consider the results from Pothoff and Roy (1964), which consist of dental measurements from the center of the pituitary to the pteryomaxillary fissure for 11 girls and 16 boys at ages 8, 10, 12, and 14. The subjects are the individual children, and there are four repeated measurements on each. You can load these data into two different SAS data sets using the code to the right. The first data set, FORGLM, will be appropriate for use with PROC GLM, while the second, FORMIXED, will be used with PROC MIXED.

The analysis of this example entertains models for both the expected value of the observations and for their within-subject variance-covariance matrix. The models for the expected value of the observations fall within the classical general linear model framework, which models the mean of the responses as a linear function of known explanatory variables. These explanatory variables can be either classification (ANOVA) or continuous (regression) type variables, and they comprise the *fixed effects* of the model (refer to Searle 1971). Regarding the variability of the data, assume that data from different subjects are statistically independent and that the variance-covariance matrix is the same for each subject.

```
data forglm(keep=person gender y1-y4)
  formixed(keep=person gender age y);
input person gender$ y1-y4;
output forglm;
y=y1; age=8; output formixed;
y=y2; age=10; output formixed;
y=y3; age=12; output formixed;
y=y4; age=14; output formixed;
datalines;
1 F 21.0 20.0 21.5 23.0
2 F 21.0 21.5 24.0 25.5
3 F 20.5 24.0 24.5 26.0
4 F 23.5 24.5 25.0 26.5
5 F 21.5 23.0 22.5 23.5
6 F 20.0 21.0 21.0 22.5
7 F 21.5 22.5 23.0 25.0
8 F 23.0 23.0 23.5 24.0
9 F 20.0 21.0 22.0 21.5
10 F 16.5 19.0 19.0 19.5
11 F 24.5 25.0 28.0 28.0
12 M 26.0 25.0 29.0 31.0
13 M 21.5 22.5 23.0 26.5
14 M 23.0 22.5 24.0 27.5
15 M 25.5 27.5 26.5 27.0
16 M 20.0 23.5 22.5 26.0
17 M 24.5 25.5 27.0 28.5
18 M 22.0 22.0 24.5 26.5
19 M 24.0 21.5 24.5 25.5
20 M 23.0 20.5 31.0 26.0
21 M 27.5 28.0 31.0 31.5
22 M 23.0 23.0 23.5 25.0
23 M 21.5 23.5 24.0 28.0
24 M 17.0 24.5 26.0 29.5
25 M 22.5 25.5 25.5 26.0
26 M 23.0 24.5 26.0 30.0
27 M 22.0 21.5 23.5 25.0
;
```

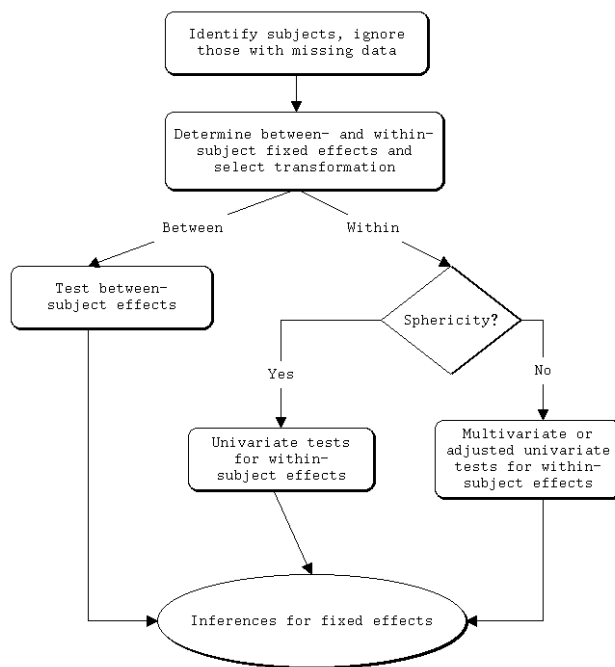
The repeated measures aspect of the data makes it interesting because observations on the same subject are usually correlated and often exhibit heterogeneous variability. If such correlation and heterogeneity are not present, a standard ordinary least squares analysis in PROC GLM is appropriate, because it assumes the observations are uncorrelated and have constant variance. When these properties are present, though, you should use a methodology that accounts for them, especially with regards to inferences about the fixed effects.

Both PROC GLM and PROC MIXED offer repeated measures analyses that account for within-subject covariability, and the following two sections compare their methods. The first section overviews and compares their overall analysis strategies, and the second applies and contrasts these

strategies to the dental data example.

## Analysis Strategies

Figure 1 depicts the traditional repeated measures strategy implemented in PROC GLM. The first thing to notice about PROC GLM's analysis is that it requires the data to be balanced within subjects; that is, it does not use any data from subjects that have missing observations. After identifying the subjects with complete data, you need to select a working mean model in terms of *between-subject* and *within-subject* fixed effects. Between-subject effects are those whose levels remain constant within subjects, whereas within-subject effects change within subjects.



**Figure 1.** Repeated Measures Analysis in PROC GLM

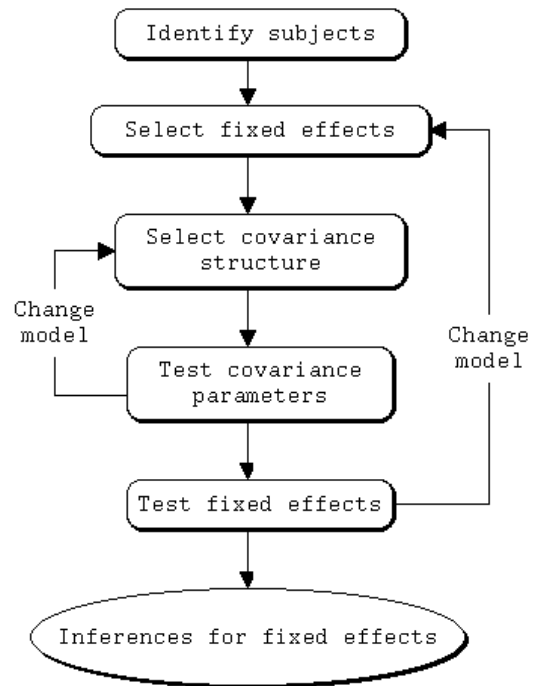
In the dental data example, GENDER is a between-subject effect and AGE and AGE\*GENDER are within-subject effects. The distinction between the two types of effects must be made in PROC GLM because you must place the between-subject effects on the MODEL statement and the main within-subject effect on the REPEATED statement. In this case AGE becomes the REPEATED effect, and PROC GLM automatically generates AGE\*GENDER.

