Fitting growth models with SAS PROC MIXED

Introduction: Variance Component & Repeated Measures Models

May 29, 2001

Steve Gregorich

Middle Schools Data

- 591 students assessed in 6th, 7th, 8th, and 9th grade
  6th grade N=591
  7th grade N=544
  8th grade N=522
  9th grade N=403

- Outcome
  Average score on self-reported ***opportunities for sex*** scale

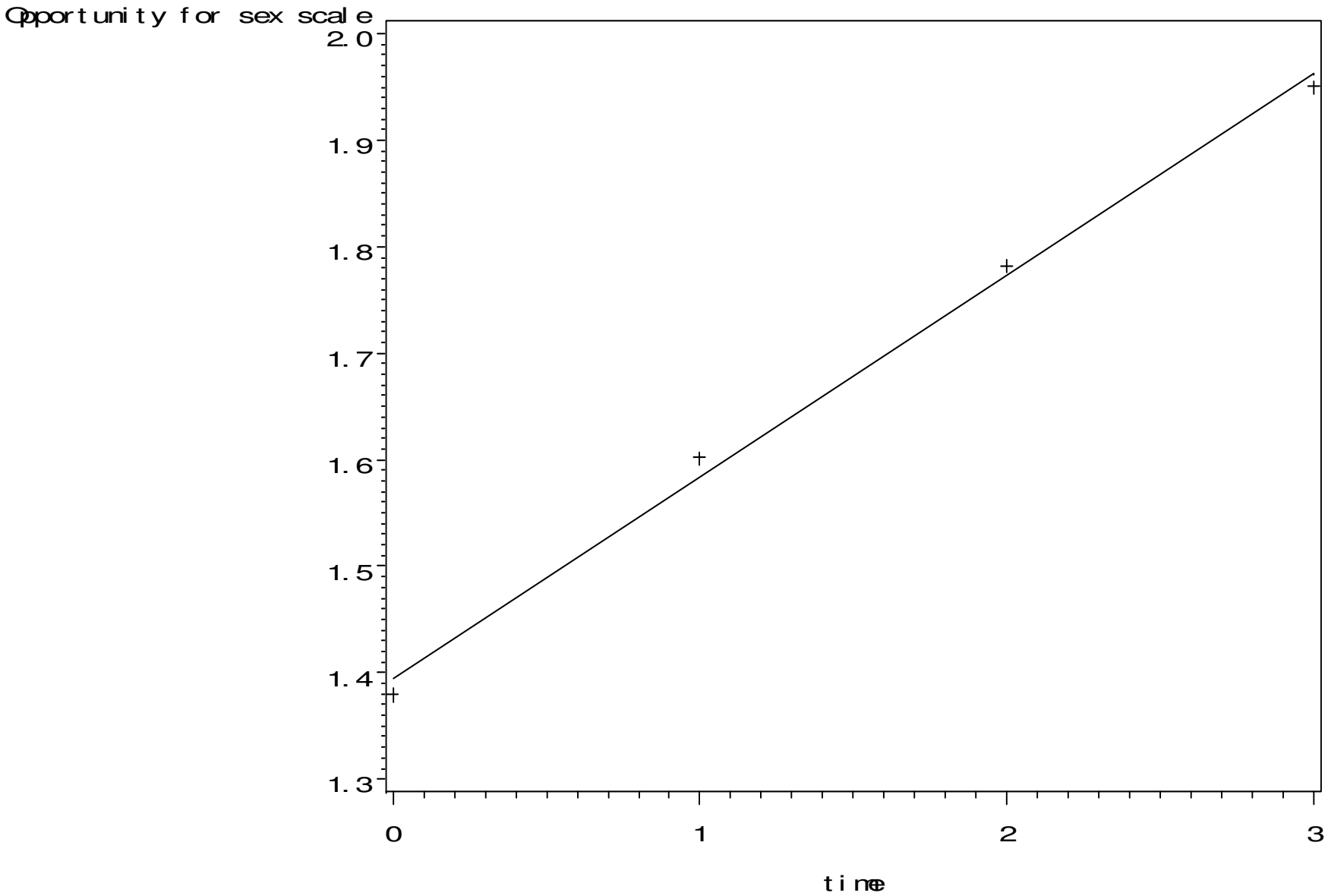
In the past 3 months how often have you
  been alone with someone you are very attracted to?
  been alone kissing and touching someone you really like?
  laid down on a couch or bed alone with someone you really like?
  been at a party where there were no adults in the house?

