Practice with Fixed and Random Effects of Time in Modeling Within-Person Change

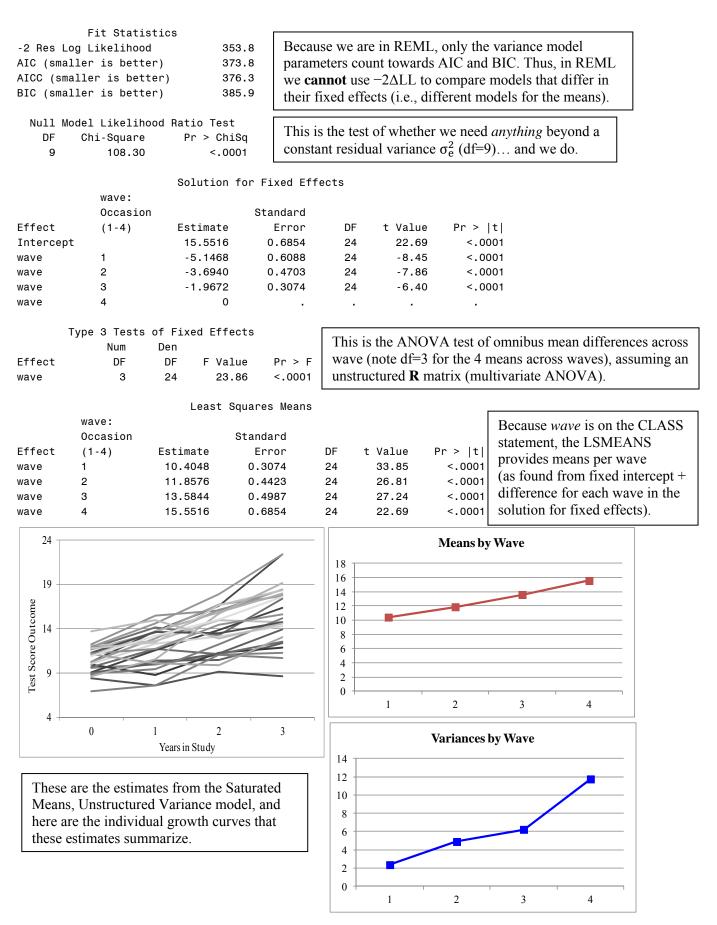
The models for this example come from Hoffman (2015) chapter 5. We will be examining the extent to which change in a test score outcome across four annual occasions can be described with fixed and random linear effects of time (indexed by years in study, in which 0 is baseline) in a sample of 25 persons.

SAS Syntax and Output for Data Manipulation:

```
* Location for files to be saved - CHANGE THIS TO YOUR DIRECTORY;
%LET filesave=C:\Dropbox\17_CLDP944\CLDP944_Example05;
LIBNAME filesave "&filesave.";
* Import data into work library, center time;
DATA work.Chapter5; SET filesave.SAS_Chapter5;
time = wave - 1; LABEL time= "time: Time in Study (0=1)"; RUN;
```