The second box in Figure 1 also indicates that you must select a transformation in a PROC GLM repeated measures analysis. This is because PROC GLM performs its calculations on a set of contrast variables numbering one less than the number of repeated measures variables. Several possible transformations are described in SAS Institute Inc.

(1989), the default being the contrast of the levels of the repeated effect with its last level.

Continuing in Figure 1, PROC GLM performs a standard significance test for the between-subject effects. However, two different kinds of tests are available for the within-subject effects: *univariate* and *multivariate*. The univariate tests are appropriate when the within-subject variance-covariance matrix of the observations has a certain structural form known as Type H (Huynh and Feldt 1970). PROC GLM performs a statistical test for this structure known as the *sphericity test*. When the sphericity test does not have a significant *p*-value, you should use the univariate tests for within-subject effects because under the Type H assumption they will usually be more powerful than the multivariate tests.

When the sphericity test is significant, PROC GLM offers you two ways to test the significance of the within-subject effects. The first way is to adjust the univariate tests themselves, and GLM prints two such adjustments: G-G (Greenhouse and Geisser 1959) and the less conservative H-F (Huynh and Feldt 1976). The second way involves four different multivariate tests: Wilks' Lambda, Pillai's Trace, Hotelling-Lawley Trace, and Roy's Greatest Root (refer to SAS Institute Inc. 1989). These tests are all based on a completely general (unstructured) within-subject variance-covariance matrix.



**Figure 2.** Repeated Measures Analysis in PROC MIXED

Finally, inferences on the fixed effects provide answers to your research questions through tests of linear hypothe-

sis. These fixed-effect inferences can involve not only the default Type I and III  $F$ -tests but also customized tests available with TEST, CONTRAST, ESTIMATE, MEANS, and LSMEANS statements. However, you should be aware that these tests are sometimes limited in scope in PROC GLM because of the dichotomy between the within-subject and between-subject effects.

Figure 2 portrays the repeated measures analysis strategy for PROC MIXED. The first difference from Figure 1 is that you can use all available data in the PROC MIXED analysis instead of ignoring subjects with missing data. The reason for this generalization is that PROC MIXED uses a likelihood-based estimation method but PROC GLM uses a method of moments that requires complete data.

Secondly, you do not have to make an initial distinction between within- and between-subject fixed effects in the PROC MIXED approach as you do in PROC GLM. You simply determine the entire mean model and place all fixed effects on the MODEL statement. Furthermore, you do not have to select a transformation in a PROC MIXED analysis.

The PROC MIXED mean specification is actually more general than the one in PROC GLM in two ways:

1. You can omit between-within interaction effects from the PROC MIXED mean model but you cannot in PROC GLM.
2. You can use continuous variables in within-subject effects in PROC MIXED, whereas all within-subject effects must consist of classification variables in PROC GLM.

The first point is true because PROC GLM creates and tests the between-within interactions automatically, whereas you can simply drop them from the MODEL statement in PROC MIXED. Regarding the second point, PROC GLM does allow you to specify contrasts that may be appropriate for a continuous within-subject effect, but the number of the contrasts is carried as high as possible, thus effectively fitting all possible degrees of freedom for the effect. PROC MIXED allows you to actually specify and fit a reduced model. For example, in the dental data, you may want to fit just a linear trend in AGE, removing the quadratic and cubic terms.

Another key difference between Figures 2 and 3 is that you must explicitly specify a covariance structure in PROC MIXED. The covariance structure specification in PROC MIXED is important because the test statistics for the fixed effects are functions of it, and PROC MIXED can produce invalid results if the structure is misspecified. Consequently, you should compare several covariance structures and select one that is reasonable. PROC MIXED provides you with a variety of possible structures to choose from in addition to the Type H and unstructured matrices used by PROC GLM. These include compound symmetry, autoregressive and other time series structures, random coefficients models, and spatial correlations. A strategy for covariance structure selection process is provided in Wolfinger (1993) and is indicated in Figure 2 by the loop back after testing the covariance parameters.

After selecting a covariance structure, you should check the significance of your fixed effects. If some are not significant, you can drop them from the model and again make sure the covariance structure is appropriate. This is indicated in Figure 2 by the loop back to "Select fixed effects."

With a final mean-variance model in hand, you are equipped to draw inferences about your fixed effects by performing hypothesis tests and constructing relevant confidence intervals. The CONTRAST, ESTIMATE, and LSMEANS statements in PROC MIXED are tailored to this purpose; they have the advantage over their counterparts in PROC GLM in that all standard error estimates account for the estimated covariance structure, whereas those in PROC GLM often do not. Furthermore, the fact that all fixed effects are together on the MODEL statement in PROC MIXED allows you to specify comparisons between them, but the separation between within-subject and between-subject effects makes this task difficult in PROC GLM.

## Example

This section applies the aforementioned analysis strategies to the Pothoff-Roy dental data described in the Introduction. Refer also to PROC MIXED Example 16.2 (SAS Institute Inc. 1992), Wolfinger, Tobias, and Sall (1991), and Latour, Latour, and Wolfinger (1994).

The SAS code from the Introduction creates two different SAS data sets, FORGLM and FORMIXED, containing the dental data. In FORGLM, the observations are stored in multivariate form, but in FORMIXED, they are strung out into one long response variable Y. This latter structure enables you to specify time-varying covariates in PROC MIXED, but this example does not consider them to facilitate comparison with PROC GLM.