0= never, **1**= 1 or 2 times, **2**= 3 to 5 times, **3**= 6 or more times.

Research Questions

What is the trajectory of OPTUN4SX across time?

What type & magnitude of dependencies exist among repeated measures?


Opportunity for sex scale
2.0
1.9
1.8
1.7
1.6
1.5
1.4
1.3
0
1
2
3
time

Variance Component Models

- How much variation in OPTUN4SX scores is attributable to individual? How much to repeated measures within students?
- What is the intra-student correlation of OPTUN4SX scores?

The Unconditional Variance Components Model

$OPTUN4SX_{ij}$ = grand_mean + student_level_resid + time_level_resid

$$= \beta_{00} + \mathbf{u}_{0j} + \mathbf{r}_{ij}$$

$$\mathbf{u}_{0j} \sim N(0,\tau_{00}) \qquad \mathbf{r}_{ij} \sim N(0, \sigma^2)$$

$COV(\mathbf{u}_{0j}, \mathbf{r}_{ij}) = 0$

Approach

- Fit a model with no explanatory variables, only a random intercept
- Implicitly--not explicitly--an intercept is estimated for each student
- β_{00}, $\mathbf{u}_{0j}$, and $\mathbf{r}_{ij}$ explained
- The student-level variance component, τ_{00}, represents the variance of student OPTUN4SX scores at time 0 around the grand mean.
- The time-level variance component, σ^2, represents the variance of repeated OPTUN4SX scores around their student mean.

Additional Model Assumptions

- Total, between-student, and within-student OPTUN4SX variances are constant across schools.
- Students are independent
- Compound-symmetric block-diagonal covariance structure

$$
\begin{bmatrix}
\tau_{00}+\sigma^2 & \tau_{00} & \tau_{00} & 0 & 0 & . & 0 & 0 & 0 & 0 \\
\tau_{00} & \tau_{00}+\sigma^2 & \tau_{00} & 0 & 0 & . & 0 & 0 & 0 & 0 \\
\tau_{00} & \tau_{00} & \tau_{00}+\sigma^2 & 0 & 0 & . & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \tau_{00}+\sigma^2 & \tau_{00} & . & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \tau_{00} & \tau_{00}+\sigma^2 & . & 0 & 0 & 0 & 0 \\
. & . & . & . & . & . & . & . & . & . \\
0 & 0 & 0 & 0 & 0 & . & \tau_{00}+\sigma^2 & \tau_{00} & \tau_{00} & \tau_{00} \\
0 & 0 & 0 & 0 & 0 & . & \tau_{00} & \tau_{00}+\sigma^2 & \tau_{00} & \tau_{00} \\
0 & 0 & 0 & 0 & 0 & . & \tau_{00} & \tau_{00} & \tau_{00}+\sigma^2 & \tau_{00} \\
0 & 0 & 0 & 0 & 0 & . & \tau_{00} & \tau_{00} & \tau_{00} & \tau_{00}+\sigma^2
\end{bmatrix}
$$

PROC MIXED Syntax and Results

```
proc mixed noclprint covtest;
  class id ;
  model optun4sx = / solution;
  random intercept / subject=id;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
Intercept	ID	0.2568	0.02232	11.51	<.0001
Residual		0.4141	0.01518	27.27	<.0001

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	1.6904	0.02537	595	66.62	<.0001

-2 Res Log Likelihood	4719.6

Results Summary

Fixed Effect

- Grand mean coursework score = 1.69

Variance Components

- Between-student variation in mean OPTUN4SX scores = 0.2568
- Within-student variation in OPTUN4SX scores = 0.4141

Intra-student correlation

ρ = 0.2568 / (0.2568 + 0.4141)
= 0.38

Growth Model:
Random Intercept with Fixed Effect of Time

$$OPTUN4SX_{ij} = \beta_{00} + time \times \beta_1 + \mathbf{u}_{0j} + \mathbf{r}_{ij}$$

$$\mathbf{u}_{0j} \sim N(0,\tau_{00}) \qquad \mathbf{r}_{ij} \sim N(0, \sigma^2) \qquad COV(\mathbf{u}_{0j}, \mathbf{r}_{ij}) = 0$$

PROC MIXED Syntax and Results

