Conference on Applied Statistics in Agriculture   2000 - 12th Annual Conference Proceedings

A SIMULATION STUDY TO EVALUATE PROC MIXED ANALYSIS OF REPEATED MEASURES DATA

LeAnna Guerin

Walter W. Stroup

A SIMULATION STUDY TO EVALUATE PROC MIXED ANALYSIS
OF REPEATED MEASURES DATA

by LeAnna Guerin and Walter W. Stroup
Department of Biometry, University of Nebraska, Lincoln, NE 68583-0712

1. Abstract

Experiments with repeated measurements are common in pharmaceutical trials, agricultural research, and other biological disciplines. Many aspects of the analysis of such experiments remain controversial. With increasingly sophisticated software becoming available, e.g. PROC MIXED, data analysts have more options from which to choose, and hence more questions about the value and impact of these options. These dilemmas include the following. MIXED offers a number of different correlated error models and several criteria for choosing among competing models. How do the model selection criteria behave? How is inference affected if the correlated error model is misspecified? Some texts use random between subject error effects in the model in addition to correlated errors. Others use only the correlated error structure. How does this affect inference? MIXED has several ways to determine degrees of freedom, including a new option to use Kenward and Roger's procedure. The Kenward-Roger procedure also corrects test statistics and standard errors for bias. How do the various degree-of-freedom options compare? When is the bias serious enough to worry about and how well does the Kenward-Roger option work? Some models are prone to convergence problems. When are these most likely to occur and how should they be addressed? We present the results of several simulation studies conducted to help understand the impact of various decisions on the small sample behavior of typical situations that arise in animal health and agricultural settings.

2. Introduction

Repeated measures experiments are used in a wide variety of applications, including animal health and most agricultural disciplines. Data analysts use several methods to analyze repeated measures data, also known as longitudinal data. Before the advent of PROC MIXED, two of the most common methods used were univariate analysis of variance (ANOVA) and multivariate analysis of variance (MANOVA). Univariate ANOVA is often called "split-plot in time" analysis because of its similarity to the analysis of split-plot experiments. Split-plot-in-time ANOVA assumes independent errors, an unrealistic assumption in most longitudinal data. Pre-PROC MIXED statistical computing packages for ANOVA, e.g. SAS PROC GLM, have limited corrections, such as adjustments to p-values, but lack systematic facilities to handle correlated errors. MANOVA allows for correlated errors among repeated measures, but is beyond the statistical training level of most biological researchers. More importantly, MANOVA's assumed correlated error structure leads to inefficient estimates and low power, and its ability to handle missing data is limited and draconian. Beginning the 1990's, comprehensive mixed model software, notably SAS PROC MIXED, became available. PROC MIXED allows one to use a comprehensive array of correlated error models. For this reason, it has become a standard tool for

Applied Statistics in Agriculture

analyzing longitudinal data.

However, PROC MIXED's versatility comes at a price. To use PROC MIXED effectively, one must

▸ select an appropriate correlated error model.

▸ select an appropriate method to approximate denominator degrees for freedom for hypothesis testing and interval estimation.

▸ depending on the correlated error structure, adjust standard errors and test-statistics to account for the fact that MIXED defaults are biased.

Recently, SAS released Version 8, which makes more options available than in past versions of PROC MIXED. For choosing an appropriate correlated error model, there are four model selection criteria available in Version 8. Keselman et al. (1998) compared the two criteria available in the previous version (release 6.12), the Akaike Information Criterion (AIC) and the Schwarz Bayesian Information Criterion (BIC). In their simulation study, they found that 1) neither the AIC nor BIC consistently chose the right model, 2) the AIC had a somewhat higher percentage of correct choices, but also noted that the BIC was incorrectly computed. Version 8 corrects the BIC. Also, Keselman et. al. noted that many of the correlated error models they used were similar and they did not distinguish between AIC or BIC choosing a model that was wrong but close versus a model that was wrong and not close. They speculated that choosing an incorrect correlated error model that is similar in structure to the "true" model would not seriously affect inference, but did not investigate this issue.

The standard approach to computing test statistics and standard errors in mixed model analysis is to substitute estimated variance and covariance components into the theoretical form of the information matrix (see section 3 below for a more complete explanation). Kacker and Harville (1984) found that this approach produces downward bias in standard errors and upward bias in test statistics, except in the case of balanced data in conjunction with i.i.d. random effects. Kenward and Roger (1997) obtained a bias correction based on a Taylor Series approximation. The Kenward-Roger correction is available in Version 8. Kenward and Roger also obtained an approximation for denominator degrees of freedom that is more general than the Satterthwaite approximation available in older versions of MIXED. Both degree-of-freedom approximations are available in Version 8, as well as the default "containment" method.

The purpose of this paper is to extend the work begun by Keselman, et. al. to address a number of questions about the small sample behavior of mixed model analysis of repeated measures data. A simulation study was conducted to answer the following specific questions:

1. How do the four model selection criteria computed by Version 8 of SAS PROC MIXED compare in terms of their ability to "choose the right" correlated error model?

2. How do Version 8 options to approximate degrees of freedom -- default, Satterthwaite, and Kenward-Roger -- compare in terms of their ability to control Type I error?

3. The Kenward-Roger option also adjusts standard errors and test-statistics. How does this procedure perform?

4. How robust is mixed model analysis to misspecification of the correlated error model?

This paper is organized as follows. Section 3 reviews repeated measures designs and pertinent theory and methods for their analysis using mixed models. Section 4 describes how the simulation study was designed and conducted. Section 5 discusses the simulation study results.

Section 6 summarizes the main conclusions and topics for future work.

3.   Background: Repeated Measures Design and Analysis using Mixed Models

Repeated measures designs have the following general structure, illustrated in Figure 1:

▸ There are 2 or more treatments. Let I (≥ 2) denote the number of treatments.

▸ Experimental units (subjects) are randomly assigned to each treatment. The number of subjects per treatment, denoted J_i, i=1, 2,..., I, need not be equal. Subjects may be assigned to treatments using any reasonable design - completely randomized, complete or incomplete block, row-column, etc. For simplicity, the simulation study in this paper uses a completely randomized design.

▸ Each subject is observed at each of K times. The times need not be equally spaced, but are typically regularly spaced. Often, their timing reflects the biology of the subjects under study, e.g. growth stage.

The following graph suggests the typical focus of repeated measures experiments. In this example, 2 treatments,"test" and "placebo" are compared.
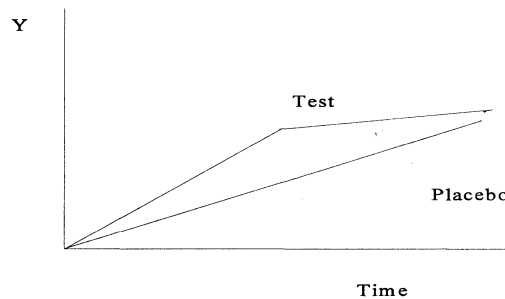
Of interest are:

▸ Treatment x time interaction
   Are changes over time the same (parallel on the graph) for all treatments?

▸ Time effects
   Assuming negligible treatment x time interaction, how does the mean response change over time? Assuming non-negligible treatment x time interaction, how does response change over time for each treatment?

▸ Treatment effects
   How do the mean responses to treatments differ? Again, this may be averaged over all times (treatment main effects) assuming negligible treatment x time interaction, or specific to each time (simple effects of treatment given time), otherwise.

Assuming experimental units are assigned to treatments using a completely randomized design, the model equation is:

$$y_{ijk} = \mu + \alpha_i + s_{ij} + \tau_k + \gamma_{ik} + e_{ijk},$$

(3.1)

where y_{ijk} denotes the observation for the jth subject, or experimental unit, on the ith treatment at the kth time, μ, α_i, τ_k, and γ_{ik} are the intercept, treatment main effect, time main effect, and

Applied Statistics in Agriculture

treatment x time interaction, respectively, s_{ij} is the ij^{th} subject effect, assumed i.i.d. $N(0,\sigma_S^2)$, and e_{ijk} is within subjects error, i.e. random variation among repeated measurements on each subject. Within subjects errors are potentially correlated. That is, the vector $\mathbf{e_{ij}}$, defined as $\mathbf{e_{ij}}' = (e_{ij1}, e_{ij2}, ..., e_{ijK})$ is assumed to be distributed $MVN(\mathbf{0}, \Sigma)$, where Σ is the covariance matrix among the e_{ijk} within the ij^{th} subject. The vectors $\mathbf{e_{ij}}$ are assumed to be mutually independent. Also, $\mathbf{e_{ij}}$ and $\mathbf{s_{ij}}$ are assumed to be independent. If a more complex design is used to assign subjects to treatments, it is reflected in the model, e.g. by adding a block effect if a randomized block design is used.

Alternatively, one can express (3.1) by replacing the intercept, time, and treatment effects with a treatment-time cell mean, $\mu_{ik} = \mu + \alpha_i + \tau_k + \gamma_{ik}$. Thus, the *observations*, $\mathbf{y_{ij}}' = (y_{ij1}, y_{ij2}, ... y_{ijK}) \sim MVN(\mathbf{\mu_i}, J\sigma_S^2 + \Sigma)$, where $\mathbf{\mu_i}'$ is the vector of means at the K times for the ith treatment, i.e. $\mathbf{\mu_i}' = (\mu_{i1}, \mu_{i2}, ... , \mu_{iK})$. Among the more common covariance models for correlated errors are:

▸ Independent Errors, $\Sigma = I\sigma^2$.