## PROC GLM analysis

You can perform a PROC GLM repeated measures analysis with the following code:

```
proc glm data=forglm;
  class gender;
  model y1-y4=gender / nouni;
  repeated age 4 (8 10 12 14) / printe;
run;
```

GENDER is the only classification effect, and it becomes the single between-subject fixed effect on the MODEL statement. All four repeated measures variables are listed on the left hand side of the MODEL equation. The NOUNI option suppresses the printing of one-way ANOVAs for each of the four variables. The REPEATED statement contains the within-subject repeated measures effect AGE. After the effect, you can specify its four levels as shown here, or PROC GLM sets them to 1,2,3,4 by default.

The AGE effect is effectively treated as a classification variable, and PROC GLM automatically includes it as such in the model along with its interaction with all of the between-subject effects. No transformation for AGE is specified here, and so PROC GLM uses the one contrasting ages 8, 10,

and 12 with age 14. The polynomial transformation is considered later in this paper.

Finally, the PRINTE option on the REPEATED statement is important because it instructs PROC GLM to carry out the test for sphericity, which is central in Figure 1.

The output from this code is as follows.

**Output 1. PROC GLM Levelization Results**

```

General Linear Models Procedure
Class Level Information

Class   Levels   Values
-----
GENDER      2   F M

Number of observations in data set = 27

General Linear Models Procedure
Repeated Measures Analysis of Variance
Repeated Measures Level Information

Dependent Variable      Y1      Y2      Y3      Y4
Level of AGE            8      10     12     14

```

Output 1 prints the levels of the CLASS effect GENDER as well as the levels of the REPEATED factor AGE, which is also effectively considered to be a class variable for the analysis.

**Output 2. PROC GLM Sphericity Tests**

```

Test for Sphericity: Mauchly's Criterion = 0.4998695
Chisquare Approximation = 16.449181 with 5 df
Prob > Chisquare = 0.0057

Applied to Orthogonal Components:
Test for Sphericity: Mauchly's Criterion = 0.7353334
Chisquare Approximation = 7.2929515 with 5 df
Prob > Chisquare = 0.1997

```

Output 2 is a partial listing of the results from the PRINTE option. The omitted portions are printouts for partial correlation coefficients from the error sums of squares and crossproducts matrix for both the original four repeated measures and for the three variables resulting from a transformation of the four variables.

Output 2 shows the results of two sphericity tests (Mauchly 1940; Anderson 1958). The second of these tests is an important part of a PROC GLM approach to repeated measures because it helps you to determine what type of significance test to use for your within-subject effects. The test checks whether or not the within-subject variance-covariance matrix has a Type H covariance structure. The connection is that the univariate repeated measures  $F$ -statistics have  $F$ -distributions under the null hypothesis only if the within-subject variance-covariance matrix is of Type H (Huynh and Feldt 1970).

Mauchly's test is carried out in terms of a statistic  $W$ . The theory is that  $-\rho n_1 \log W$  has an approximate  $\chi^2$ -distribution with  $f$  degrees of freedom, where  $n_1$  is the number of subjects minus the rank of the between-subject design matrix,

$$\rho = 1 - (2p^2 + p + 2) / (6pn_1)$$

$$f = p(p + 1) / 2 - 1$$

and  $p$  is the number of repeated measurements minus 1. The correction factor  $\rho$  is not necessary for the asymptotic validity of the test, since it converges to 1 as the number of subjects goes to infinity and the number of repeated measures stays constant; however,  $\rho$  does improve the  $\chi^2$  approximation in small samples.

For the dental data,  $n_1 = 27 - 2 = 25$  and  $p = 3$ . So using  $W = 0.7353334$  from Output 2,

$$n_1 \log W = 7.6857804$$

$$\rho = 0.9488889$$

$$\chi^2 = 7.2929516$$

on  $f = 5$  degrees of freedom.

To obtain  $n_1 \log W$ , PROC GLM uses the fact that a covariance matrix is of Type H if and only if its quadratic form with an orthogonal contrast matrix equals a constant times the identity matrix (Huynh and Feldt 1970). It therefore applies an orthogonal transformation to the original variables and then performs a likelihood ratio test on the covariance structure for the transformed variables, comparing the simple null model ( $\sigma^2$  times an identity matrix) with the unstructured form.

The appendix shows how you can reproduce PROC GLM's calculations using PROC IML and PROC MIXED, although a simpler and more direct approach is described in the next subsection.

For these data, the  $p$ -value of 0.1997 is not significant, indicating that you cannot reject the null hypothesis that the matrix is of Type H. So as indicated in Figure 1, the standard univariate tests for AGE and AGE\*GENDER are appropriate.

**Output 3. PROC GLM Multivariate Tests for Within-Subject Effects**

```

Manova Test Criteria and Exact F Statistics for
the Hypothesis of no AGE Effect
H = Type III SS&CP Matrix for AGE   E = Error SS&CP Matrix

S=1      M=0.5      N=10.5

Statistic      Value      F      Num DF      Den DF      Pr > F
Wilks' Lambda      0.194794      31.691      3      23      0.0001
Pillai's Trace      0.805206      31.691      3      23      0.0001
Hotelling-Lawley Trace      4.133622      31.691      3      23      0.0001
Roy's Greatest Root      4.133622      31.691      3      23      0.0001

Manova Test Criteria and Exact F Statistics for
the Hypothesis of no AGE*GENDER Effect
H = Type III SS&CP Matrix for AGE*GENDER
E = Error SS&CP Matrix

S=1      M=0.5      N=10.5

Statistic      Value      F      Num DF      Den DF      Pr > F
Wilks' Lambda      0.739887      2.6953      3      23      0.0696
Pillai's Trace      0.260113      2.6953      3      23      0.0696
Hotelling-Lawley Trace      0.351557      2.6953      3      23      0.0696
Roy's Greatest Root      0.351557      2.6953      3      23      0.0696

```

Output 3 prints the multivariate tests for the within-subject effects AGE and AGE\*GENDER, although, according to Mauchly's test in Output 2, it is not necessary to make a multivariate assumption about these data. These tests are

still valid but are less powerful than the univariate tests given the Type H assumption (Muller et al. 1992).