The **ANSWER KEY** for both the model for the means (via saturated means) and the model for the variance (via unstructured **R** matrix of all possible variances and covariances) is possible to estimate in balanced data:

```
TITLE1 'Ch 5: Saturated Means, Unstructured Variance Model';
TITLE2 'ANSWER KEY for both sides of the model';
PROC MIXED DATA=work.Chapter5 COVTEST NOCLPRINT IC METHOD=REML;
       CLASS PersonID wave;
       MODEL outcome = wave / SOLUTION DDFM=Satterthwaite;
       REPEATED wave / R RCORR TYPE=UN SUBJECT=PersonID;
       LSMEANS wave;
RUN; TITLE1; TITLE2;
            Dimensions
Covariance Parameters
                                 10
Columns in X
                                  5
Columns in Z
                                  0
Subjects
                                 25
Max Obs Per Subject
                                  4
           Estimated R Matrix for PersonID 1
 Row
           Col1
                        Col2
                                    Col3
                                                Col4
          2.3618
                      2.7867
                                  1.9566
                                              2.4204
   1
   2
          2.7867
                      4.8900
                                  4.0440
                                              5.5525
                                                          Because this model uses REPEATED
   3
          1.9566
                      4.0440
                                  6.2172
                                              7.7994
   4
          2.4204
                      5.5525
                                  7.7994
                                             11.7437
                                                          only (no RANDOM statement), the R
                                                          matrix holds the total variances and
     Estimated R Correlation Matrix for PersonID 1
                                                          covariances over waves directly.
 Row
           Col1
                        Col2
                                   Col3
                                                Col4
                                                          Likewise, RCORR holds the total
   1
          1.0000
                      0.8200
                                  0.5106
                                              0.4596
                                                          correlations over waves directly.
                                              0.7327
                      1.0000
  2
          0.8200
                                  0.7334
  3
          0.5106
                      0.7334
                                  1.0000
                                              0.9128
   4
          0.4596
                      0.7327
                                  0.9128
                                              1.0000
                 Covariance Parameter Estimates
                                   Standard
                                                    Ζ
Cov Parm
            Subject
                       Estimate
                                      Error
                                                Value
                                                             Pr Z
UN(1,1)
            PersonID
                            2.3618
                                        0.6818
                                                    3.46
                                                              0.0003
            PersonID
                            2.7867
                                                    3.11
                                                              0.0019
UN(2,1)
                                        0.8971
UN(2,2)
            PersonID
                            4.8900
                                        1.4116
                                                    3.46
                                                              0.0003
UN(3,1)
            PersonID
                            1.9566
                                                    2.23
                                                              0.0259
                                        0.8783
UN(3,2)
            PersonID
                            4.0440
                                        1.3958
                                                    2.90
                                                              0.0038
                                                              0.0003
UN(3,3)
            PersonID
                            6.2172
                                                    3.46
                                        1.7947
            PersonID
                            2.4204
                                                    2.05
                                                              0.0408
UN(4,1)
                                        1.1831
UN(4,2)
            PersonID
                            5.5525
                                        1.9176
                                                    2.90
                                                              0.0038
UN(4,3)
            PersonID
                            7.7994
                                        2.3615
                                                    3.30
                                                              0.0010
            PersonID
                           11.7437
                                        3.3901
                                                    3.46
                                                              0.0003
UN(4,4)
```