▸ Compound Symmetry (CS),

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & \rho & . & . & \rho \\ \rho & 1 & \rho & \rho & . & \rho \\ . & . & . & . & . & . \\ \rho & . & . & . & \rho & 1 \end{bmatrix}$$

where ρ= correlation between observations on the ij^{th} subject. Note that the compound symmetry model can be reexpressed as $\Sigma^* = J\sigma_S^2 + I\sigma^2$, and hence the correlation among repeated measurements is the same as the interclass correlation $\dfrac{\sigma_s^2}{\sigma^2 + \sigma_s^2}$ in the independent errors model. Thus, σ_S^2 in the independent errors model and ρ in the compound symmetry model are confounded; the two are actually equivalent expressions of the same model.

▸ First-order Autoregressive [AR(1)]

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & . & . & \rho^{k-1} \\ \rho & 1 & \rho & \rho^2 & . & \rho^{k-2} \\ . & . & . & . & . & . \\ \rho^{k-1} & . & . & . & \rho & 1 \end{bmatrix}$$

AR(1) assumes correlation between errors (e_{ijk}) w time periods apart is ρ^w.

▶ Toeplitz (TOEP)

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdot & \cdot & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \rho_2 & \cdot & \rho_{k-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_{k-1} & \cdot & \cdot & \cdot & \rho_1 & 1 \end{bmatrix}$$

▶ First-order Antedependence [ANTE(1)]

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_1 & \sigma_1\sigma_3\rho_1\rho_2 & \cdot & \cdot & \sigma_1\sigma_k\rho_1\rho_2\cdots\rho_{k-1} \\ \cdot & \sigma_2^2 & \sigma_2\sigma_3\rho_2 & \sigma_2\sigma_4\rho_2\rho_3 & \cdot & \sigma_2\sigma_k\rho_1\rho_2\cdots\rho_{k-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \sigma_{k-1}\sigma_k\rho_{k-1} & \sigma_k^2 \end{bmatrix}$$

▶ Unstructured (UN)

$\Sigma = [\sigma_{ij}]$, where $\sigma_{ii} = \sigma_i^2$. Note that the unstructured covariance model is conceptually similar to the correlation structure assumed when one uses MANOVA to analyze longitudinal data, the distinction being that in MANOVA, the covariance matrix among the *observations*, y_{ijk}, is unstructured, whereas with $\Sigma=[\sigma_{ij}]$ the covariance among the *errors* is unstructured.

Note that ANTE(1) and UN allow for heterogeneous variances at each time of observation. There are modifications of CS, AR(1), and TOEP that also allow for heterogeneity at each time. Equally spaced times of observation are implicitly assumed for AR(1) and TOEP, whereas CS, ANTE(1), and UN allow for unequal spacing.

The repeated measures model (3.1) is a special case of the mixed model

$$\mathbf{y} = X\boldsymbol{\beta} + Z\mathbf{U} + \mathbf{e}$$

where $\begin{bmatrix} U \\ e \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix} \right)$. Hence, $\mathbf{y} \sim MVN(X\boldsymbol{\beta}, V)$, where $V=ZGZ'+R$.

X describes the treatment-time design, $\boldsymbol{\beta}$ is the vector of fixed treatment-time effects, Z describes the between-subjects design, $\mathbf{u}$ is the vector of random subject effects and $\mathbf{e}$ is the vector of random errors. Consistent with (3.1), $G=I\sigma_S^2$ and R is block diagonal, with each block equal to Σ, the within subject covariance matrix described above.

PROC MIXED obtains estimates of $\boldsymbol{\beta}$ and $\mathbf{U}$ by solving the mixed model equations:

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \end{bmatrix} \begin{bmatrix} b \\ u \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix} ,$$

where **b** and **u** denote the solutions for β and **U**, respectively. Note that the solution for β is equal to the generalized least squares (GLS) solution, $\mathbf{b} = (XV^{-1}X)^{-}X'V^{-1}\mathbf{y}$.

Inference for the mixed model is based on **predictable functions**, i.e. functions of the form K'β + M'**U**, where K'β is an **estimable function**. In this paper we focus exclusively on inference on estimable functions, i.e. cases where M=0. For known G and R, inference proceeds as follows:

▸   **Confidence Interval**: use the formula $K'b \pm Z_{\alpha} \sqrt{K'(X'V^{-1}X)^{-}K}$, where Z_{α} is the

standard normal table value for the 1-α level of confidence.

▸   **Test H$_0$: K'β = 0**: use the Wald statistic, $(K'b)' [K(XV^{-1}X)^{-}K]^{-1} (K'b)$, which is approximately $\sim \chi^2_{rank(K)}$.

For unknown G and R, PROC MIXED substitutes **restricted maximum likelihood** (REML) estimates of the variance and covariance components, e.g. REML estimates of σ_S^2 and the components of Σ in repeated measures models, directly into the mixed model equations and the formula for prediction error variance. Inference proceeds as follows.

▸   **Confidence Interval**: use $K'b \pm t_{(\alpha, \, \nu)} \sqrt{K'(X'\hat{V}^{-1}X)^{-}K}$, where ν = d.f. to estimate

K'(X'V^{-1}X)$^{-}$ K. Note that ν may be "obvious" by inspection or require approximation.

▸   **Test H$_0$: K'β = 0**: use the Wald statistic divided by its degrees of freedom,

$$\frac{(K'b)' \, [K'(X'\hat{V}^{-1}X)^{-}K]^{-1} \, (K'b)}{rank(K)}$$, which is approximately $\sim F_{[rank(K), \, \nu]}$.

Kacker and Harville (1984) showed that, except for *balanced, variance components only* models, these procedures are biased. Specifically, standard errors obtained from K'(X'$\hat{V}^{-1}$X)$^{-}$ K underestimate the true standard errors based on known K'(X'V^{-1}X)$^{-}$ K. Therefore, except for balanced, independent errors (and hence compound symmetry) models, PROC MIXED computes standard errors that are biased downward and test statistics that are biased upward for repeated measures experiments. The bias can be exacerbated by misspecifying Σ.

Kenward and Roger (1997) obtained a bias correction for standard errors and test statistics that is implemented in SAS Version 8. Also, Version 8 computes two approximations for the denominator degrees of freedom, ν, in addition to the default containment method: the Satterthwaite procedure and the more general procedure of Kenward and Roger. A major focus of the simulation study was to assess the severity of the bias caused by using estimated variance components for various Σ and for various misspecifications of Σ, and the Kenward-Roger procedure's effectiveness in correcting bias.

Because results can be affected by choice of covariance model, Σ, an important aspect of mixed model analysis of repeated measures data involves choosing a suitable Σ. Selection is based on evaluating the restricted likelihood. One can either compare covariance structures that constitute nested subsets using likelihood ratio tests (see Littell, et. al. (1996) for details) or compare model fitting criteria. Model fitting criteria are obtained by applying a penalty function to the restricted likelihood to account for differences in the number of parameters for competing

Σ. Version 8 computes four information criteria:
1. AIC: Akaike (1974) information criterion
2. BIC: Schwarz (1978) Bayesian information criterion (BIC):
3. CAIC: Bozdogan (1987) "corrected" Akaike information criterion
4. HQIC: Hannan and Quinn (1979) information criterion

SAS on-line documentation for Version 8 (1999) gives computational details for these information criteria. As noted earlier, Keselman et. al. (1998) compared AIC and BIC performance in SAS Release 6.12, but in Version 8, two new criteria have been added and the computation of BIC has been corrected. An objective of this study was to re-evaluate model selection criteria performance in view of the changes now available.

A final issue concerns the estimation of the variance-covariance components of G and R. As noted above, the repeated measures, the general form of V($\mathbf{y}$) = ZGZ' + R is block diagonal ($J\sigma_S^2 + \Sigma$). For some structures of Σ, notably CS and UN, the components of Σ contain the between subject variance σ_S^2. In these cases, model (3.1) reduces to

$$y_{ijk} = \mu + \alpha_i + \tau_k + \gamma_{ik} + e_{ijk}, \tag{3.2}$$

where the e_{ijk} are correlated normal random variables as described above. In other covariance structures, including all those described above except CS and UN, the components of Σ are at least in theory identifiable from σ_S^2 so that the covariance components of Σ and the between subjects variance component σ_S^2 can be estimated separately. That is, in PROC MIXED terms, separate REPEATED statement and RANDOM statement for subjects can coexist in the same MIXED program. In practice, identifiability problems often prevent the REML algorithm from converging when this is done. Some covariance models are more prone to this than others. A typical strategy for dealing with this is to drop s_{ij} from model (3.1) and use (3.2) instead. However, this strategy risks undesirable impact on the estimate of K'(X'V⁻¹X)⁻ K and hence on statistical inference. This study examined cases where non-convergence was a problem. In most cases, non-convergence was a result of MIXED defaults and could be circumvented. We were thus able to assess the impact of dropping s_{ij} from the model when it is, in theory, not necessary to do so.

4. Materials and Methods: The Simulation Study

A simulation study of PROC MIXED's mixed model analysis of repeated measures data was conducted to address the questions listed above at the end of Section 2. The study used a simulated experiment consisting of 2 treatments, 6 subjects per treatment, and all subjects observed at each of 6 time periods. The treatments were arranged in a completely randomized design (CRD) with repeated measures, depicted in Figure 1. For each of the 72 observations, correlated multivariate normal data were generated according to model (3.1). The data were generated using 13 scenarios involving three covariance structures with various covariance parameter values as follows:

a. **First-Order Autoregressive [AR(1)]**. Without loss of generality, the error variance was set to σ^2=1. There were six scenarios, representing a factorial combination of two autocorrelations (ρ=0.25 and ρ=0.75) and three subject variances (σ_S^2=0.25, 1.00 and 4.00).

b. Heterogeneous First-Order Autoregressive [ARH(1)]. The variances of the observations at the 6 time periods were $\sigma^2 = (\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2, \sigma_5^2, \sigma_6^2) = (1.50, 1.25, 1.00, 0.50, 0.25, 0.10)$. Again, there were six scenarios representing a factorial combination of two autocorrelations ($\rho = 0.25$ and $\rho = 0.75$) and three subject variances ($\sigma_s^2 = 0.025$, 1.0 and 6.0). The σ_s^2 were selected by taking $0.25 \times$ the minimum σ_k^2, 1.0, and $4 \times$ the maximum σ_k^2.

c. First-Order Antedependence [ANTE(1)]. The variances at the six time periods were the same as for the ARH(1), i.e. $\sigma^2 = (1.50, 1.25, 1.00, 0.50, 0.25, 0.10)$. The correlations were $\rho = (\rho_1, \rho_2, \rho_3, \rho_4, \rho_5) = (0.80, 0.60, 0.60, 0.40, 0.40)$. The subject variance, σ_s^2, was set to zero.

