CHAPTER 3

SIMULATION METHODOLOGY AND PARAMETER ESTIMATES FOR SIMULATION STUDY

3.1 Simulation Methodology

Using the PR data set a simulation study was carried out to determine the robustness of linear mixed effects models under misspecification of the covariance structure. The PR data set, as explained in Chapter one, refers to a data set obtained by researchers at the University of North Carolina Dental School who followed the growth of 27 children (16 males, 11 females) from age 8 until age 14. Every two years they measured the distance between the pituitary and the pterygomaxillary fissure using x-ray exposures of the side of the head. This distance was used as the response variable, with age as a quantitative predictor variable and gender as a qualitative predictor variable (Potthoff & Roy, 1964).

Before data could be simulated under the various covariance structures, the model parameters were first estimated through fitting these models to the original data. These parameter estimates were taken as the true parameter values and used to generate the simulated data sets. SAS® PROC MIXED (SAS ver. 9.1) was used to carry out the analyses. This software allows the user to choose a covariance structure for the random effects and for the errors. All of the available covariance structures can be selected for either the random effects or the errors. The classical linear mixed effect model has the error covariance structure $\boldsymbol{\omega}_i = \sigma^2\mathbf{I}$, whereas the general linear mixed effect model allows the covariance structure of both the errors and the random effects to be arbitrarily specified (Wolfinger, 1993). Wolfinger (1993) and Jennrich and Schluchter (1986) discuss the available covariance structures and methods for parameter estimation of these covariance structures. Covariance structures available include VC, CS, AR(1), TOEP and UN, as well as a number of spatial covariance structures. What is made clear by these two studies is that there are a large number of possible combinations available. The covariance structures considered included all the covariance structures discussed in Chapter two, namely VC, CS, AR(1), CSH, ARH(1), TOEP and US, which are those particularly suitable for temporal longitudinal data. SAS® PROC MIXED (SAS ver.9.1) allows the user to model the correlation between observations by means of the error covariance structure alone, or through the inclusion of random effects. Both of these methods were considered. The user can also choose the number of random effects to be included in the model. In this study the random intercept and the random intercept and slope methods were both considered.

The estimated parameter values and variance component values were used to simulate data under the assumed model. In order to simulate the data, the linear mixed effects model below was used:

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\varepsilon}_i \quad \text{for } i = 1,\ldots N$$
$$\mathbf{b}_i \sim \mathbf{N(0,\Sigma)},\ \ \boldsymbol{\varepsilon}_i \sim \mathbf{N}(\mathbf{0},\boldsymbol{\omega}_i)$$

where the parameters are defined as in Chapters one and two. For the simulation of the data, the same regressor matrices, $\mathbf{X}_i$ and $\mathbf{Z}_i$, were used as in the original model fitting. The parameter values for $\boldsymbol{\beta}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\omega}_i$ were estimated from the original data. The random effects, $\mathbf{b}_i$, and random errors, $\boldsymbol{\varepsilon}_i$, were generated from a normal distribution with a zero mean vector and with a covariance matrix equal to the estimated $\boldsymbol{\Sigma}$ and $\boldsymbol{\omega}_i$ respectively. The package R (R Development Core Team 2007) was used to generate the simulated data. The code used for this procedure is given in Appendix C1. At this stage of the simulation process, some models did not converge to produce valid parameter estimates, or in other cases warning messages were produced by the procedure due to invalid parameter estimates. These models were not considered any further.

Two hundred and fifty simulated data sets were obtained for each of the assumed models. In some cases simulated data could not be obtained from the available model parameters, as at least one of the covariance matrices was not positive definite. These problems did not appear in a warning message during the first stage, and therefore these models were still retained for later fitting purposes. Since a researcher, unaware that the covariance matrices were not positive definite, could still have considered these as appropriate models, therefore they were kept for the purposes of comparison, although these covariance structures were not considered to be robust options. Certain covariance matrices were retained for comparative purposes. In Table 3.1, discussed below, those covariance matrices marked with an asterisk (*) are non-positive definite matrices, and those marked a double asterisk (**) contain variance components equal to zero. The significance of the fixed effects estimates is indicated by the presence of either (##) or (#) beside the fixed effects estimate, indicating significance at the 1% and 5% significance levels respectively.

The maximum number of likelihood evaluations and the maximum number of iterations to take place during the optimisation procedure can be set. The default settings for these values are 150 and 50 respectively, and these defaults were used during the parameter estimation procedure. The default convergence criterion for PROC MIXED (SAS ver 9.1) was also used, which is the relative Hessian convergence criterion with tolerance equal to 1×10^{-8}. This criterion is defined as

$$\frac{\mathbf{g}_k'\mathbf{H}_k^{-1}\mathbf{g}_k}{|f_k|}\le 1\times10^{-8} \quad \text{(SAS PROC MIXED, 2003)}$$

where $\mathbf{g}_k$ is the gradient of the objective function, f_k, and $\mathbf{H}_k$ is the Hessian of the objective function. The default method of estimation is REML, which was also kept for this study.