**Output 4. PROC GLM Tests for Between-Subject Effects**

General Linear Models Procedure Repeated Measures Analysis of Variance Tests of Hypotheses for Between Subjects Effects					
Source	DF	Type III SS	F Value	Pr > F	
GENDER	1	140.46485690	9.29	0.0054	
Error	25	377.91477273			

Output 4 prints the Type III *F*-test for GENDER, the lone between-subject effect, and the test is significant at the 1% level. The assumption of sphericity is not required for this test to be valid.

**Output 5. PROC GLM Univariate Tests for Within-Subject Effects**

General Linear Models Procedure Repeated Measures Analysis of Variance Univariate Tests of Hypotheses for Within Subject Effects									
Source: AGE									
	DF	Type III	SS	Mean Square	F Value	Pr > F	Adj G - G	Pr > F H - F	
	3	209.436974	69.812325	35.35	0.0001	0.0001	0.0001	0.0001	
Source: AGE*GENDER									
	DF	Type III	SS	Mean Square	F Value	Pr > F	Adj G - G	Pr > F H - F	
	3	13.992529	4.664176	2.36	0.0781	0.0878	0.0878	0.0781	
Source: Error(AGE)									
	DF	Type III	SS	Mean Square					
	75	148.127841	1.975038						
Greenhouse-Geisser Epsilon = 0.8672 Huynh-Feldt Epsilon = 1.0156									

Output 5 prints both standard and adjusted univariate Type III *F*-tests for the within-subject effects AGE and AGE\*GENDER. The standard tests are constructed in the usual fashion by taking the ratio of the mean square of the effect with that of the error term, denoted Error(AGE).

The two adjustments are based on a degrees of freedom adjustment factor known as  $\epsilon$  satisfying  $0 < \epsilon \leq 1$  (Box 1954). Both adjustments estimate  $\epsilon$  and then multiply the numerator and denominator degrees of freedom by this estimate before determining significance levels for the *F*-tests. The first adjustment, denoted G-G, was proposed by Greenhouse and Geisser (1959). Huynh and Feldt (1976) have shown that G-G estimate of Box's  $\epsilon$  factor is too conservative, especially for small samples, and have proposed an alternative adjustment, which is denoted H-F.

The results of Output 2 indicate that the adjustments are not necessary, although they have little effect. In fact, there is no H-F adjustment for AGE\*GENDER because the H-F estimate of  $\epsilon$  is greater than 1 and is set equal to 1 when multiplying the degrees of freedom.

**PROC MIXED Analysis and Comparison with PROC GLM**

As displayed in Figure 2, several PROC MIXED runs are usually necessary in a typical repeated measures analy-

sis. To illustrate the search for an appropriate covariance structure, consider the following three sets of statements:

```
proc mixed data=formixed;
  class gender age person;
  model y = gender|age;
  repeated / type=cs sub=person;
run;
```

```
proc mixed data=formixed;
  class gender age person;
  model y = gender|age;
  repeated / type=hf sub=person;
run;
```

```
proc mixed data=formixed;
  class gender age person;
  model y = gender|age;
  repeated / type=un sub=person;
run;
```

All three PROC MIXED specifications have the same CLASS and MODEL statements. The CLASS statement specifies GENDER, AGE, and PERSON to be classification effects, and the MODEL statement specifies the mean model for the data. The GENDER|AGE specification is expanded to GENDER AGE GENDER\*AGE in the PROC MIXED fixed-effects analysis.

Note that you place all fixed effects, both between- and within-subject, on the MODEL statement in PROC MIXED. This is in contrast to PROC GLM, in which you place between-subject effects on the MODEL statement and the main within-subject effect on the REPEATED statement. The flexibility of the PROC MIXED method is evident in that you can easily remove AGE from the CLASS statement to fit a gender-specific linear regression model, and/or you can drop the AGE\*GENDER interaction. You cannot fit any of these reduced models in PROC GLM.

The only difference in the preceding three PROC MIXED specifications is with respect to the TYPE= option on the REPEATED statement. The three models have the compound symmetry (CS), Type H or Huynh-Feldt (HF), and unstructured (UN) forms for the within-subject variance covariance matrix.

The SUB= option specifies PERSON to be the subject effect, which instructs PROC MIXED to make the  $108 \times 108$  variance-covariance matrix of the entire data vector to be block diagonal with  $27 \times 4 \times 4$  blocks. Each of these blocks has the covariance structure given by the TYPE= option.

You can optionally include AGE as an effect on the PROC MIXED REPEATED statement before the slash (/). This is not necessary when the data are balanced (as in this example), when you explicitly indicate which data values are missing in your SAS data set, or when all missing data are at the end of each subject's records. If none of these conditions apply, you must include AGE as a REPEATED effect to inform PROC MIXED of the correct placement of the observed responses.

The output from these three sets of statements is as follows.

**Output 6. PROC MIXED Levelization Results**

```

The MIXED Procedure

Class Level Information

Class      Levels  Values
-----
GENDER     2      F M
AGE        4      8 10 12 14
PERSON     27     1 2 3 4 5 6 7 8 9 10 11 12 13
          14 15 16 17 18 19 20 21 22 23
          24 25 26 27
    
```

All three PROC MIXED outputs begin with the same Class Level Information table shown in Output 6. Output 6 is similar to Output 1, except that PERSON is also included as a classification effect. The PERSON effect is necessary in PROC MIXED because of its extended data structure.

Next, each PROC MIXED call prints the REML Estimation Iteration History and Covariance Parameter Estimates tables displayed in Output 7.