Figure 2 gives a summary of the 13 scenarios. A total of 500 simulated data sets were generated for each scenario.

SAS (Version 8) PROC IML code was written to generate the data. The covariance matrix Σ, was defined according to the above scenarios. A 72x1 vector of standard normal random deviates were generated using SAS's RANNOR function. Denote the vector $\mathbf{e}_{ij}` = (e_{ij1}, e_{ij2}, e_{ij3}, e_{ij4}, e_{ij5}, e_{ij6}) \sim MVN(0,1)$, and $\mathbf{e}' = (e'_{11}, e'_{12}, \ldots, e'_{26})$. A 12x1 vector of subject effects $\mathbf{s}$, was generated as $\mathbf{s} = \sigma_s \mathbf{w}$, where the elements of $\mathbf{w}$ were generated using the RANNOR function. Then, the vector of observations was calculated as $\mathbf{y} = \mathbf{s} + \Sigma^{1/2}\mathbf{e}$, where $\Sigma^{1/2}$ is the Cholevsky decomposition of Σ.

Each simulated data set was analyzed with each of eight different covariance structures, CS, CSH, AR(1), ARH(1), TOEP, TOEPH, ANTE(1) and UN, using PROC MIXED. Three denominator degree of freedom calculations, containment (SAS default), Satterthwaite's approximation and the Kenward & Roger adjustment, were calculated for each of the eight covariance structure analyses. Thus, there were 24 analyses per data set.

Analyses using PROC MIXED were compared using the following criteria:
1. Type I error rates of overall F-tests for Ho: no Trt main effect, Ho: no Time main effect, and Ho: no Trt x Time.
2. Tests of simple effects, Time 1 vs. Time k for k=(2,4,6) on Treatment 1, and Treatment 1 vs. Treatment 2 for time k=(1,3,4,6).
3. Standard errors for each of the simple effects. (not discussed here - to appear in a future paper)
4. Estimates of the covariance parameters. (also not discussed here)
5. Information Criteria: AIC, BIC, HQIC and CAIC.

Of particular interest is the observed Type I error rates using each of the denominator degree of freedom calculations, and how often each of the four information criteria choose the "true" covariance structure.

5. Simulation Results and Discussion

The discussion of the results is divided into three subsections. Subsection 5a addresses convergence for the ARH(1) model; 5b compares the performance of the model selection criteria; 5c discusses the Type I error rates for the various analyses. For purposes of this paper, not all 13 scenarios will be discussed. We have chosen representative scenarios that best illustrate our main findings for discussion here.

5a.   Convergence

As noted in section 3, in theory the parameters of $\text{Var}(\mathbf{e}) = \Sigma$ and σ_s^2 are identifiable for covariance structures other than CS and UN. The repeated measures analyses for simultaneous estimation of Σ and σ_s^2 in this study can be implemented by the following *example* SAS code:

```
proc mixed;                                                    (5a.1)
    class trt subj time;
    model y=trt time trt*time/ddfm=kr;
      /* or satterthwaite ddfm, containment ddfm, etc. as desired */
    random subj(trt);
    repeated time/type=ARH(1) subject=subj(trt);
```

For covariance types ANTE(1), ARH(1), and TOEPH, we encountered convergence problems when using this approach. For these types, *computational* identifiability can sometimes be a problem, despite the existence of *theoretical* identifiability. We focused on the ARH(1) model, although the approaches shown here can be adapted for use with the other types. Figure 3 shows a typical example. Note that convergence has failed because of an "infinite likelihood" and one or more estimates of σ_k^2 are effectively zero.

This problem can be avoided one of two ways. First, one can drop σ_s^2 from the model, i.e. drop the "random subj(trt)" from the PROC MIXED statements. This has been a typical approach for PROC MIXED users, as it results in 100% convergence. However, results in Section 5c suggest that one should avoid dropping σ_s^2 except as a last resort. Alternatively, one can use a lower bound to prevent estimates of the σ_k^2 from approaching zero too closely. We thank Russ Wolfinger (SAS Institute, Inc.) for this suggestion. Add the following SAS statement to the above code:

```
    parms 0 1 1 1 1 1 1 0/lowerb=1e-4, 1e-4, 1e-4, 1e-4, 1e-4, 1e-4;     (5a.2)
```

Note that the starting values shown here are comparable to the OLS parms suggested by SAS documentation as a possible way to avoid non-convergence. Figure 4 shows the output for this method. Note here that convergence was met, but the test statistics, standard errors and degrees of freedom are nonsensical. The problem here is that the estimates of σ_k^2 are still too close to the lower bound, resulting in certain matrices used by the Kenward-Roger adjustment being ill-conditioned. Moving the lower bound further from zero avoids this problem. For example, Figure 5 shows output when using a lower bound of 1e-3. The test statistics, standard errors and degrees of freedom are much more reasonable. In practice, data analysts would find the optimal lower bound -- i.e. the closest to zero that avoids ill-conditioning problems like those in Figure 4 -- through trial and error. However, for this simulation study, in the interest of time we used a lower bound of 1e-3 throughout. Also to save time, the PARMs starting values were set to the actual variance components used in generating the data. Note, however, that parameter estimates, test statistics, standard errors etc. were robust to starting values; convergence *mainly* depended on the lower bound, and *not* on starting values. Several data sets were examined more closely: alternative starting values resulted in at most a few more iterations, but never in failure to converge as long as the lower bound was used. Final REML estimates were unaffected by

starting values.

Table 1 shows percent convergence rates for ARH(1) analyses. The first three columns depict the 13 scenarios which were generated. The fourth column shows the percent convergence rates for the PROC MIXED default estimation procedure when including the 'random subj(trt)' (example code 5a.1). The default uses MIVQUE(0) starting values. The fifth column gives percent convergence rates when the Parms/OLS starting values are used instead of the default. The last column gives percent convergence rates when including the "random subj(trt)" and the statement "parms/lowerb=1e-3,.....,1e-3;" where the parms starting values were set to the values originally generated. The default and OLS starting values do not attain near 100% convergence, even as the subject variation decreases. However, when the parms/lowerb statement is added with a reasonable lower bound, all 13 ARH(1) analyses yielded 100% convergence.

5b.   Model Selection Criteria

Each of the four model selection criteria calculated in SAS (AIC, BIC, HQIC, and CAIC) were analyzed for all data sets generated to determine which of the eight covariance structure analyses the criteria chose most often. Table 2 gives the percent of data sets chosen by each criteria for the AR(1) generated data sets. In the interest of space, only the $\sigma_S^2=0.25$ and $\sigma_S^2=4$ scenarios are given. Note that in this table the "random subj(trt)" and "parms/lowerb=1e-3,....,1e-3" statements were used for each of the ARH(1) analyses. When the autocorrelation coefficient is $\rho=0.25$, the four criteria tend to choose CS most often (33.6% to 67.0% of data sets chosen), and AR(1), the true model, second most often. When $\rho=0.75$ the criteria choose AR(1), the true model, most often (55.8% to 91.8% of data sets chosen). For this scenario, the AIC and HQIC choose ARH(1) second most often while the BIC chooses TOEP second most often. The CAIC is noticeably more conservative throughout, with its stronger penalty. The CAIC chooses CS second most often when $\rho=0.75$.

Table 3 gives the percent of data sets for which correlation models were chosen when the data were generated as ARH(1). Again, in the interest of space, only the scenarios $\sigma_S^2=0.25$ and $\sigma_S^2=4$ are given. In this comparison, the "random subj(trt)" was dropped from the PROC MIXED statements for the ARH(1) analyses. When the generated subject variance to error variance ratio is four (i.e. $\sigma_s^2/\sigma_k^2=4$), the criteria behave sporadically. The AIC and HQIC choose UN most often and CS second, whereas the BIC and CAIC choose CS most often and AR(1) second. When $\sigma_s^2/\sigma_k^2=0.25$, ARH(1) is chosen most often for all four criteria. For true ARH(1) data, one should at least choose a model allowing for heterogeneous variances (i.e. CSH, ARH(1), TOEPH, ANTE(1), or UN). When the ARH(1) analyses fail to account for σ_s^2 and σ_s^2 is relatively large, the criteria do not consistently choose the same model (or even close to the same model). Problems develop when models which do not account for heterogeneous variances are chosen to analyze true ARH(1) data. These problems are discussed in Section 5c.

Table 4 gives the percent of data sets for the various correlation models chosen for the same four ARH(1) scenarios. However, now the ARH(1) analyses include the "random subj(trt)" and "parms/lowerb=1e-3,…,1e-3" statements. When $\sigma_s^2/\sigma_k^2=4$, all four criteria consistently choose ARH(1), the true model, most often. One exception is the CAIC, once again it tends to choose a simpler model; in this case when $\rho=0.75$ and $\sigma_s^2/\sigma_k^2=4$ it chooses the AR(1) most. Table

4 shows the importance of retaining σ_s^2 in the model for ARH(1) analyses when the errors are truly ARH(1), especially when σ_s^2/σ_k^2 is large.