```
proc mixed noclprint covtest;
  class id ;
  model optun4sx = time / solution;
  random intercept / subject=id;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
Intercept	ID	0.2773	0.02195	12.63	<.0001
Residual		0.3303	0.01212	27.25	<.0001

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	1.3876	0.02956	595	46.94	<.0001
time	0.2316	0.01183	1469	19.57	<.0001

-2 Res Log Likelihood	4384.8

Model Comparison

Fixed Effects	**intercept only**	**+ time**
intercept	1.69	1.39
time	.	0.23
Random Effects		
τ_{00}	0.26	0.28
σ^2	0.41	0.33
-2LL	4719.6	4384.8
Δ -2LL	334.8 (1)	

all estimates, $p < .001$

Growth Model:
Adding a Random Slope Term

$$OPTUN4SX_{ij} = \pi_{0j} + time \times \pi_{1j} + \mathbf{r}_{ij}$$

$$\pi_{0j} = \beta_{00} + \mathbf{u}_{0j}$$

$$\pi_{1j} = \beta_{10} + \mathbf{u}_{1j}$$

$$OPTUN4SX_{ij} = \beta_{00} + time \times \beta_{01} + \mathbf{u}_{0j} + time \times \mathbf{u}_{1j} + \mathbf{r}_{ij}$$

where,

$$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix}\right]$$

PROC MIXED Syntax and Results

```
proc mixed noclprint covtest;
  class id ;
  model optun4sx = time /solution;
  random intercept time/ subject=id type=un;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	ID	0.1612	0.02196	7.34	<.0001
UN(2,1)	ID	0.02115	0.009245	2.29	0.0221
UN(2,2)	ID	0.04612	0.007070	6.52	<.0001
Residual		0.2596	0.01186	21.90	<.0001

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	1.3845	0.02438	595	56.78	<.0001
time	0.2362	0.01422	562	16.62	<.0001

-2 Res Log Likelihood 4219.6

Model Comparison

Fixed Effects	**fixed time**	**random time**
intercept	1.39	1.39
time	0.23	0.24
Random Effects		
τ_{00}	0.28	0.16
τ_{10}	.	0.02*
τ_{11}	.	0.05
σ^2	0.33	0.26
-2LL	4384.8	4219.6
Δ -2LL	165.2 (2)	

all estimates $p < .001$ except for *, $p < .05$

Revisiting the Variance Components Approach

$\sigma_w^2 + \sigma_b^2$	.	.	.
σ_b^2	$\sigma_w^2 + \sigma_b^2$	.	.
σ_b^2	σ_b^2	$\sigma_w^2 + \sigma_b^2$	.
σ_b^2	σ_b^2	σ_b^2	$\sigma_w^2 + \sigma_b^2$

Another Approach

- Do not estimate variance components (random intercepts and slopes)
- Assume intercepts and growth rates are constant across people
- Estimate different covariance structures for the residuals
- Choose 'best' model with respect to fit and parsimony

$$OPTUN4SX_{ij} = \pi_{0j} + time \times \pi_{1j} + \mathbf{r}_{ij}$$

$$\pi_{0j} = \beta_{00}$$

$$\pi_{1j} = \beta_{10}$$

$$OPTUN4SX_{ij} = \beta_{00} + time \times \beta_{01} + \mathbf{r}_{ij}$$

where, $\mathbf{r}_{ij} \sim N(0,\Sigma)$

Unstructured (10)

σ_{11}^2	.	.	.
σ_{21}^2	σ_{22}^2	.	.
σ_{31}^2	σ_{32}^2	σ_{33}^2	.
σ_{41}^2	σ_{42}^2	σ_{43}^2	σ_{44}^2

PROC MIXED Syntax

```
proc mixed noclprint covtest;
  class id _time_;
  model optun4sx = time /solution;
  repeated _time_/ subject=id type=un r=6
                   rcorr=6;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	ID	0.3113	0.01804	17.26	<.0001
UN(2,1)	ID	0.2206	0.02019	10.93	<.0001
UN(2,2)	ID	0.5836	0.03538	16.50	<.0001
UN(3,1)	ID	0.2229	0.02272	9.81	<.0001
UN(3,2)	ID	0.3767	0.03304	11.40	<.0001
UN(3,3)	ID	0.7549	0.04659	16.20	<.0001
UN(4,1)	ID	0.1568	0.02557	6.13	<.0001
UN(4,2)	ID	0.3195	0.03752	8.52	<.0001
UN(4,3)	ID	0.4987	0.04570	10.91	<.0001
UN(4,4)	ID	0.9245	0.06451	14.33	<.0001