**Output 7. PROC MIXED REML Estimation Results**

TYPE=CS:

```

REML Estimation Iteration History

Iteration  Evaluations  Objective  Criterion
-----
0          1 286.70313978
1          1 239.62082619  0.00000000

Convergence criteria met.

Covariance Parameter Estimates (REML)

Cov Parm      Ratio  Estimate  Std Error  Z
-----
DIAG CS       1.66345582  3.28538826  1.07194118  3.06
Residual      1.00000000  1.97503788  0.32252234  6.12

Covariance Parameter Estimates (REML)

Pr > |Z|
-----
0.0022
0.0001
    
```

TYPE=HF:

```

REML Estimation Iteration History

Iteration  Evaluations  Objective  Criterion
-----
0          1 286.70313978
1          2 237.93345023  0.00000481
2          1 237.93287471  0.00000000

Convergence criteria met.

Covariance Parameter Estimates (REML)

Cov Parm      Estimate  Std Error  Z  Pr > |Z|
-----
DIAG VAR(1)   5.02640775  1.34783547  3.73  0.0002
VAR(2)        4.39507681  1.35710319  3.24  0.0012
VAR(3)        6.17387297  1.63565028  3.77  0.0002
VAR(4)        5.28477803  1.61323942  3.28  0.0011
HF            1.97502054  0.32251667  6.12  0.0001
Residual      1.00002271  .          .    .
    
```

TYPE=UN:

```

REML Estimation Iteration History

Iteration  Evaluations  Objective  Criterion
-----
0          1 286.70313978
1          1 230.24709436  0.00000000

Convergence criteria met.

Covariance Parameter Estimates (REML)

Cov Parm      Estimate  Std Error  Z  Pr > |Z|
-----
DIAG UN(1,1)  5.41545455  1.53172185  3.54  0.0004
UN(2,1)       2.71681818  1.09623989  2.48  0.0132
UN(2,2)       4.18477273  1.18363247  3.54  0.0004
UN(3,1)       3.91022727  1.41775367  2.76  0.0058
UN(3,2)       2.92715909  1.19304751  2.45  0.0141
UN(3,3)       6.45573864  1.82595863  3.54  0.0004
UN(4,1)       2.71022727  1.17209851  2.31  0.0208
UN(4,2)       3.31715909  1.12903016  2.94  0.0033
UN(4,3)       4.13073864  1.40356157  2.94  0.0033
UN(4,4)       4.98573864  1.41017984  3.54  0.0004
Residual      1.00000000  .          .    .
    
```

A key difference is now evident between PROC MIXED and PROC GLM. By default, PROC MIXED uses restricted maximum likelihood (REML) to estimate all unknown variance-covariance parameters (Jennrich and Schluchter 1986), whereas PROC GLM uses method of moments estimators on transformed data. The advantage of the likelihood-based approach of PROC MIXED is that it can accommodate data that are missing at random (Rubin 1976), in contrast with PROC GLM, which must ignore data from all subjects which have missing repeated measures.

The dental data have no missing records, and in this case REML and PROC GLM's method of moments produce identical estimates. PROC MIXED also offers regular maximum likelihood and MIVQUE0 estimation methods, the latter being a method of moments procedure that is also equivalent to PROC GLM's procedure for balanced data.

Output 7 displays the REML estimates of the three covariance structures. The compound symmetry structure has two unknown parameters and has the following form:

$$\begin{bmatrix} \sigma^2 + \sigma_1 & \sigma_1 & \sigma_1 & \sigma_1 \\ & \sigma^2 + \sigma_1 & \sigma_1 & \sigma_1 \\ & & \sigma^2 + \sigma_1 & \sigma_1 \\ & & & \sigma^2 + \sigma_1 \end{bmatrix}$$

Here the REML estimate of  $\sigma_1$  is 3.29 and that for  $\sigma^2$  is 1.98. The Huynh-Feldt or Type H structure has five unknown parameters in the following form:

$$\begin{bmatrix} \sigma_1 & \frac{\sigma_1 + \sigma_2}{2} - \lambda & \frac{\sigma_1 + \sigma_3}{2} - \lambda & \frac{\sigma_1 + \sigma_4}{2} - \lambda \\ & \sigma_2 & \frac{\sigma_2 + \sigma_3}{2} - \lambda & \frac{\sigma_2 + \sigma_4}{2} - \lambda \\ & & \sigma_3 & \frac{\sigma_3 + \sigma_4}{2} - \lambda \\ & & & \sigma_4 \end{bmatrix}$$

The REML estimates for  $\sigma_1$ - $\sigma_4$  are labeled VAR(1)-VAR(4) in Output 7 and that for  $\lambda$  is labeled HF. The unstructured matrix is the most general form possible and has ten unknown parameters:

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ & & \sigma_{33} & \sigma_{34} \\ & & & \sigma_{44} \end{bmatrix}$$

Their REML estimates are indexed as UN(row,column) in Output 7.

The standard errors for all of the REML estimates in Output 7 are asymptotically valid and are obtained from the inverse of the second derivative matrix of the restricted likelihood function. Such estimates of precision are another benefit of a likelihood-based approach, although the asymptotic Wald tests printed at the end of Output 7 can be unreliable in small samples.

How should you select one of these three structures? The Model Fitting Information tables in Output 8 present likelihood-based criteria which are particularly useful in comparing different covariance structure models. As one approach you can compare the information criteria of Akaike (AIC) and Schwarz (BIC) and select the model with the largest value. Both AIC and BIC favor the compound symmetry structure for these data.

**Output 8.** PROC MIXED REML Fitting Information

TYPE=CS:

Model Fitting Information for Y	
Description	Value
Observations	108.0000
Variance Estimate	1.9750
Standard Deviation Estimate	1.4054
REML Log Likelihood	-211.704
Akaike's Information Criterion	-213.704
Schwarz's Bayesian Criterion	-216.309
-2 REML Log Likelihood	423.4085
Null Model LRT Chi-Square	47.0823
Null Model LRT DF	1.0000
Null Model LRT P-Value	0.0000

TYPE=HF:

Model Fitting Information for Y	
Description	Value
Observations	108.0000
Variance Estimate	1.0000
Standard Deviation Estimate	1.0000
REML Log Likelihood	-210.860
Akaike's Information Criterion	-215.860
Schwarz's Bayesian Criterion	-222.373
-2 REML Log Likelihood	421.7206
Null Model LRT Chi-Square	48.7703
Null Model LRT DF	4.0000
Null Model LRT P-Value	0.0000

TYPE=UN:

Model Fitting Information for Y	
Description	Value
Observations	108.0000
Variance Estimate	1.0000
Standard Deviation Estimate	1.0000
REML Log Likelihood	-207.017
Akaike's Information Criterion	-217.017
Schwarz's Bayesian Criterion	-230.043
-2 REML Log Likelihood	414.0348
Null Model LRT Chi-Square	56.4560
Null Model LRT DF	9.0000
Null Model LRT P-Value	0.0000

You can also construct restricted likelihood ratio tests by subtracting the values of -2 REML Log Likelihood and comparing the result with a  $\chi^2$ -distribution with degrees of freedom equal to the difference in the number of parameters.