Table 5 shows the percent of ANTE(1) data sets for which a given correlation model was chosen by the information criteria. The first table shows the ARH(1) analyses without the "random subj(trt)". The second table shows the ARH(1) analyses with the adequate parms/lowerb statements (5a.2). In both tables, the ARH(1) is chosen most often and ANTE(1) second most often, with the exception again of the CAIC, which chose the AR(1) second most often. We would not expect much change among the criteria when comparing the two ways to model the ARH(1) since no subject variance was generated for the ANTE(1) data.

5c.   Type I Error Rates

Table 6 gives the empirical rejection rates, hereafter denoted α_e, of the F-tests of H_o: no Time x Trt interaction at nominal $\alpha=0.05$, for four AR(1) scenarios. Each of the AR(1) scenarios were analyzed with eight different covariance structures in conjunction with the three denominator degree of freedom calculations listed in Section 3. Note that the ARH(1) analyses were modeled two different ways. The first ARH(1) analysis was modeled omitting σ_s^2 (i.e. the "random subj(trt)" statement in PROC MIXED). The second was modeled including "random subj(trt)" and the "parms/lowerb=1e-3,...,1e-3" statements. Since time only allowed us to run 500 simulated data sets for each scenario, empirical rejection rates within the 0.03 to 0.07 range form a 95% confidence interval about 0.05, and are hence " acceptable" at the nominal 5% level. Unless otherwise noted in the tables, 100% convergence for a given scenario and analysis was obtained.

When the correlation among errors is large (i.e. $\rho=0.75$), the CS analyses produce excessive empirical Type I errors: $\alpha_e=0.10$ versus the nominal $\alpha=0.05$. This is expected since the large autocorrelation is not modeled in the CS. As the covariance model of analysis becomes more complex, the Kenward-Roger adjustment becomes more important for controlling Type I errors. For example, when the fourth AR(1) scenario (i.e. $\rho=0.75$, $\sigma_s^2=4$) is analyzed as UN, the SAS default ddfm and Satterthwaite ddfm both give empirical rejection rates of $\alpha_e=0.22$ at the nominal 5% level. The KR ddfm gives an empirical rejection rate of $\alpha_e=0.06$.

Table 7 gives time x trt empirical rejection rates at the nominal 5% level for four ARH(1) scenarios and the ANTE(1) scenario. This table underlines the importance of the KR adjustment. However, the KR adjustment is not a "cure all." For example, when the ARH(1) scenarios are modeled with CS, which does not model correlation or heterogeneous variances, the empirical Type I errors for the KR are still excessive ($\alpha_e=0.07$ to $\alpha_e=0.10$), and are in fact act similar to those observed for the default and Satterthwaite ddfm.

Empirical rejection rates for time comparisons on treatment 1 for two AR(1) scenarios ($\rho=0.25$, $\sigma_s^2=0.25$ and $\rho=0.75$, $\sigma_s^2=0.25$) are given in Table 8. The tests are, Ho: no Time 1 vs. Time k (k=2,4,6) effect on trt. 1 at $\alpha=0.05$. Again, the AR(1) scenarios were analyzed the same covariance structure-ddfm combinations noted previously. The Type I errors throughout Table 8 are fairly stable, with a few cases where the KR adjustment controls Type I error better than the other two calculations. The exception is the AR(1) scenario with $\rho=0.75$ and $\sigma_s^2/\sigma_k^2=0.25$ when analyzed as CS and CSH: the Type I errors are excessive regardless of the ddfm method used.

Again, CS and CSH do not model the high autocorrelation of the data in this scenario..

Table 9 shows empirical rejection rates for treatment 1 vs. treatment 2 contrasts for a given time for the same AR(1) scenarios as in Table 8. The tests are, Ho: no trt 1 vs. trt2 effect at time k (k=1,3,4,6) at the nominal α=0.05. This table illustrates the fact that treatment simple effects within each time period are not affected by the KR adjustment.

Time contrasts on treatment 1 empirical rejection rates for two ARH(1) scenarios (ρ=0.25, σ_s^2/σ_k^2=4 and ρ=0.75, σ_s^2/σ_k^2=4) are given in Table 10. Type I errors are severely out of control when the ARH(1) data are modeled by covariance structures which do not account for heterogeneous variance, i.e. CS, AR(1) and TOEP. The KR adjustment provides little if any "damage control" when analyzing the ARH(1) data as CS, CSH, AR(1), TOEP, TOEPH and ANTE(1). The KR adjustment also provides no damage control when analyzing the ARH(1) data as ARH(1) without subj(trt) in the model. However, the KR adjustment significantly improves Type I errors for the ARH(1) analyses with the "random subj(trt)" and "parms/lowerb=1e-3,...,1e-3" statements. The observed Type I error rates for the UN analyses are quite good at the nominal 5% level fairly consistently throughout the entire simulation study.

Table 11 gives empirical rejection rates for treatment comparisons across time for two ARH(1) scenarios (ρ=0.25, σ_s^2/σ_k^2=0.25 and ρ=0.75, σ_s^2/σ_k^2=0.25). This table shows the importance of modeling ARH(1) data with a covariance structure that accounts for heterogeneous variances. All three ddfm calculations yield unreasonable Type I error rates for covariance model analyses which do not account for heterogeneous variances, i.e. CS, AR(1) and TOEP. Once again, the Type I errors for the UN analyses are fairly consistent at the nominal 5% level.

An example of the variance component estimates for an ARH(1) scenario analyzed as ARH(1) with and without σ_s^2 are given in Table 12. For the ARH(1) analysis without σ_s^2, it can be seen that the estimates of the variances of the observations at each of the time periods as well as the estimate of the correlation are severely inflated. For example, in generating the data, the variance for the first time period was set equal to σ_1^2=1.50. The ARH(1) analysis with the "random subj(trt)" and the "parms/lowerb=1e-3,...,1e-3" statements yields a mean estimate of σ_1^2=1.50, whereas the ARH(1) analysis dropping the "random subj(trt)" statement yields means estimates of σ_1^2=10.15. Similar results are obtained for the remaining estimates of the variance components. The critical importance of retaining σ_s^2 in the model can clearly be seen here.

Time comparisons on treatment 1 empirical rejection rates for the ANTE(1) scenario are given in Table 13, for the tests of Ho: No time 1 vs. time k (k=2,4,6) effect on trt 1 at nominal α=0.05. The Type I errors are out of control for all covariance models of analyses other than the ANTE(1) and UN, emphasizing the importance of modeling for the heterogeneous correlations and variances which were generated by the ANTE(1) scenario. Table 14 gives treatment 1 vs. treatment 2 comparisons across time. Results similar to those from Table 13 are obtained for these contrasts.

6.   Summary and Conclusions

Four main questions about PROC MIXED's analysis of repeated measures data were posed in section 2. We return to these questions with a summary of the answers this simulation study provides.

1.   How do the four model selection criteria computed by Version 8 of SAS PROC MIXED compare in terms of their ability to "choose the right" correlated error model?

The four criteria vary in their ability to "choose the right" model. In general, the BIC and even more strongly, the CAIC, tend to choose simpler models, whereas AIC and, especially, HQIC tend to choose more complex models. In terms of type I error control, assuming the Kenward-Roger adjustment is used, it is better to err in the direction of a somewhat more complex model. More complex models tend to have inflated type I error rates only if one fails to use the KR adjustment, whereas excessively simple models have inflated type I error rates that ddfm adjustments cannot correct. However, because complex models reduce power, erring too far in the direction of complex models is also bad. In this sense, the AIC appears to be the most desirable compromise in practice.

2.   How do Version 8 options to approximate degrees of freedom -- default, Satterthwaite, and Kenward-Roger -- compare in terms of their ability to control Type I error?

In all scenarios studied, the Kenward-Roger adjustment was either superior to, or at worst equal to, the Satterthwaite and default ddfm options. For the more complex covariance models, e.g. ARH(1), ANTE(1), and UN, type I error rate inflation was extremely severe unless the KR option was used. These results confirm those obtained by Kenward and Roger (1997) and extend them to a larger class of applications. Based on these results, we recommend using the Kenward-Roger option as standard operating procedure and recommend strongly *against* the use of either the default or Satterthwaite alternatives. Note that for balanced variance component-only models, the Satterthwaite procedure is a special case of Kenward-Roger and hence the two options produce identical results.

3.   The Kenward-Roger option also adjusts standard errors and test-statistics. How does this procedure perform?

Consistent with our conclusions for (2), the Kenward-Roger adjustment is *essential* for the more complex covariance structures. In general, we found that the impact is greatest on the overall test for Trt x Time interaction and for the simple effects of Time at given treatment levels. For simpler covariance structures, e.g. AR(1), the observed bias was trivial, but using the KR adjustment certainly did no harm.

4.   How robust is mixed model analysis to misspecification of the correlated error model?

There is a conventional wisdom or "oral tradition" among those who work with mixed models and repeated measures that most of the gain from modeling the covariance comes from "getting close." That is, if the data are AR(1), one pays dearly for failing to model autocorrelation (i.e. using CS) or for grossly over modeling it (i.e. UN) but similar models (e.g. TOEP) have trivial impact on the analysis. However, this "oral tradition" has never been systematically studied to our knowledge prior to this simulation study. In general, our results tend to support conventional wisdom. This study focused on type I error rates, so the effects of under modeling tend to be more obvious. We are currently extending this work by studying the impact of model misspecification on power, where the undesirability of overmodelling should be more obvious.