If an unstructured \mathbf{R} matrix was not possible to estimate, I'd still examine the answer key for the model for the means (via a saturated means model), but estimate a random intercept only (which should always be possible):

```
TITLE1 "Saturated Means, Random Intercept Variance Model -- MEANS ANSWER KEY";
PROC MIXED DATA=work.Chapter5 COVTEST NOCLPRINT IC METHOD=REML;
       CLASS PersonID wave;
       MODEL outcome = wave / SOLUTION DDFM=Satterthwaite;
       RANDOM INTERCEPT / G V VCORR TYPE=UN SUBJECT=PersonID;
       REPEATED wave / R TYPE=VC SUBJECT=PersonID;
       LSMEANS wave;
RUN;
           Estimated V Matrix for PersonID 1
 Row
            Col1
                         Co12
                                     Co13
                                                  Col4
                       4.0933
                                   4.0933
          6.3032
                                                4.0933
   1
   2
          4.0933
                       6.3032
                                   4.0933
                                                4.0933
   3
          4.0933
                       4.0933
                                   6.3032
                                                4.0933
                                                6.3032
   4
          4.0933
                       4.0933
                                   4.0933
     Estimated V Correlation Matrix for PersonID 1
 Row
                                     Col3
                                                  Co14
            Co11
                         Co12
          1.0000
                       0.6494
                                   0.6494
                                                0.6494
   1
   2
          0.6494
                       1.0000
                                   0.6494
                                                0.6494
   3
          0.6494
                       0.6494
                                   1.0000
                                                0.6494
          0.6494
                       0.6494
                                   0.6494
                                                1.0000
   4
                 Covariance Parameter Estimates
                                    Standard
                                                      7
Cov Parm
            Subject
                        Estimate
                                       Error
                                                  Value
                                                              Pr > Z
            PersonID
                             4.0933
                                                      3.04
                                                                 0.0012
UN(1,1)
                                          1.3443
wave
            PersonID
                             2.2099
                                          0.3683
                                                      6.00
                                                                 <.0001
           Fit Statistics
-2 Res Log Likelihood
                                 412.5
AIC (smaller is better)
                                 416.5
AICC (smaller is better)
                                 416.7
BIC (smaller is better)
                                 419.0
  Null Model Likelihood Ratio Test
    DF
          Chi-Square
                           Pr > ChiSq
     1
               49.51
                               <.0001
                          Solution for Fixed Effects
             wave:
             Occasion
                                      Standard
             (1-4)
Effect
                          Estimate
                                         Error
                                                     DF
                                                           t Value
                                                                       Pr > |t|
                                        0.5021
Intercept
                           15.5516
                                                   42.4
                                                              30.97
                                                                         <.0001
wave
             1
                           -5.1468
                                         0.4205
                                                     72
                                                             -12.24
                                                                         <.0001
wave
             2
                           -3.6940
                                         0.4205
                                                     72
                                                              -8.79
                                                                         <.0001
             3
                           -1.9672
                                         0.4205
                                                     72
                                                              -4.68
                                                                         <.0001
wave
             4
wave
                                 0
                                                      .
                                                                .
                                                                          .
        Type 3 Tests of Fixed Effects
                                                    This is the ANOVA test of omnibus mean differences across
              Num
                       Den
               DF
                                          Pr > F
                        DF
                              F Value
                                                    wave (note df=3 for the 4 means across waves), assuming a
Effect
                3
                        72
wave
                                55.82
                                          <.0001
                                                    random intercept only (CS V matrix; univariate ANOVA).
                            Least Squares Means
          wave:
                                   Standard
          Occasion
Effect
                       Estimate
                                      Error
                                                  DF
                                                        t Value
                                                                    Pr > |t|
          (1 - 4)
          1
                        10.4048
                                     0.5021
                                                42.4
                                                          20.72
                                                                      <.0001
wave
                                                42.4
          2
                        11.8576
                                     0.5021
                                                          23.62
                                                                      <.0001
wave
                                     0.5021
                                                42.4
                                                          27.05
wave
          З
                        13.5844
                                                                      <.0001
          4
                        15.5516
                                     0.5021
                                                42.4
                                                          30.97
                                                                      <.0001
wave
```