These tests are appropriate for structures which are nested within each other, or in other words, one is a special case of the other.

For example, the restricted likelihood ratio test between Type H and unstructured has a  $\chi^2$ -statistic equal to  $421.7206 - 414.0348 = 7.6858$ . Note that this equals the uncorrected Mauchly sphericity test statistic  $n_1 \log W$  shown after Output 2. Using  $10 - 5 = 5$  degrees of freedom, this test favors the Type H structure.

However, compound symmetry is a special case of Type H, and the restricted likelihood ratio test between these two structures has a  $\chi^2$ -statistic equal to  $423.4085 - 421.7206 = 1.6879$  on  $5 - 2 = 3$  degrees of freedom. With  $p > .1$  this test favors compound symmetry, agreeing with AIC and BIC.

Finally, the final three lines of the first table in Output 8 carry out a restricted likelihood ratio test of the compound symmetry covariance model versus the simple structure used in standard ordinary least squares ( $\sigma^2$  times an identity matrix). This test reveals that the compound symmetry model fits considerably better than the simple null model, and therefore compound symmetry is the best fitting structure among the ones considered.

Having selected the compound symmetry structure, you are now ready to consider tests of fixed effects. For comparison purposes, Output 9 displays these tests for all three of the covariance structures.

**Output 9.** PROC MIXED Tests of Fixed Effects

TYPE=CS:

Tests of Fixed Effects					
Source	NDF	DDF	Type III F	Pr > F	
GENDER	1	25	9.29	0.0054	
AGE	3	75	35.35	0.0001	
GENDER*AGE	3	75	2.36	0.0781	

TYPE=HF:

Tests of Fixed Effects					
Source	NDF	DDF	Type III F	Pr > F	
GENDER	1	25	9.39	0.0052	
AGE	3	75	35.35	0.0001	
GENDER*AGE	3	75	2.36	0.0781	

TYPE=UN:

Tests of Fixed Effects					
Source	NDF	DDF	Type III F	Pr > F	
GENDER	1	25	9.29	0.0054	
AGE	3	25	34.45	0.0001	
GENDER*AGE	3	25	2.93	0.0532	

The TYPE=CS and TYPE=UN Type III  $F$ -test for the between-subject effect GENDER is identical to that from Output 4. However, the TYPE=HF test is different.

This difference is due to the way PROC MIXED computes its  $F$ -statistics, which is not the same as PROC GLM. PROC

MIXED uses a general Wald-type quadratic form

$$F = \frac{\hat{\beta}'L[L'(X'\hat{V}^{-1}X)-L]^{-1}L'\hat{\beta}}{\text{rank}(L)}$$

where  $\hat{\beta}$  is the estimate of the fixed-effects parameter vector,  $L$  is a Type III coefficient matrix,  $X$  is the design matrix of the fixed effects, and  $\hat{V}$  is the estimated variance covariance matrix of the data (refer to SAS Institute Inc. 1992). In contrast, PROC GLM uses the traditional ratios of mean squares to construct its  $F$ -tests.

The advantage of the PROC MIXED method is that it can accommodate any type of covariance structure, but a disadvantage is that the denominator degrees of freedom for the tests must be assigned using other criteria. For REPEATED analyses, PROC MIXED obtains the degrees of freedom by partitioning the residual degrees of freedom into between- and within-subject parts and assigning them to between- and within-subject effects, respectively. For this example, the residual degrees of freedom equals  $108 - 8 = 100$ , and the between-subject component equals  $27 - 2 = 25$ .

For the within-subject effects AGE and GENDER\*AGE, the PROC MIXED  $F$ -tests using TYPE=CS and TYPE=HF are identical to the PROC GLM unadjusted univariate tests in Output 5. These tests are appropriate here because you cannot reject the hypothesis of sphericity.

The TYPE=UN PROC MIXED  $F$ -tests for AGE and AGE\*GENDER are not the same as any of the PROC GLM multivariate or adjusted univariate tests, even though they do take into account the fact that the within-subject variance-covariance matrix is unstructured. The PROC GLM test statistic most closely related to the  $F$ -statistic in PROC MIXED is the Hotelling-Lawley Trace. Wright (1995) performs a simulation study indicating that the  $F$ -test from PROC MIXED under TYPE=UN can be too liberal in small samples, whereas the Hotelling-Lawley Trace test does well, as does a modification of it due to McKeon (1974). Consequently, in Release 6.11 of PROC MIXED, you can specify the HLPS and HLM options on the REPEATED statement. These options only apply when you are using TYPE=UN, and their small-sample properties in missing data situations have yet to be investigated.

### Additional Analyses

Since AGE is a measure of time, it can often be informative to break its effect into polynomial contrasts. The following PROC GLM specification allows you to do this:

```
proc glm data=forglm;
  class gender;
  model y1-y4=gender / nouni ss1;
  repeated age 4 (8 10 12 14) polynomial /
    summary;
run;
```

The SS1 option on the MODEL statement requests Type I or sequential sums of squares, which are usually more appropriate than Type III when comparing polynomial effects. The POLYNOMIAL option sets up the contrasts, which for

this example consist of linear, quadratic, and cubic terms. The SUMMARY option produces univariate ANOVAs for each of them.

The comparable PROC MIXED code is as follows:

```
proc mixed data=formixed;
  class gender person;
  model y = gender|age|age|age / htype=1;
  repeated / type=un sub=person;
run;
```

Note that AGE has been dropped from the CLASS statement, making it a continuous variable, and it is included three times on the MODEL statement to construct the polynomial terms. These terms are not orthogonalized for this analysis, although you can do this by constructing orthogonal polynomial variables in the SAS data set and using them as MODEL effects.