One aspect of model misspecification needs to be stressed. This concerns the question of modeling the between subject effects (s_{ij} in model 3.1) separately from the covariance among

repeated measures. For many practitioners, it is common practice *not* to model s_{ij} separately, but rather to absorb them into the correlated error structure. One practical reason for doing this is to avoid convergence problems such as those described in detail for the ARH(1) case. However, our results should make it clear that absorbing the variance by dropping the between subjects random effect is a form of model misspecification that can have a severe impact on accurate control of type I error. For this reason, except for obvious models such as CS, CSH, and UN, where the variance and covariance terms are mathematically confounded, we recommend that s_{ij} be modeled separately - i.e. separate random and repeated statements as shown in SAS code (5a.1) - if at all possible unless one can show that $\sigma_S^2 \approx 0$, independently of the components of Σ. Judicious use of lower and upper bound options in PROC MIXED makes this possible in all the cases we observed.

This misspecification issue has implications for repeated measure analysis of generalized linear models, e.g. the GEE option in PROC GENMOD. This option compels one to absorb the between subjects random effects into the covariance structure, much like dropping the "random subj(trt);" statement in the ARH(1) analysis. We are currently investigating whether doing this has a similar affect on GEE type I error control as we observed in this study.

Other topics for further investigation include the following. The first is a more careful examination of the type I error rate for each model selection criterion by tracing the observed rejection rates for the models actually selected for each simulated data set by each criterion. The second is documenting the observed standard errors with and without Kenward-Roger adjustments and comparing them to theory. Finally, we are following up our study of type I error control with a similar study addressing power.

In summary, this study has clearly demonstrated PROC MIXED's ability to analyze repeated measures data, subject to a few caveats. First, use the Kenward-Roger option - always. Second, when in doubt, err toward the more complex model -- but not too far! Third, retain the separate random between subjects term - do not absorb it into the repeated statement (covariance structure) unless theory dictates (as in CS) or there is compelling evidence that $\sigma_S^2 \approx 0$.

REFERENCES

Akaike, H. (1974), "A new look at the statistical model identification." *IEEE Transaction on Automatic Control*, **AC-19**: 716 -723.

Bozdogan, H. (1987), "Model selection and Akaike's information criterion (aic): the general theory and its analytical extensions." *Psychometrika*, **52**: 345-370.

Carlin, B.P. 1996. *Bayes and Empirical Bayes Methods and Applications*. New York: Chapman & Hall.

Diggle, P.J., K.-Y. Liang, and S.L. Zeger. 1994. *Analysis of Longitudinal Data*. New York: Oxford University Press.

Hannan, E.J. and Quinn, B.G. (1979), "The Determination of the order of an autoregression," *J. Royal Statist. Soc., Ser. B*, **41**: 190-195.

Kacker, R.N. and D.A. Harville. 1984 "Approximations for standard errors of estimators of fixed and random effects in mixed linear models." *J. Amer. Statist. Assoc.* **79**: 853-862.

Kenward, M.G. and J.H. Roger/ 1997. "Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics* **53**: 983-997.

Keselman, H.J., J. Algina, R.K. Kowalchuk, R.D. Wolfinger. 1998. "A comparison of two approaches for selecting covariance structures in the analysis of repeated measurements." *Commun. Statist. - Simula.* **27**(3): 591-604.

Liang,K.-Y. and S.L. Zeger. 1986. "Longitudinal data analysis using generalized linear models." *Biometrika* **73**: 13-22.

Littell, R.C., G.A. Milliken, W.W. Stroup, and R.D. Wolfinger. 1996. *Sas System for Mixed Models*. Cary, NC: SAS Institute.

SAS Institute, Inc. 1997. *SAS/STAT Software: Changes and Enhancements through Release 6.12*. Cary NC: SAS Institute.

SAS Institute, Inc. 1999. *SAS OnLine Doc, Version 8*. Cary NC: SAS Institute.

Schwarz, G. (1978), "Estimating the dimension of a model," *Annals of Statistics*, **6**: 461-464.

Wolfinger, R.D. 1999. Personal communication.

Applied Statistics in Agriculture

Figure 1. Repeated Measures Experiment: Generic Description

2 treatments: P=placebo; T=treated
6 subjects per treatment; 6 times of observation

subject	1	2	3	4	5	6	7	8	9	10	11	12
time ⇓ trt ⇒	P	P	P	P	P	P	T	T	T	T	T	T
1												
2												
3												
4												
5												
6												

Figure 2. Covariance Models and Parameter Values for Simulated Data Scenarios

Scenario	Covariance Model	ρ	Subject/Error Variance Ratio	Error Variance
1	AR(1)	0.25	$\sigma_s^2 / \sigma^2 = 4$	$\sigma^2 = 1$
2		0.75		
3		0.25	$\sigma_s^2 / \sigma^2 = 1$	
4		0.75		
5		0.25	$\sigma_s^2 / \sigma^2 = 0.25$	
6		0.75		
7	ARH(1)	0.25	$\sigma_s^2 = 6 = 4X$ maximum σ_i^2	σ_i^2, (i=1.2,3,4,5,6) = (1.50, 1.25, 1.00, 0.50, 0.25, 0.10)
8		0.75		
9		0.25	$\sigma_s^2 = 1$	
10		0.75		
11		0.25	$\sigma_s^2 = 0.025$ $= 0.25X$ minimum σ_i^2	
12		0.75		
13	ANTE(1)	ρ_i (i=1,2,3,4,5) = (0.8,0.6,0.6,0.4,0.4)	σ_s^2 not applicable	σ_i^2, (i=1.2,3,4,5,6) = (1.50, 1.25, 1.00, 0.50, 0.25, 0.10)

Applied Statistics in Agriculture

Figure 3. Example SAS PROC MIXED Failure to Converge Using ARH(1) Covariance Model

Default SAS Code:

```
proc mixed;
 class trt subj time;
 model y=trt|time/ddfm=kr;
 random subj(trt);
 repeated /type=arh(1) subject=subj(trt);
```

Results:

```
                        Iteration History

        Iteration     Evaluations     -2 Res Log Like        Criterion

               0               1        306.04081621


        WARNING: Stopped because of infinite likelihood.


                   Covariance Parameter Values
                       At Last Iteration

              Cov Parm        Subject        Estimate

              SUBJ(TRT)                         6.2934
              Var(1)          SUBJ(TRT)         0.8036
              Var(2)          SUBJ(TRT)         0.8024
              Var(3)          SUBJ(TRT)         1.4419
              Var(4)          SUBJ(TRT)        1.11E-12
              Var(5)          SUBJ(TRT)        1.11E-12
              Var(6)          SUBJ(TRT)         0.06172
              ARH(1)          SUBJ(TRT)         0.4079
```

Figure 4. SAS - PROC MIXED ARH(1) Example with Inadequate Lower Bound

SAS Code:

```
proc mixed;
 class trt subj time;
 model y=trt|time/ddfm=kr;
 parms 0 1 1 1 1 1 1 0 / lowerb=0,1e-4,1e-4,1e-4,1e-4,1e-4,1e-4;
 random subj(trt);
 repeated /type=arh(1) subject=subj(trt);
```

Results:

```
                    Convergence criteria met.

                 Covariance Parameter Estimates

         Cov Parm        Subject         Estimate

         SUBJ(TRT)                         5.7302
         Var(1)          SUBJ(TRT)         1.4295
         Var(2)          SUBJ(TRT)         1.0437
         Var(3)          SUBJ(TRT)         0.6029
         Var(4)          SUBJ(TRT)        0.000100
         Var(5)          SUBJ(TRT)         0.08502
         Var(6)          SUBJ(TRT)         0.1460
         ARH(1)          SUBJ(TRT)         0.6852


               Type 3 Tests of Fixed Effects

                    Num     Den
         Effect      DF      DF      F Value    Pr > F

         TRT          1     10.8      Infty    <.0001
         TIME         1      0        0.00      .
         TRT*TIME     5      17       0.51     0.7616


                         Estimates

                            Standard
     Label            Estimate      Error       DF    t Value    Pr > |t|

     time1 vs time2 @ trt1    0.03104         0       17     Infty     <.0001
     time1 vs time3 @ trt1    0.03509    181.02        0      0.00     <.0001
     time1 vs time4 @ trt1   -0.1228     181.02        0     -0.00     <.0001
     time1 vs time5 @ trt1   -0.4234     0.004046      0   -104.66     1.0000
     time1 vs time6 @ trt1   -0.3892     0.01341       0    -29.02     1.0000
     trt1 vs trt2 @ time 1    1.6224          0       17     Infty     <.0001
     trt1 vs trt2 @ time 2    1.5138          0       17     Infty     <.0001
     trt1 vs trt2 @ time 3    1.8407          0       17     Infty     <.0001
     trt1 vs trt2 @ time 4    2.0658          0       17     Infty     <.0001
     trt1 vs trt2 @ time 5    2.0931          0       17     Infty     <.0001
     trt1 vs trt2 @ time 6    1.8802          0       17     Infty     <.0001
```

Applied Statistics in Agriculture

Figure 5. SAS - PROC MIXED ARH(1) Example with Adequate Lower Bound

SAS Code:

```
proc mixed;
 class trt subj time;
 model y=trt|time/ddfm=kr;
 parms 0 1 1 1 1 1 0 / lowerb=0,1e-3,1e-3,1e-3,1e-3,1e-3,1e-3;
 random subj(trt);
 repeated /type=arh(1) subject=subj(trt);
```