5.1: Empty Means, Random Intercept Model

<u>5.1: I</u>	Empty Means,	Random In	tercept M	odel			
		ork.Chapter			del"; IC METHOD=REML;	Level 1: Level 2:	$y_{ti} = \beta_{0i} + e_{ti}$ $\beta_{0i} = \gamma_{00} + U_{0i}$
	CLASS Perso	-		.	•.		$p_{0i} - p_{00} + C_{0i}$
	MODEL outcom				alte; ECT=PersonID;	Composite	$: y_{ti} = (\gamma_{00} + U_{0i}) + e_{ti}$
	REPEATED wa				-	composite	$f_{ti} = (f_{00} + G_{01}) + G_{ti}$
RUN;		,	0000000	1-10100			
	Estimated	R Matrix for	PersonID	1			
Row	Col1	Col2	Col3	Col4			
1	7.0554						
2		7.0554					
3			7.0554				
4				7.0554			
	Estimated	G Matrix					
		PersonID:					
		Person ID					
Row	Effect	number	Col1				
1	Intercept	1	2.8819				
		V Matrix for			Because this model	l uses the REPE	ATED and
Row	Col1	Col2	Col3	Col4	RANDOM stateme	ents, the V matri	x holds the
1	9.9373	2.8819	2.8819	2.8819	total variances and	· · · · · · · · · · · · · · · · · · ·	
2	2.8819	9.9373	2.8819	2.8819	(from putting G an		
3 4	2.8819 2.8819	2.8819 2.8819	9.9373 2.8819	2.8819 9.9373	the Z matrix). Like	•	<u> </u>
4	2.0019	2.0019	2.0019	9.9373	total correlations or		
E	Estimated V Cor	relation Matr	ix for Per	sonID 1			
Row	Col1	Col2	Col3	Col4			
1	1.0000	0.2900	0.2900	0.2900	VCORR provides t	he ICC as [.] IntV	ar/TotalVar
2	0.2900	1.0000	0.2900	0.2900	· contripionaes ·		
3	0.2900	0.2900	1.0000	0.2900			
4	0.2900	0.2900	0.2900	1.0000			
	Cov	ariance Paran	neter Estim Standard	ates Z			
Cov Pa	arm Subject	Estimate	Error	Value	Pr > Z		
UN(1,	,					dom intercept	variance in G
wave	PersonID	7.055	54 1.1	521 6.		idual variance	
AIC (s AICC	Fit Stati s Log Likelihoo smaller is bett (smaller is bet smaller is bett	d 5 er) 5 ter) 5	502.2 506.2 506.3 508.7				
Nul: Df	L Model Likelih ⁻ Chi-Square		niSa Thi		of whether we need the set (a_2, d_{1-1}) and u		
-	9.79	0.0	018 Inte		ice (so df=1) and w	/C UU.	
		olution for F Standarc	I				
Effect				t Value 29.81	Pr > t < 0001 This is	aamma00	
Inter	cept 12.849	6 0.4311	∠4	29.01	<.0001 This is	yammaoo	

5.3: Fixed Linear Time, Random Intercept Model

```
TITLE1 "Eq 5.3: Fixed Linear Time, Random Intercept Model";
PROC MIXED DATA= work.Chapter5 COVTEST NOCLPRINT IC
                                                                      Level 1:
                                                                                  y_{ti} = \beta_{0i} + \beta_{1i} (Time_{ti}) + e_{ti}
METHOD=REML;
  CLASS PersonID wave;
  MODEL outcome = time / SOLUTION DDFM=Satterthwaite;
                                                                      Level 2:
                                                                                  \beta_{0i} = \gamma_{00} + U_{0i}
  RANDOM INTERCEPT / G V VCORR TYPE=UN SUBJECT=PersonID;
  REPEATED wave / R TYPE=VC SUBJECT=PersonID;
                                                                                   \beta_{1i} = \gamma_{10}
  ESTIMATE "Intercept at Time 0" int 1 time 0;
  ESTIMATE "Intercept at Time 1" int 1 time 1;
                                                                      Composite: y_{ti} = (\gamma_{00} + U_{0i}) + \gamma_{10} (Time_{ti}) + e_{ti}
  ESTIMATE "Intercept at Time 2" int 1 time 2;
  ESTIMATE "Intercept at Time 3" int 1 time 3;
RUN;
```

Note the two different versions of the "time" variable in the syntax. Both are necessary here because they do different things. "**Wave**" is treated as a **categorical** predictor, and its role is to structure the **R** matrix in the event of missing data. Therefore, "wave" goes on the CLASS and REPEATED statements. In contrast, "**time**" is treated as a **continuous** predictor, and its role is to index linear effects of time (and it is centered such that wave 1 = time 0). Accordingly, in the ESTIMATE statements, only one value after "time" is needed.