### Output 10. Polynomial Contrast Results

General Linear Models Procedure Repeated Measures Analysis of Variance Analysis of Variance of Contrast Variables				
AGE.N represents the nth degree polynomial contrast for AGE				
Contrast Variable: AGE.1				
Source	DF	Type I SS	F Value	Pr > F
MEAN	1	235.35601852	99.45	0.0001
GENDER	1	12.11415194	5.12	0.0326
Error	25	59.16732955		
Contrast Variable: AGE.2				
Source	DF	Type I SS	F Value	Pr > F
MEAN	1	1.44675926	1.39	0.2497
GENDER	1	1.19954756	1.15	0.2935
Error	25	26.04119318		
Contrast Variable: AGE.3				
Source	DF	Type I SS	F Value	Pr > F
MEAN	1	0.38935185	0.15	0.6974
GENDER	1	0.67882997	0.27	0.6081
Error	25	62.91931818		

The MIXED Procedure Tests of Fixed Effects				
Source	NDF	DDF	Type I F	Pr > F
GENDER	1	25	9.29	0.0054
AGE	1	25	99.45	0.0001
AGE*GENDER	1	25	5.12	0.0326
AGE*AGE	1	25	1.39	0.2497
AGE*AGE*GENDER	1	25	1.15	0.2935
AGE*AGE*AGE	1	25	0.15	0.6974
AGE*AGE*AGE*GENDER	1	25	0.27	0.6081

Output 10 lists the relevant results from the PROC GLM and PROC MIXED analyses. The Type I  $F$ -tests for the linear, quadratic, and cubic contrasts are identical. A typical analysis usually considers the higher order terms first and then drops them from the model if they are not significant. The Type I sums of squares are constructed under the assumption that higher order terms are not present, and so they allow you to perform this backwards selection process using only one model fit.



One fact to keep in mind is that all of the Type I  $F$ -tests use variance-covariance estimates from the full mean model. In order to gain parsimony and power in your model, it is usually sensible to re-estimate these parameters under a reduced mean model. However, you cannot perform this reduction in PROC GLM, but you can in PROC MIXED.

For this example, it appears that the cubic and quadratic terms are not needed, and so the following PROC MIXED code fits the reduced model:

```
proc mixed data=formixed;
  class gender person;
  model y = gender|age / htype=1;
  repeated / type=un sub=person;
run;
```

Output 11 shows the updated tests of fixed effects. Note that both between- and within-subject effect tests are affected by the new estimate of the unstructured covariance matrix, and so PROC MIXED can potentially lead to different inferences than those based on PROC GLM.

**Output 11. PROC MIXED Linear Contrast Results**

Tests of Fixed Effects				
Source	NDF	DDF	Type I F	Pr > F
GENDER	1	25	9.40	0.0051
AGE	1	25	116.81	0.0001
AGE*GENDER	1	25	7.40	0.0117

In addition to being more flexible than PROC GLM in its mean specification, PROC MIXED can be more informative in carrying out final inferences on fixed effects. Such inferences are usually the goal of the entire analysis, as shown in Figures 1 and 2.

Returning to the case where AGE is considered a classification variable, suppose you wish to compare GENDER means adjusted for AGE and AGE\*GENDER and averaged across the repeated measures. You can add

```
lsmeans gender / pdiff;
```

to both PROC GLM and PROC MIXED specifications, but the results differ considerably, as shown in Output 12.

In contrast to PROC MIXED, PROC GLM does not average across the repeated measures and does not compute standard errors accounting for the appropriate covariance structure. Furthermore, you can allow AGE to be a continuous effect in a PROC MIXED LSMEANS analysis but you cannot in PROC GLM.

## Summary

Table 1 summarizes this paper's comparison of the GLM and MIXED procedures; it is organized roughly according to Figures 1 and 2. To conclude, PROC GLM provides more extensive results for the traditional univariate and multivariate approaches to repeated measures. However, PROC MIXED offers you a richer class of both mean and variance-covariance models, and you can apply these to more general data structures and obtain more general inferences on the fixed effects.

**Output 12. GENDER LSMEANS Results**

General Linear Models Procedure Least Squares Means			
GENDER	Y1 LSMEAN	Pr >  T	H0: LSMEAN1=LSMEAN2
F	21.1818182		
M	22.8750000	0.0750	
GENDER	Y2 LSMEAN	Pr >  T	H0: LSMEAN1=LSMEAN2
F	22.2272727		
M	23.8125000	0.0590	
GENDER	Y3 LSMEAN	Pr >  T	H0: LSMEAN1=LSMEAN2
F	23.0909091		
M	25.7187500	0.0141	
GENDER	Y4 LSMEAN	Pr >  T	H0: LSMEAN1=LSMEAN2
F	24.0909091		
M	27.4687500	0.0007	

The MIXED Procedure Least Squares Means					
Level	LSMEAN	Std Error	DDF	T	Pr >  T
GENDER F	22.64772727	0.58613896	25	38.64	0.0001
GENDER M	24.96875000	0.48600075	25	51.38	0.0001
Differences of Least Squares Means					
Level 1	Level 2	Difference	Std Error	DDF	T
GENDER F	GENDER M	-2.32102273	0.76141685	25	-3.05
Differences of Least Squares Means					
Pr >  T					
0.0054					

## Acknowledgements

We'd like to thank Randy Tobias for insightful comments and Mike Cybrynski for help with the process flow diagrams.