Results:

```
                    Convergence criteria met.

                 Covariance Parameter Estimates

        Cov Parm         Subject        Estimate

        SUBJ(TRT)                         5.6966
        Var(1)           SUBJ(TRT)        1.4781
        Var(2)           SUBJ(TRT)        1.0817
        Var(3)           SUBJ(TRT)        0.6315
        Var(4)           SUBJ(TRT)        0.001000
        Var(5)           SUBJ(TRT)        0.09292
        Var(6)           SUBJ(TRT)        0.1537
        ARH(1)           SUBJ(TRT)        0.6998


                 Type 3 Tests of Fixed Effects

                        Num      Den
        Effect           DF       DF     F Value    Pr > F

        TRT               1     10.9        1.70    0.2188
        TIME              5     16.9        3.05    0.0383
        TRT*TIME          5     16.9        0.48    0.7848


                           Estimates

                              Standard
    Label               Estimate      Error       DF    t Value    Pr > |t|

    time1 vs time2 @ trt1   0.03104    0.3639     16.7      0.09      0.9330
    time1 vs time3 @ trt1   0.03509    0.4476     16.3      0.08      0.9385
    time1 vs time4 @ trt1  -0.1228     0.4924     10.3     -0.25      0.8080
    time1 vs time5 @ trt1  -0.4234     0.4873     11.7     -0.87      0.4024
    time1 vs time6 @ trt1  -0.3892     0.5023     12.9     -0.77      0.4525
    trt1 vs trt2 @ time 1   1.6224     1.5465     14.9      1.05      0.3109
    trt1 vs trt2 @ time 2   1.5138     1.5031     13.7      1.01      0.3314
    trt1 vs trt2 @ time 3   1.8407     1.4524     12.2      1.27      0.2287
    trt1 vs trt2 @ time 4   2.0658     1.3781     10        1.50      0.1647
    trt1 vs trt2 @ time 5   2.0931     1.3892     10.3      1.51      0.1619
    trt1 vs trt2 @ time 6   1.8802     1.3965     10.5      1.35      0.2064
```

Table 1. Percent Convergence Rates for ARH(1) Model Analyses

Generated Covariance Structure	ρ	σ_s^2/σ^2	With subj(trt) Default SV*	OLS**	With subj(trt) parms/lowerb
AR(1)	0.25	4	18	9	100
AR(1)	0.75	4	12	53	100
AR(1)	0.25	1	62	66	100
AR(1)	0.75	1	34	91	100
AR(1)	0.25	0.25	94	91	100
AR(1)	0.75	0.25	56	87	100
ARH(1)	0.25	4X max σ_i^2	8		100
ARH(1)	0.75	4X max σ_i^2	6		100
ARH(1)	0.25	1	36		100
ARH(1)	0.75	1	18		100
ARH(1)	0.25	0.25 min σ_i^2	71		100
ARH(1)	0.75	0.25 min σ_i^2	12		100
ANTE(1)	***	na	26		100

*   SAS Default Starting Values (MIVQUE(0))
**   Ordinary Least Squares estimates for starting values
*** ρ =(0.80, 0.60, 0.60, 0.40, 0.40)

Table 2. Percent of Covariance Models Chosen by Four Model Fitting Criteria, AR(1) Data

				Criterion			
Covariance Structure	ρ	σ_S^2 / σ^2	Model Analysis	AIC	HQIC	BIC	CAIC
AR(1)	0.25	4	CS	39.6	34.4	49.6	67.0
			CSH	3.4	4.2	2.2	0.6
			AR(1)	28.6	24.6	33.8	29.2
			ARH(1)	11.4	14.4	5.8	2.0
			TOEP	7.8	9.2	5.2	1.0
			TOEPH	2.4	2.6	1.0	0.2
			ANTE(1)	0.6	0.8	0.2	0.0
			UN	6.2	9.8	2.2	0.0
AR(1)	0.75	4	CS	1.4	1.0	1.6	4.8
			CSH	0.6	0.6	0.8	0.4
			AR(1)	63.8	55.8	79.2	90.0
			ARH(1)	13.8	16.0	7.0	2.0
			TOEP	9.6	11.2	7.2	2.6
			TOEPH	3.2	3.8	2.0	0.0
			ANTE(1)	2.0	2.6	0.8	0.0
			UN	5.6	9.0	1.4	0.2
AR(1)	0.25	0.25	CS	40.4	33.6	48.8	62.4
			CSH	2.4	2.8	1.2	0.2
			AR(1)	31.4	28.8	37.2	34.6
			ARH(1)	6.6	8.6	4.4	1.6
			TOEP	8.6	10.6	5.2	1.0
			TOEPH	2.6	3.8	1.0	0.0
			ANTE(1)	1.8	2.8	0.6	0.2
			UN	6.2	9.0	1.6	0.0
AR(1)	0.75	0.25	CS	1.0	0.8	1.8	4.0
			CSH	0.2	0.2	0.0	0.0
			AR(1)	66.0	57.2	82.2	91.8
			ARH(1)	11.6	13.2	6.2	1.6
			TOEP	10.8	12.4	6.8	2.4
			TOEPH	2.2	3.6	1.0	0.2
			ANTE(1)	2.8	4.2	1.0	0.0
			UN	5.4	8.4	1.0	0.0

Table 3. Percent of Covariance Models Chosen by Four Model Fitting Criteria, ARH(1) Data

Covariance Structure	ρ	σ_S^2 / σ^2	Model Analysis	Criterion			
				AIC	HQIC	BIC	CAIC
ARH(1)	0.25	4X max σ_i^2	CS	24.0	19.2	41.4	65.0
			CSH	6.2	5.0	5.6	2.4
			AR(1)	17.2	13.4	23.8	27.6
			ARH(1) *	2.6	2.4	3.8	2.4
			TOEP	2.2	2.0	1.8	0.2
			TOEPH	1.4	1.6	1.0	0.0
			ANTE(1)	13.0	11.0	10.0	1.6
			UN	33.4	45.4	12.6	0.8
ARH(1)	0.75	4X max σ_i^2	CS	0.0	0.0	1.0	2.4
			CSH	0.4	0.2	0.4	0.2
			AR(1)	30.8	23.6	50.6	81.2
			ARH(1) *	9.0	6.6	8.2	5.2
			TOEP	4.6	4.4	3.0	1.6
			TOEPH	2.8	3.6	2.4	0.2
			ANTE(1)	35.0	37.2	27.8	8.4
			UN	17.4	24.4	6.6	0.8
ARH(1)	0.25	0.25X min σ_i^2	CS	0.4	0.0	2.0	12.6
			CSH	24.0	21.0	26.2	24.8
			AR(1)	0.4	0.2	1.2	8.2
			ARH(1) *	42.4	38.2	53.0	51.6
			TOEP	0.0	0.0	0.0	0.0
			TOEPH	11.0	12.8	6.0	1.0
			ANTE(1)	11.2	11.6	7.6	1.4
			UN	10.6	16.2	4.0	0.4
ARH(1)	0.75	0.25X min σ_i^2	CS	0.0	0.0	0.0	0.2
			CSH	0.0	0.0	0.0	3.2
			AR(1)	0.2	0.0	1.0	8.6
			ARH(1) *	73.8	64.8	84.2	85.2
			TOEP	0.0	0.0	0.0	0.0
			TOEPH	0.0	0.0	0.0	0.0
			ANTE(1)	12.0	14.2	6.8	2.4
			UN	10.8	18.2	4.2	0.4

* ARH(1) without random subj(trt) effect

Applied Statistics in Agriculture

Table 4. Percent of Covariance Models Chosen by Four Model Fitting Criteria, ARH(1) Data

Covariance Structure	ρ	σ_S^2 / σ_k^2	Model Analysis	Criterion AIC	HQIC	BIC	CAIC
ARH(1)	0.25	4X max σ_k^2	CS	7.4	4.6	16.6	39.4
			CSH	2.0	1.8	2.6	1.2
			AR(1)	5.2	2.8	7.2	15.2
			ARH(1) †	67.6	65.6	65.4	43.0
			TOEP	0.2	0.4	0.8	0.2
			TOEPH	0.6	0.8	0.8	0.0
			ANTE(1)	2.2	2.0	1.6	0.6
			UN	14.8	22.0	5.0	0.4
ARH(1)	0.75	4X max σ_k^2	CS	0.0	0.0	0.0	1.6
			CSH	0.2	0.2	0.2	0.2
			AR(1)	14.4	10.0	29.4	57.4
			ARH(1) †	57.6	55.6	54.0	35.8
			TOEP	2.4	1.8	1.0	1.4
			TOEPH	2.4	3.0	2.0	0.4
			ANTE(1)	13.2	14.6	8.4	2.6
			UN	9.8	14.8	5.0	0.6
ARH(1)	0.25	0.25X min σ_k^2	CS	0.4	0.2	2.2	12.8
			CSH	26.4	22.8	31.0	33.0
			AR(1)	1.0	0.8	1.6	9.6
			ARH(1) †	37.6	35.2	45.4	41.0
			TOEP	0.0	0.0	0.0	0.0
			TOEPH	12.6	13.8	7.8	1.2
			ANTE(1)	11.4	11.4	8.4	2.0
			UN	10.6	15.8	3.6	0.4
ARH(1)	0.75	0.25X min σ_k^2	CS	0.0	0.0	0.0	0.0
			CSH	1.6	1.2	3.4	3.4
			AR(1)	0.4	0.2	1.4	12.0
			ARH(1) †	64.2	57.0	77.8	79.2
			TOEP	0.0	0.0	0.0	0.0
			TOEPH	11.5	11.6	7.4	3.4
			ANTE(1)	12.0	13.8	7.4	2.0
			UN	10.4	16.2	2.6	0.0