Row 1 2 3 4	Estimated Coll 2.1725	R Matrix for Col2 2.1725	PersonID 1 Col3 2.1725	Col4 2.1725	After controlling for the fixed linear effect of time, the residual variance was reduced from $\sigma_e^2 = 7.06$ in the empty means, random intercept model to $\sigma_e^2 = 2.17$ in this model. This is a reduction of $(7.06 - 2.17) / 7.06 = .69$ (or 69% of the residual variance is accounted for by a fixed linear time).
	Estimated	G Matrix			
		PersonID:			However, the random intercept variance actually increased
		Person ID			from 2.88 to 4.10. This is because of how $\tau_{U_0}^2$ is found:
Row	Effect	number	Col1		true $\tau_{U_0}^2$ = observed $\tau_{U_0}^2 - (\sigma_e^2/n)$
1	Intercept	1	4.1026		So reducing σ_e^2 will make $\tau_{U_0}^2$ increase.
	Estimated	V Matrix for	PersonID 1		
Row	Col1	Col2	Col3	Col4	
1	6.2751	4.1026	4.1026	4.1026	
2	4.1026	6.2751	4.1026	4.1026	
3	4.1026	4.1026	6.2751	4.1026	
4	4.1026	4.1026	4.1026	6.2751	
Es	stimated V Cor	relation Matri	Lx for Perso	onID 1	
Row	Col1	Col2	Col3	Col4	
1	1.0000	0.6538	0.6538	0.6538	
2	0.6538	1.0000	0.6538	0.6538	
3	0.6538	0.6538	1.0000	0.6538	
4	0.6538	0.6538	0.6538	1.0000	
	Cov	ariance Parame	eter Estimat	es	
			Standard	Ζ	
Cov Par	rm Subject	Estimate	Error	Value	Pr > Z
UN(1,1)	-	4.1026	5 1.344	11 3	3.05 0.0011 Random intercept variance in G
wave	PersonID	2.1725	0.357		5.08 <.0001 Residual variance in R
	Fit Stati	stics			
-2 Res	Log Likelihoo		5.1 Are	wa allaw	red to examine the $-2\Delta LL$ to see if adding a fixed linear effect
	naller is bett				•
AICC (s	smaller is bet	ter) 41	9.2	me mpro	oved model fit in REML? If not, what do we do instead?
BIC (sr	naller is bett	er) 42	21.5		
			L		
Null	Model Likelih	ood Ratio Test	This	tosts wh	ether we need the random intercept
DF	Chi-Square				*
1	51.12	<.00		ance (80 (df=1) and we (still) do.

	Solut	ion for Fix	ed Effec	ts				
		Standard						
Effect	Estimate	Error	DF	t Value	Pr > t			
Intercept	10.2745	0.4743	34.7	21.66	<.0001	this	is	gamma00
time	1.7167	0.1318	74	13.02	<.0001	this	is	gamma10

	Type 3 Tests	or Fixed	ETTECTS	
	Num	Den		
Effect	DF	DF I	F Value	Pr > F

LIICOL			Varue	11 - 1
time	1	74	169.57	<.0001

Estimates -> These are the predicted outcome means from a fixed linear time model

		Scanuaru			
Label	Estimate	Error	DF	t Value	Pr > t
Intercept at Time O	10.2745	0.4743	34.7	21.66	<.0001
Intercept at Time 1	11.9912	0.4361	25.1	27.50	<.0001
Intercept at Time 2	13.7080	0.4361	25.1	31.43	<.0001
Intercept at Time 3	15.4247	0.4743	34.7	32.52	<.0001

Ctondond

5.5: Random Linear Time Model

Note that the "time" variable gets included in the RANDOM statement, not "wave"—including "wave" would result in model non-convergence, because it would try to estimate a random slope variance for each possible difference between waves (instead of a single variance for a continuous random slope through all the waves).

Row 1 2 3 4	Col1 0.6986	d R Matrix fo Col2 0.6986 timated G Mat PersonID:	or PersonID 1 Col3 0.6986 trix	Col4 0.6986	After adding a random linear effect of time, the residual variance is smaller, but it is not correct to say that it has been reduced. Random effects do not explain variance; they simply re-allocate it. Here, this means that part of what was residual is now individual differences in the linear effect of time as a new pile of variance in the G matrix below.
		Person ID			
Row	Effect	number	Col1	Col2	
1	Intercept	1	2.2624	0.05454	The G matrix provides the variances and covariances of the
2	time	1	0.05454	0.9089	individual random effects. Now G is a 2x2 matrix because we have 2 random effects (intercept, linear slope).
	Estimated	d G Correlat:	ion Matrix		we have 2 fundom enfocts (intercept, initial stope).
		PersonID:			The GCORR matrix provides the correlation(s) among the
		Person ID			individual random effects. Here, the individual intercepts
Row	Effect	number	Col1	Col2	· · · ·
1	Intercept	1	1.0000	0.03803	and slopes are correlated $r = .04$.
2	time	1	0.03803	1.0000	