Unstructured Covariance Matrix

0.31	.	.	.
0.22	0.58	.	.
0.22	0.38	0.75	.
0.16	0.32	0.50	0.92

Unstructured Correlation Matrix

1.00			
0.52	1.00		
0.46	0.57	1.00	
0.29	0.44	0.60	1.00

-2LL = 4114.4 (10 parameters)

Compound Symmetric (Exchangeable; 2)

$\sigma^2 + \sigma_1^2$	.	.	.
σ_1^2	$\sigma^2 + \sigma_1^2$	.	.
σ_1^2	σ_1^2	$\sigma^2 + \sigma_1^2$	.
σ_1^2	σ_1^2	σ_1^2	$\sigma^2 + \sigma_1^2$

PROC MIXED Syntax

```
proc mixed noclprint covtest;
  class id _time_;
  model optun4sx = time /solution;
  repeated _time_/ subject=id type=cs r=6
                   rcorr=6;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
CS	ID	0.2773	0.02195	12.63	<.0001
Residual		0.3303	0.01212	27.25	<.0001

Compound Symmetric Covariance Matrix

0.61	.	.	.
0.28	0.61	.	.
0.28	0.28	0.61	.
0.28	0.28	0.28	0.61

Compound Symmetric Correlation Matrix

1.00	.	.	.
0.46	1.00	.	.
0.46	0.46	1.00	.
0.46	0.46	0.46	1.00

-2LL = 4384.8 (- 4114.4 = 270.4 w/ 8 df)

(note: same as VC with fixed time effect)

Compound Symmetric w/ Heterogeneous Variances (5)

σ_1^2	.	.	.
ρ	σ_2^2	.	.
ρ	ρ	σ_3^2	.
ρ	ρ	ρ	σ_4^2

PROC MIXED Syntax

```
proc mixed noclprint covtest;
  class id _time_;
  model optun4sx = time /solution;
  repeated _time_/ subject=id type=csh r=6
                   rcorr=6;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
Var(1)	ID	0.3230	0.01881	17.18	<.0001
Var(2)	ID	0.5688	0.03369	16.88	<.0001
Var(3)	ID	0.7232	0.04333	16.69	<.0001
Var(4)	ID	0.9504	0.06554	14.50	<.0001
CSH	ID	0.4806	0.02280	21.08	<.0001

Compound Symmetric Het Covariance Matrix

0.32	.	.	.
0.21	0.57	.	.
0.23	0.31	0.72	.
0.27	0.35	0.40	0.95

Compound Symmetric Het Correlation Matrix

1.00	.	.	.
0.48	1.00	.	.
0.48	0.48	1.00	.
0.48	0.48	0.48	1.00

-2LL = 4192.7 (- 4114.4 = 78.3 w/ 5 df)

1st-Order Autoregressive (2)

σ^2	.	.	.
ρ	σ^2	.	.
ρ^2	ρ	σ^2	.
ρ^3	ρ^2	ρ	σ^2

PROC MIXED Syntax

```
proc mixed noclprint covtest;
  class id _time_;
  model optun4sx = time /solution;
  repeated _time_/ subject=id type=ar(1) r=6
                   rcorr=6;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
AR(1)	ID	0.5440	0.01930	28.18	<.0001
Residual		0.6071	0.02279	26.64	<.0001

1st-Order Autorgrssv Covariance Matrix

0.61			
0.33	0.61		
0.18	0.33	0.61	
0.10	0.18	0.33	0.61

1st-Order Autorgrssv Correlation Matrix

1.00	.	.	.
0.54	1.00	.	.
0.30	0.54	1.00	.
0.16	0.30	0.54	1.00

-2LL = 4334.5 (- 4114.4 = 220.1 w/ 8 df)

1st-Order Autoregressive w/ Heterogeneous Variances (5)

σ_1^2	.	.	.
ρ	σ_2^2	.	.
ρ^2	ρ	σ_3^2	.
ρ^3	ρ^2	ρ	σ_4^2

PROC MIXED Syntax