† ARH(1) with random subj(trt) effect, lowerb option

Table 5. Percent of Covariance Models Chosen by Four Model Fitting Criteria

Covariance Structure of Data	ρ	σ_s^2/σ_k^2	Model Analysis	Criterion			
				AIC	HQIC	BIC	CAIC
ANTE(1)	*	na	CS	0.2	0.2	0.6	2.2
			CSH	3.4	3.6	4.2	4.2
			AR(1)	2	0.6	8.8	27.8
			*ARH(1)	42.6	36.8	55	56.4
			TOEP	0.6	0.6	0.6	0.6
			TOEPH	0	0	0	0
			ANTE(1)	30.2	31.4	23.8	8.4
			UN	21	26.8	7	0.4
ANTE(1)	*	na	CS	0.2	0.2	0.4	2.0
			CSH	2.8	2.6	4.0	5.0
			AR(1)	2.2	1.0	8.8	28.6
			*ARH(1)	42.0	37.0	55.4	54.6
			TOEP	0.6	0.6	0.6	0.6
			TOEPH	0.0	0.0	0.0	0.0
			ANTE(1)	31.4	32.2	24.4	8.8
			UN	20.8	26.4	6.4	0.4

* ρ =(0.80, 0.60, 0.60, 0.40, 0.40)

Applied Statistics in Agriculture

Table 6. Empirical Rejection Rates for Test of Ho: No Time x Trt Effect at nominal $\alpha = 0.05$

Method of Analysis		AR(1) Data Generated			
	ρ	0.25	0.75	0.25	0.75
Covariance Structure	DDFM option $\quad \sigma_s^2 / \sigma^2$	0.25	0.25	4	4
CS	all	0.06	0.10	0.06	0.10
CSH	Def*	0.11	0.10	0.07	0.09
	Satt**	0.07	0.07	0.06	0.07
	KR***	0.05	0.06	0.04	0.05
AR(1)	Def	0.07	0.07	0.07	0.07
	Satt	0.05	0.06	0.05	0.05
	KR	0.05	0.03	0.04	0.02
ARH(1) Without subj(trt)	Def	0.09	0.09	0.07	0.07
	Satt	0.07	0.07	0.06	0.06
	KR	0.04	0.03	0.05	0.03
ARH(1) With subj(trt)	Def	0.15	0.14	0.18	0.16
parms/lowerb	Satt	0.09	0.09	0.11	0.11
	KR	0.06	0.04	0.10	0.07
TOEP	Def	0.12	0.10	0.12	0.11
	Satt	0.08	0.08	0.08	0.08
	KR	0.05	0.06	0.05	0.05
TOEPH	Def	0.16 a	0.15 b	0.14	0.13
	Satt	0.10 a	0.10 b	0.09	0.09
	KR	0.05 a	0.05 b	0.04	0.04
ANTE(1)	Def	0.14	0.14	0.14	0.12
	Satt	0.08	0.09	0.09	0.10
	KR	0.04	0.03	0.04	0.03
UN	Def	0.20	0.22	0.20	0.22
	Satt	0.20	0.22	0.20	0.22
	KR	0.07	0.06	0.07	0.06

Notes: *   model y=trt/time
      **   model y=trt/time/ddfm=satterthwaite
    ***   model y=trt/time/ddfm=KR

    a   99.6% convergence
    b   98% convergence

Table 7. Empirical Rejection Rates for Test of Ho: No Time x Trt Effect at nominal $\alpha = 0.05$

Method of Analysis		Covariance Model and Parameters for Data Generated					ANTE(1)
		ARH (1)					
Covariance Structure	DDFM	ρ	0.25	0.75	0.25	0.75	**
		σ_s^2 / σ^2	0.25	0.25	4	4	na
CS	Def		0.07	0.10	0.07	0.10	0.09
	Satt		0.08	0.10	0.07	0.10	0.09
	KR		0.08	0.10	0.07	0.10	0.09
CSH	Def		0.12	0.10	0.07	0.09	0.08
	Satt		0.07	0.07	0.06	0.08	0.07
	KR		0.05	0.06	0.04	0.06	0.06
AR(1)	Def		0.09	0.08	0.09 c	0.08 c	0.07
	Satt		0.08	0.07	0.08 c	0.07 c	0.06
	KR		0.07	0.06	0.06 c	0.03 c	0.05
ARH(1) Without subj(trt)	Def		0.11	0.10	0.08	0.08	0.09
	Satt		0.08	0.08	0.07	0.06	0.07
	KR		0.05	0.06	0.04	0.04	0.06
ARH(1) With subj(trt parms/lowerb	Def		0.14	0.13	0.16	0.14	0.11
	Satt		0.09	0.10	0.10	0.10	0.08
	KR		0.06	0.07	0.09	0.06	0.06
TOEP	Def		0.10	0.10	0.11	0.11	0.12
	Satt		0.08	0.07	0.08	0.07	0.09
	KR		0.07	0.06	0.06	0.06	0.06
TOEPH	Def		0.16 a	b	0.13	0.13	d
	Satt		0.10 a	b	0.08	0.07	d
	KR		0.04 a	b	0.04	0.04	d
ANTE(1)	Def		0.16	0.16	0.14	0.13	0.16
	Satt		0.10	0.11	0.10	0.10	0.11
	KR		0.05	0.06	0.05	0.05	0.06
UN	Def		0.21	0.22	0.21	0.22	0.24
	Satt		0.21	0.22	0.21	0.22	0.24
	KR		0.06	0.06	0.06	0.06	0.07

a 90% convergence
b 65% convergence
c 99% convergence
d 70% convergence

Applied Statistics in Agriculture

Table 8. Empirical Rejection Rates For Test of Ho: No Time 1 x Time k (k=2,4,6) Effect Given Trt 1
AR(1) Data, $\sigma_S^2/\sigma^2 = 0.25$, at nominal $\alpha = 0.05$

Method of Analysis		$\rho=0.25$			$\rho=0.75$		
Covariance Structure	DDFM	Time 1 v 2	Time 1 v 4	Time 1 v 6	Time 1 v 2	Time 1 v 4	Time 1 v 6
CS	All	0.03	0.06	0.06	0.02	0.10	0.16
CSH	Def	0.04	0.07	0.08	0.02	0.09	0.15
	Satt	0.03	0.06	0.07	0.02	0.08	0.14
	KR	0.03	0.06	0.07	0.01	0.08	0.13
AR(1)	Def	0.06	0.06	0.05	0.06	0.07	0.09
	Satt	0.05	0.06	0.05	0.06	0.06	0.07
	KR	0.05	0.06	0.04	0.04	0.05	0.05
ARH(1)	Def	0.06	0.05	0.04	0.06	0.05	0.05
Without subj(trt)	Satt	0.04	0.05	0.04	0.05	0.05	0.05
	KR	0.05	0.04	0.04	0.05	0.05	0.04
ARH(1)	Def	0.07	0.07	0.07	0.06	0.07	0.09
With subj(trt)	Satt	0.04	0.06	0.06	0.05	0.05	0.07
parms/lowerb	KR	0.04	0.06	0.06	0.07	0.05	0.06
TOEP	Def	0.05	0.06	0.08	0.06	0.06	0.08
	Satt	0.05	0.05	0.05	0.06	0.06	0.05
	KR	0.05	0.05	0.05	0.06	0.06	0.05
TOEPH	Def	0.05	0.07	0.07	0.05	0.06	0.09
	Satt	0.04	0.06	0.05	0.04	0.05	0.06
	KR	0.04	0.06	0.04	0.04	0.05	0.06
ANTE(1)	Def	0.07	0.05	0.04	0.09	0.05	0.05
	Satt	0.05	0.04	0.04	0.05	0.04	0.04
	KR	0.05	0.04	0.04	0.04	0.04	0.04
UN	Def	0.05	0.05	0.05	0.05	0.04	0.06
	Satt	0.05	0.05	0.05	0.05	0.04	0.06
	KR	0.05	0.05	0.05	0.05	0.04	0.06

Table 9. Empirical Rejection Rates for Test of Ho: No Trt 1 vs Trt 2 Effect at Time k
AR(1) Data, $\sigma_S^2/\sigma^2 = 0.25$, at nominal $\alpha = 0.05$

Method of Analysis		$\rho=0.25$				$\rho=0.75$			
Covariance Structure	DDFM	Time 1	Time 3	Time 4	Time 6	Time 1	Time 3	Time 4	Time 6
CS	Def	0.05	0.04	0.06	0.05	0.05	0.06	0.06	0.06
	Satt	0.05	0.04	0.06	0.05	0.05	0.05	0.06	0.06
	KR	0.05	0.04	0.06	0.05	0.05	0.05	0.06	0.06
CSH	Def	0.07	0.07	0.06	0.09	0.07	0.08	0.09	0.08
	Satt	0.05	0.05	0.05	0.05	0.04	0.06	0.06	0.06
	KR	0.05	0.05	0.05	0.05	0.04	0.06	0.06	0.06
AR(1)	Def	0.05	0.05	0.06	0.05	0.05	0.06	0.07	0.07
	Satt	0.05	0.04	0.06	0.05	0.05	0.05	0.06	0.06
	KR	0.05	0.04	0.06	0.05	0.05	0.05	0.06	0.06
ARH(1) Without subj(trt)	Def	0.07	0.07	0.06	0.08	0.07	0.07	0.08	0.08
	Satt	0.05	0.04	0.05	0.05	0.05	0.05	0.06	0.07
	KR	0.05	0.04	0.05	0.05	0.05	0.05	0.06	0.07
ARH(1) With subj(trt parms/lowerb	Def	0.07	0.06	0.06	0.08	0.07	0.06	0.07	0.08
	Satt	0.05	0.04	0.05	0.05	0.05	0.04	0.05	0.07
	KR	0.05	0.04	0.05	0.05	0.05	0.04	0.05	0.07
TOEP	Def	0.05	0.04	0.06	0.05	0.05	0.06	0.06	0.06
	Satt	0.05	0.04	0.06	0.05	0.05	0.05	0.05	0.06
	KR	0.05	0.04	0.06	0.05	0.05	0.05	0.05	0.06
TOEPH	Def	0.07	0.07	0.06	0.08	0.07	0.08	0.07	0.08
	Satt	0.05	0.04	0.05	0.04	0.04	0.05	0.05	0.06
	KR	0.05	0.04	0.05	0.04	0.04	0.05	0.05	0.06
ANTE(1)	Def	0.07	0.07	0.06	0.09	0.07	0.08	0.08	0.09
	Satt	0.05	0.06	0.05	0.05	0.05	0.05	0.06	0.07
	KR	0.05	0.06	0.05	0.05	0.05	0.05	0.06	0.07
UN	Def	0.05	0.06	0.05	0.05	0.05	0.05	0.06	0.07
	Satt	0.05	0.06	0.05	0.05	0.05	0.05	0.06	0.07
	KR	0.05	0.06	0.05	0.05	0.05	0.05	0.06	0.07