## Appendix

This appendix shows how you can reproduce the sphericity test calculations from PROC GLM using PROC IML and PROC MIXED. The first step is to select a contrast transformation of the data, which reduces the number of repeated measures by one. PROC GLM offers you several such transformations, the default being the contrast of the first 3 levels of AGE with the last level. This one is used in the following IML and DATA step statements:

Table 1: Summary of the Repeated-Measures Comparison

<b>PROC GLM</b>	<b>PROC MIXED</b>
Requires balanced data; ignores subjects with missing observations	Allows data that are missing at random
Handles between- and within-subject effects differently with regard to syntax and tests	Handles between- and within-subject effects similarly
Requires a dimension-reducing orthogonal transformation for the repeated measures variables	Analyzes the data in their original form
Assumes a full ANOVA (cell means) model for within-subject effects	Allows a full ANOVA and/or a reduced mean model for within-subject effects
Assumes covariates are constant within a subject	Allows covariates to vary within a subject
Automatically performs a sphericity test with the PRINTE option	Can produce sphericity test results either by running both TYPE=UN and TYPE=HF or by using TYPE=UN on transformed data
Assumes either a Type H or unstructured within-subject covariance matrix	Allows a wide variety of within-subject covariance structures, including CS, AR(1), HF, FA, UN, spatial, and random coefficients
Estimates covariance parameters using a method of moments	Estimates covariance parameters using restricted maximum likelihood, maximum likelihood, and MIVQUE0
Is computationally fast and prints all significance tests in one run	Can be computationally intensive and requires different runs for different covariance structures
Computes $F$ -statistics that are ratios of mean squares	Computes $F$ -statistics that are Wald-type quadratic forms
Computes standard, G-G, and H-F univariate repeated measures tests	Computes only standard univariate repeated measures tests (using TYPE=CS or TYPE=HF)
Computes four multivariate repeated measures tests: Wilk's Lambda, Pillai's Trace, Hotelling-Lawley Trace, and Roy's Greatest Root	Computes a Wald-type $F$ (using TYPE=UN) and two versions of Hotelling-Lawley Trace
Computes LSMEANS only for each separate variable	Computes LSMEANS which are averaged across repeated measures and whose standard errors reflect the appropriate covariance structure

```

proc iml;
  t={1 0 0 -1,
      0 1 0 -1,
      0 0 1 -1};
  call gsorth(p1,p2,l,t`);
  m=p1`;
  varname='m1' || 'm2' || 'm3' || 'm4';
  create orth from m [colname=varname];
  append from m;
  close orth;
quit;

```

```

data trans;
  array v[4] v1-v4;
  do i = 1 to 4;
    set formixed;
    v[i] = y;
  end;
  do i=1 to 3;
    set orth point=i;
    age=i;
    y = v1*m1 + v2*m2 + v3*m3 + v4*m4;
    output;
  end;
  keep person gender age y;
run;

```

You should next execute the following statements:

```

proc mixed data=trans;
  class gender age person;
  model y = gender|age;
  repeated / type=un sub=person;
run;

```

Output 13 displays the model fitting information from this PROC MIXED run. The Null Model LRT Chi-Square value equals  $n_1 \log W$  and the Null Model LRT DF is 5.

**Output 13.** PROC MIXED REML Fitting Information for the Transformed Data

Model Fitting Information for Y	
Description	Value
Observations	81.0000
Variance Estimate	1.0000
Standard Deviation Estimate	1.0000
REML Log Likelihood	-135.855
Akaike's Information Criterion	-141.855
Schwarz's Bayesian Criterion	-148.808
-2 REML Log Likelihood	271.7105
Null Model LRT Chi-Square	7.6858
Null Model LRT DF	5.0000
Null Model LRT P-Value	0.1744

Finally, you can also reproduce the Mauchly test by using regular maximum likelihood with the METHOD=ML option on the PROC MIXED statement. In this case just change the correction factor to

$$\rho_M = n_1/n - (2p^2 + p + 2)/(6pn_1)$$

where  $n$  is the number of subjects.

SAS is a registered trademark or trademark of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

## References

- Anderson, T.W. (1958), *An Introduction to Multivariate Statistical Analysis*, New York: John Wiley & Sons, Inc.
- Greenhouse, S.W. and Geisser, S. (1959), "On Methods in the Analysis of Profile Data," *Psychometrika*, 32, 95-112.
- Huynh, H. and Feldt, L.S. (1970), "Conditions under Which Mean Square Ratios in Repeated Measurements Designs Have Exact F-Distribution," *Journal of the American Statistical Association*, 65, 1582-1589.
- Huynh, H. and Feldt, L.S. (1976), "Estimation of the Box Correction for Degrees of Freedom from Sample Data in the Randomized Block and Split Plot Designs," *Journal of Educational Statistics*, 1, 69-82.
- Jennrich, R.I., and Schluchter, M.D. (1986), "Unbalanced Repeated-Measures Models with Structured Covariance Matrices," *Biometrics*, 42, 805-820.
- Latour, D., Latour, K., and Wolfinger, R.D. (1994), *Getting Started with PROC MIXED*, Software Sales and Marketing Department, SAS Institute Inc., Cary, NC.
- Mauchly, J.W. (1940), "Significance Test for Sphericity of a Normal n-Variate Distribution," *The Annals of Mathematical Statistics*, 11, 204-209.
- McKeon, J.J. (1974), "F approximations to the distribution of Hotelling's  $T_0^2$ ," *Biometrika*, 61, 381-383.
- Muller, K.E., LaVange, L.M., Ramey, S.L., and Ramey, C.T. (1992), "Power Calculations for General Linear Multivariate Models Including Repeated Measures Applications," *Journal of the American Statistical Association*, 87, 1209-1226.
- Rubin, D.B. (1976), "Inference and missing data," *Biometrika*, 63, 581-592.
- SAS Institute Inc. (1989), *SAS/STAT User's Guide: Version 6, Fourth Edition, Volume 2*, Cary NC: SAS Institute Inc.
- SAS Institute Inc. (1992), SAS Technical Report P-229, *SAS/STAT Software: Changes and Enhancements, Release 6.07*, Cary NC: SAS Institute Inc.
- Searle, S.R. (1971), *Linear Models*, New York: John Wiley & Sons, Inc.
- Wolfinger, R.D. (1993), Covariance structure selection in general mixed models. *Communications in Statistics, Simulation and Computation*, 22(4), 1079-1106.
- Wolfinger, R.D., Tobias, R.D., and Sall, J. (1991), "Mixed Models: A Future Direction," *Proceedings of the Sixteenth Annual SAS Users Group Conference*, SAS Institute Inc., Cary, NC, 1380-1388.
- Wright, P.S. (1995), "Adjusted F tests for repeated measures with the MIXED procedure," to appear in *Proceedings of the Twentieth Annual SAS Users Group Conference*, SAS Institute Inc., Cary, NC.