```
proc mixed noclprint covtest;
  class id _time_;
  model optun4sx = time /solution;
  repeated _time_/ subject=id type=arh(1) r=6
                   rcorr=6;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
Var(1)	ID	0.3201	0.01817	17.62	<.0001
Var(2)	ID	0.5954	0.03557	16.74	<.0001
Var(3)	ID	0.7369	0.04380	16.83	<.0001
Var(4)	ID	0.8971	0.05962	15.05	<.0001
ARH(1)	ID	0.5580	0.01976	28.24	<.0001

1st-Order Autorgrssv Het Covariance Matrix

0.32	.	.	.
0.24	0.60	.	.
0.15	0.40	0.74	.
0.09	0.23	0.45	0.90

1st-Order Autorgrsv Het Correlation Matrix

1.00	.	.	.
0.56	1.00	.	.
0.31	0.56	1.00	.
0.17	0.31	0.56	1.00

-2LL = 4169.1 (- 4114.4 = 54.7 w/ 5 df)

Toeplitz (4)

σ^2	.	.	.
σ_1	σ^2	.	.
σ_2	σ_1	σ^2	.
σ_3	σ_2	σ_1	σ^2

PROC MIXED Syntax

```
proc mixed noclprint covtest;
  class id _time_;
  model optun4sx /solution;
  repeated _time_/ subject=id type=toep r=6
                   rcorr=6;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
TOEP(2)	ID	0.3283	0.02242	14.64	<.0001
TOEP(3)	ID	0.2461	0.02304	10.68	<.0001
TOEP(4)	ID	0.1464	0.02887	5.07	<.0001
Residual		0.6056	0.02332	25.97	<.0001

Toeplitz Covariance Matrix

0.61	.	.	.
0.33	0.61	.	.
0.25	0.33	0.61	.
0.15	0.25	0.33	0.61

Toeplitz Correlation Matrix

1.00	.	.	.
0.54	1.00	.	.
0.41	0.54	1.00	.
0.24	0.41	0.54	1.00

-2LL = 4297.5 (- 4114.4 = 183.1 w/ 6 df)

Toeplitz w/ Heterogeneous Variances (7)

σ_1^2	.	.	.
ρ_1	σ_2^2	.	.
ρ_2	ρ_1	σ_3^2	.
ρ_3	ρ_2	ρ_1	σ_4^2

PROC MIXED Syntax

```
proc mixed noclprint covtest;
  class id _time_;
  model optun4sx = time /solution;
  repeated _time_/ subject=id type=toeph r=6
                   rcorr=6;
```

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
Var(1)	ID	0.3167	0.01808	17.52	<.0001
Var(2)	ID	0.5931	0.03574	16.60	<.0001
Var(3)	ID	0.7375	0.04437	16.62	<.0001
Var(4)	ID	0.9039	0.06064	14.91	<.0001
TOEPH(1)	ID	0.5558	0.02150	25.85	<.0001
TOEPH(2)	ID	0.4458	0.02796	15.94	<.0001
TOEPH(3)	ID	0.2934	0.04212	6.96	<.0001

Toeplitz Het. Covariance Matrix

0.32	.	.	.
0.24	0.59	.	.
0.22	0.37	0.74	.
0.16	0.33	0.45	0.90

Toeplitz Het. Correlation Matrix

1.00	.	.	.
0.56	1.00	.	.
0.45	0.56	1.00	.
0.29	0.44	0.56	1.00

-2LL = 4119.8 (- 4114.4 = 5.4 w/ 3 df)

Combined VC and RM model

```
proc mixed noclprint covtest;
  class id _time_;
  model optun4sx = time / solution;
  random           time / subject= id
                          type   = simple g;
  repeated _time_       / subject= id
                          type   = toeph
                          r=6  rcorr=6;
```

(Note. solutions for models with random intercepts were improper)

Results

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
time	ID	0.01495	0.01110	1.35	0.0890
Var(1)	ID	0.3123	0.01810	17.26	<.0001
Var(2)	ID	0.5752	0.03729	15.43	<.0001
Var(3)	ID	0.6850	0.05703	12.01	<.0001
Var(4)	ID	0.7907	0.1023	7.73	<.0001
TOEPH(1)	ID	0.5407	0.02564	21.08	<.0001
TOEPH(2)	ID	0.4475	0.02978	15.03	<.0001
TOEPH(3)	ID	0.3198	0.04964	6.44	<.0001

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	1.3617	0.02280	1502	59.74	<.0001
time	0.2428	0.01385	562	17.53	<.0001

-2 Res Log Likelihood 4130.5

(note. fit is worse than RM TOEPH model)