Table 10. Empirical Rejection Rates For Test of Ho: No Time 1 x Time k (k=2,4,6) Effect Given Trt 1
ARH(1) Data, $\sigma_S^2/\sigma^2 = 4X$ max σ_k^2, at nominal $\alpha = 0.05$

Method of Analysis		$\rho=0.25$			$\rho=0.75$		
Covariance Structure	DDFM	Time 1 v 2	Time 1 v 4	Time 1 v 6	Time 1 v 2	Time 1 v 4	Time 1 v 6
CS	All	0.11	0.11	0.08	0.04	0.13	0.15
CSH	Def	0.10	0.10	0.08	0.03	0.12	0.13
	Satt	0.09	0.09	0.07	0.03	0.11	0.12
	KR	0.08	0.08	0.06	0.03	0.10	0.12
AR(1)	Def	0.14	0.10	0.07	0.14	0.10	0.09
	Satt	0.14	0.10	0.07	0.14	0.09	0.07
	KR	0.12	0.09	0.07	0.07	0.04	0.03
ARH(1)	Def	0.09	0.01	0.00	0.12	0.05	0.03
Without subj(trt)	Satt	0.09	0.01	0.00	0.11	0.05	0.03
	KR	0.08	0.01	0.00	0.09	0.04	0.02
ARH(1)	Def	0.06	0.07	0.08	0.08	0.07	0.08
With subj(trt	Satt	0.05	0.06	0.07	0.06	0.06	0.06
parms/lowerb	KR	0.04	0.06	0.06	0.05	0.04	0.05
TOEP	Def	0.13	0.10	0.09	0.13	0.09	0.08
	Satt	0.13	0.10	0.07	0.13	0.08	0.06
	KR	0.13	0.10	0.07	0.13	0.08	0.06
TOEPH	Def	0.11	0.08	0.07	0.11	0.07	0.06
	Satt	0.10	0.06	0.05	0.10	0.05	0.04
	KR	0.09	0.05	0.04	0.09	0.04	0.04
ANTE(1)	Def	0.08	0.01	0.01	0.09	0.04	0.04
	Satt	0.05	0.01	0.01	0.05	0.03	0.04
	KR	0.04	0.01	0.00	0.05	0.02	0.03
UN	Def	0.05	0.05	0.06	0.05	0.04	0.06
	Satt	0.05	0.05	0.06	0.05	0.04	0.06
	KR	0.05	0.05	0.06	0.05	0.04	0.06

Table 11. Empirical Rejection Rates for Test of Ho: No Trt 1 vs Trt 2 Effect at Time k
ARH(1) Data, $\sigma_s^2/\sigma^2 = 0.25X$ min σ_k^2, at nominal $\alpha = 0.05$

Method of Analysis		$\rho=0.25$				$\rho=0.75$			
Covariance Structure	DDFM	Time 1	Time 3	Time 4	Time 6	Time 1	Time 3	Time 4	Time 6
CS	Def	0.16	0.09	0.02	0.00	0.16	0.11	0.03	0.00
	Satt	0.16	0.09	0.02	0.00	0.15	0.11	0.02	0.00
	KR	0.16	0.09	0.02	0.00	0.15	0.11	0.02	0.00
CSH	Def	0.07	0.09	0.06	0.09	0.06	0.09	0.09	0.08
	Satt	0.04	0.06	0.04	0.05	0.04	0.07	0.06	0.06
	KR	0.04	0.06	0.04	0.05	0.04	0.07	0.06	0.06
AR(1)	Def	0.15	0.09	0.02	0.00	0.16	0.10	0.03	0.00
	Satt	0.15	0.09	0.02	0.00	0.15	0.09	0.02	0.00
	KR	0.15	0.09	0.02	0.00	0.15	0.09	0.02	0.00
ARH(1) Without subj(trt)	Def	0.07	0.09	0.06	0.09	0.06	0.08	0.08	0.08
	Satt	0.05	0.06	0.05	0.05	0.05	0.05	0.06	0 07
	KR	0.05	0.06	0.05	0.05	0.05	0.05	0.06	0.07
ARH(1) With subj(trt parms/lowerb	Def	0.07	0.09	0.05	0.08	0.07	0.08	0.08	0.08
	Satt	0.05	0.06	0.04	0.05	0.05	0.05	0.06	0.07
	KR	0.05	0.06	0.04	0.05	0.05	0.05	0.06	0.07
TOEP	Def	0.15	0.08	0.02	0.00	0.15	0.10	0.03	0.00
	Satt	0.15	0.08	0.02	0.00	0.15	0.09	0.02	0.00
	KR	0.15	0.08	0.02	0.00	0.15	0.09	0.02	0.00
TOEPH	Def	0.07	0.10	0.06	0.08	a	a	a	a
	Satt	0.04	0.06	0.04	0.04				
	KR	0.04	0.06	0.04	0.04				
ANTE(1)	Def	0.07	0.08	0.06	0.09	0.07	0.09	0.08	0.09
	Satt	0.05	0.07	0.04	0.05	0.05	0.06	0.06	0.07
	KR	0.05	0.07	0.04	0.05	0.05	0.06	0.06	0.07
UN	Def	0.05	0.07	0.04	0.05	0.05	0.06	0.06	0.07
	Satt	0.05	0.07	0.04	0.05	0.05	0.06	0.06	0.07
	KR	0.05	0.07	0.04	0.05	0.05	0.06	0.06	0.07

a 65% convergence

Table 12. An Example of Covariance Parameter Estimates for ARH(1) Analysis

Covariance Parameter	True Covariance Parameters for ARH(1) Generated Data	ARH(1) with subj(trt) & parms/lowerb	ARH(1) w/o subj(trt)
σ_S^2	6.00	6.08	na
σ_1^2	1.50	1.50	10.15
σ_2^2	1.25	1.24	9.57
σ_3^2	1.00	1.04	8.04
σ_4^2	0.50	0.53	6.38
σ_5^2	0.25	0.23	5.31
σ_6^2	0.10	0.12	5.01
ρ	0.25	0.24	0.92

Table 13. Empirical Rejection Rates For Test of Ho: No Time 1 x Time k (k=2,4,6) Effect Given Trt 1
ANTE(1) Data, at nominal α = 0.05

Method of Analysis

Covariance Structure	DDFM	Time 1 v 2	Time 1 v 4	Time 1 v 6
CS	all	0.01	0.12	0.11
CSH	Def	0.00	0.08	0.09
	Satt	0.00	0.06	0.07
	KR	0.00	0.06	0.07
AR(1)	Def	0.05	0.10	0.07
	Satt	0.05	0.10	0.07
	KR	0.03	0.09	0.06
ARH(1)	Def	0.02	0.07	0.09
Without subj(trt)	Satt	0.02	0.06	0.07
	KR	0.02	0.06	0.07
ARH(1)	Def	0.02	0.08	0.09
With subj(trt)	Satt	0.02	0.06	0.07
parms/lowerb	KR	0.02	0.06	0.07
TOEP	Def	0.04	0.08	0.06
	Satt	0.04	0.08	0.05
	KR	0.04	0.08	0.05
TOEPH	Def	a	a	a
ANTE(1)	Def	0.08	0.06	0.06
	Satt	0.06	0.05	0.05
	KR	0.05	0.05	0.05
UN	Def	0.06	0.05	0.05
	Satt	0.06	0.05	0.05
	KR	0.06	0.05	0.05

a 70% convergence

Applied Statistics in Agriculture

Table 14. Empirical Rejection Rates for Test of Ho: No Trt 1 vs Trt 2 Effect at Time k, (k=1,3,4,6)
ANTE(1) Data, at nominal α = 0.05

Method of Analysis					
Covariance Structure	DDFM	Time 1	Time 3	Time 4	Time 6
CS	all	0.19	0.09	0.02	0.00
CSH	Def	0.08	0.08	0.06	0.05
	Satt	0.06	0.06	0.04	0.04
	KR	0.06	0.06	0.04	0.04
AR(1)	Def	0.19	0.09	0.02	0.00
	Satt	0.18	0.09	0.01	0.00
	KR	0.18	0.09	0.01	0.00
ARH(1)	Def	0.10	0.08	0.04	0.05
Without subj(trt)	Satt	0.09	0.06	0.03	0.03
	KR	0.09	0.06	0.03	0.03
ARH(1)	Def	0.10	0.08	0.04	0.05
With subj(trt)	Satt	0.09	0.06	0.03	0.03
parms/lowerb	KR	0.09	0.06	0.03	0.03
TOEP	Def	0.18	0.09	0.02	0.00
	Satt	0.18	0.08	0.01	0.00
	KR	0.18	0.08	0.01	0.00
TOEPH	all	a	a	a	a
ANTE(1)	Def	0.08	0.07	0.06	0.07
	Satt	0.05	0.05	0.04	0.05
	KR	0.05	0.05	0.04	0.05
UN	Def	0.05	0.05	0.04	0.05
	Satt	0.05	0.05	0.04	0.05
	KR	0.05	0.05	0.04	0.05

a 70% convergence