Level 1: $y_{ti} = \beta_{0i} + \beta_{1i} (Time_{ti}) + e_{ti}$ Level 2: $\beta_{0i} = \gamma_{00} + U_{0i}$ $\beta_{1i} = \gamma_{10} + U_{1i}$

Composite: $y_{ti} = (\gamma_{00} + U_{0i}) + (\gamma_{10} + U_{1i}) (Time_{ti}) + e_{ti}$

Estimated V Matrix for PersonID 1
Row Col1 Col2 Col3 Col4
1 2.9611 2.3170 2.3715 2.4260
2 2.3170 3.9790 4.2438 5.2073 The V matrix holds the total variances and
3 2.3715 4.2438 6.8148 7.9885 covariances over waves (from putting G and R
4 2.4260 5.2073 7.9885 11.4684 back together through the Z matrix). Likewise,
VCORR holds the total correlations over
Estimated V Correlation Matrix for PersonID 1
Row Col1 Col2 Col3 Col4 waves. Note that all of these are now predicted
1 1.0000 0.6750 0.5279 0.4163 to differ as a function of which wave it is (see
2 0.6750 1.0000 0.8150 0.7709 table 5.2 for a description of how this works).
3 0.5279 0.8150 1.0000 0.9036
4 0.4163 0.7709 0.9036 1.0000
Covariance Parameter Estimates
Standard Z
Cov Parm Subject Estimate Error Value Pr Z
UN(1,1) PersonID 2.2624 0.8003 2.83 0.0023 Random intercept variance in G
UN(2,1) PersonID 0.05454 0.3507 0.16 0.8764 Random intercept-slope covariance in G
UN(2,2) PersonID 0.9089 0.3040 2.99 0.0014 Random linear slope variance in G wave PersonID 0.6986 0.1397 5.00 <.0001 Residual variance in R
wave PersonID 0.6986 0.1397 5.00 <.0001 Residual variance in R
Fit Statistics
-2 Res Log Likelihood $_{366.7}$ Are we allowed to examine the $-2\Delta LL$ to see if adding a random linear
AIC (smaller is better) 374.7 effect of time improved model fit in REML? If so, how many model
AICC (smaller is better) 375.2 parameters have we added?
BIC (smaller is better) 379.6
Null Model Likelihood Ratio Test This tests whether we need <i>anything</i> in the G matrix (so df=3). Note this
3 99.47 <.0001 does NOT tell us if we need the random linear time slope specifically!
Solution for Fixed Effects
Standard
Effect Estimate Error DF t Value $Pr > t $
Intercept 10.2745 0.3318 24 30.97 <.0001 this is gamma00
time 1.7167 0.2048 24 8.38 <.0001 this is gamma10
Type 3 Tests of Fixed Effects
Num Den
Effect DF DF F Value Pr > F
time 1 24 70.26 <.0001
Estimates \rightarrow These are the predicted outcome means from a random linear time model
Standard
Label Estimate Error DF t Value Pr > t
Intercept at Time 0 10.2745 0.3318 24 30.97 <.0001
Intercept at Time 1 11.9912 0.3736 24 32.09 <.0001
Intercept at Time 2 13.7080 0.5030 24 27.25 <.0001
Intercept at Time 3 15.4247 0.6711 24 22.98 <.0001
We can use the model estimates to calculate 95% random effects confidence intervals that describe the
predicted range of individual random effects:

Random Effect 95% CI = fixed effect $\pm (1.96*\sqrt{\text{Random Variance}})$

Intercept 95% CI =
$$\gamma_{00} \pm \left(1.96^* \sqrt{\tau_{U_0}^2}\right) = 10.27 \pm \left(1.96^* \sqrt{2.26}\right) = 7.32 \text{ to } 13.22$$

Linear Time Slope 95% CI = $\gamma_{10} \pm \left(1.96^* \sqrt{\tau_{U_1}^2}\right) = 1.72 \pm \left(1.96^* \sqrt{0.91}\right) = -0.15 \text{ to } 3.59$

Last but not least: there may still be residual covariances after modeling individual differences in the linear effect of time (i.e., adding a random linear slope to the **G** matrix). We can test alternative **R** matrix assumptions besides VC (which assumes no residual covariance/correlation over time) to see if this is the case:

```
TITLE1 "Random Linear Time Model + AR1 R Matrix";
PROC MIXED DATA= work.Chapter5 COVTEST NOITPRINT NOCLPRINT IC METHOD=REML;
       CLASS PersonID wave;
       MODEL outcome = time / SOLUTION DDFM=Satterthwaite;
       RANDOM INTERCEPT time / G GCORR V VCORR TYPE=UN SUBJECT=PersonID;
       REPEATED / R RCORR TYPE=AR(1) SUBJECT=PersonID;
RUN:
           Estimated R Matrix for PersonID 1
                                    Co13
                                                Co14
 Row
            Col1
                        Co12
                                            0.000012
   1
          0.7193
                     0.01841
                                0.000471
   2
         0.01841
                     0.7193
                                0.01841
                                            0.000471
   3
        0.000471
                     0.01841
                                 0.7193
                                            0.01841
        0.000012
                   0.000471
                                 0.01841
                                             0.7193
   4
     Estimated R Correlation Matrix for PersonID 1
 Row
           Col1
                        Co12
                                    Co13
                                                Co14
         1.0000
                     0.02560
                                0.000655
                                            0.000017
   1
  2
        0.02560
                     1.0000
                                0.02560
                                            0.000655
   З
        0.000655
                     0.02560
                                 1.0000
                                            0.02560
   4
        0.000017
                    0.000655
                                 0.02560
                                              1.0000
           Fit Statistics
-2 Res Log Likelihood
                                366.7
                                          The -2LL is not smaller than the random linear time model, so adding
AIC (smaller is better)
                                376.7
AICC (smaller is better)
                                377.4
                                          an AR1 correlation to the R matrix does not improve model fit.
BIC (smaller is better)
                                382.8
TITLE1 "Random Linear Time Model + TOEP2 R Matrix";
PROC MIXED DATA=example5 COVTEST NOITPRINT NOCLPRINT IC METHOD=REML;
       CLASS PersonID wave;
       MODEL outcome = time / SOLUTION DDFM=Satterthwaite;
       RANDOM INTERCEPT time / G GCORR V VCORR TYPE=UN SUBJECT=PersonID;
       REPEATED wave / R RCORR TYPE=TOEP(2) SUBJECT=PersonID;
RUN;
           Estimated R Matrix for PersonID 1
 Row
            Co11
                        Co12
                                    Co13
                                                Co14
  1
          0.7127
                     0.01259
   2
         0.01259
                     0.7127
                                 0.01259
                     0.01259
                                             0.01259
   3
                                 0.7127
                                 0.01259
   4
                                             0.7127
     Estimated R Correlation Matrix for PersonID 1
 Row
            Col1
                        Co12
                                    Co13
                                                Co14
          1.0000
                     0.01766
   1
  2
         0.01766
                     1.0000
                                 0.01766
   3
                     0.01766
                                 1.0000
                                             0.01766
                                              1.0000
   4
                                 0.01766
           Fit Statistics
-2 Res Log Likelihood
                                366.7
                                          The -2LL is not smaller than the random linear time model, so adding a
AIC (smaller is better)
                                376.7
AICC (smaller is better)
                                377.4
                                          lag-1 covariance to the R matrix does not improve model fit, either.
BIC (smaller is better)
                                382.8
```