To estimate the degrees of freedom for the t-test of the fixed effects parameter estimates, the Satterthwaite degrees of freedom option was used, specified in the MODEL statement of PROC MIXED (SAS ver.9.1). This method bases the degrees of freedom on the chi-squared distribution which best approximates the distribution of $\mathbf{h}'(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})\mathbf{h}$, where $\mathbf{h}'\boldsymbol{\beta}$ is a linear combination of the fixed effects parameters (Duchateau & Janssen, 1997). This method for estimating the degrees of freedom for the t-statistic is recommended by Verbeke and Molenberghs (2000), but they also note that for longitudinal data, since different subjects contribute independent information, resulting in large degrees of freedom, the different estimation methods for the degrees of freedom lead to very similar p-values.

Table 3.1: Parameter estimates under the various covariance structures or reasons for exclusion from simulation study. Estimates from successfully fitted models were used to generate the simulated data.

Models with successful parameter estimation				
$\boldsymbol{\omega_i}$	$\boldsymbol{\Sigma}$	Estimated $\boldsymbol{\omega_i}$	Estimated $\boldsymbol{\Sigma}$	$\boldsymbol{\beta}$
VC OLS Estimates	None	$\begin{pmatrix} 5.0938 & & & \\ & 5.0938 & & \\ & & 5.0938 & \\ & & & 5.0938 \end{pmatrix}$		$\begin{pmatrix} 17.3727\#\# \\ -1.0321 \\ 0.4795\#\# \\ 0.3048\# \end{pmatrix}$
	Intercept Only	$\begin{pmatrix} 1.9221 & & & \\ & 1.9221 & & \\ & & 1.9221 & \\ & & & 1.9221 \end{pmatrix}$	3.2986	$\begin{pmatrix} 17.3727\#\# \\ -1.0321 \\ 0.4795\#\# \\ 0.3048\# \end{pmatrix}$
	VC	$\begin{pmatrix} 1.8646 & & & \\ & 1.8646 & & \\ & & 1.8646 & \\ & & & 1.8646 \end{pmatrix}$	$\begin{pmatrix} 2.4168 & \\ & 0.007747 \end{pmatrix}$	$\begin{pmatrix} 17.3727\#\# \\ -1.0321 \\ 0.4795\#\# \\ 0.3048\# \end{pmatrix}$
	CS	$\begin{pmatrix} 1.9880 & & & \\ & 1.9880 & & \\ & & 1.9880 & \\ & & & 1.9880 \end{pmatrix}$	$\begin{pmatrix} 0.02168 & 0.02168 \\ & 0.02168 \end{pmatrix}^{**}$	$\begin{pmatrix} 17.3727\#\# \\ -1.0321 \\ 0.4795\#\# \\ 0.3048\# \end{pmatrix}$
	CSH	$\begin{pmatrix} 1.7163 & & & \\ & 1.7163 & & \\ & & 1.7163 & \\ & & & 1.7163 \end{pmatrix}$	$\begin{pmatrix} 5.7821 & -0.2893 \\ & 0.03249 \end{pmatrix}$	$\begin{pmatrix} 17.3727\#\# \\ -1.0321 \\ 0.4795\#\# \\ 0.3048\# \end{pmatrix}$
	ARH(1)	$\begin{pmatrix} 1.7163 & & & \\ & 1.7163 & & \\ & & 1.7163 & \\ & & & 1.7163 \end{pmatrix}$	$\begin{pmatrix} 5.7821 & -0.2893 \\ & 0.03249 \end{pmatrix}$	$\begin{pmatrix} 17.3727\#\# \\ -1.0321 \\ 0.4795\#\# \\ 0.3048\# \end{pmatrix}$
	UN	$\begin{pmatrix} 1.7162 & & & \\ & 1.7162 & & \\ & & 1.7162 & \\ & & & 1.7162 \end{pmatrix}$	$\begin{pmatrix} 5.7864 & -0.2896 \\ & 0.03252 \end{pmatrix}$	$\begin{pmatrix} 17.3727\#\# \\ -1.0321 \\ 0.4795\#\# \\ 0.3048\# \end{pmatrix}$
CS	None	$\begin{pmatrix} 5.2207 & 3.2986 & 3.2986 & 3.2986 \\ & 5.2207 & 3.2986 & 3.2986 \\ & & 5.2207 & 3.2986 \\ & & & 5.2207 \end{pmatrix}$		$\begin{pmatrix} 17.3727\#\# \\ -1.0321 \\ 0.4795\#\# \\ 0.3048\# \end{pmatrix}$
	CSH	$\begin{pmatrix} 3.6510 & 1.9348 & 1.9348 & 1.9348 \\ & 3.6510 & 1.9348 & 1.9348 \\ & & 3.6510 & 1.9348 \\ & & & 3.6510 \end{pmatrix}$	$\begin{pmatrix} 3.8517 & -0.2896 \\ & 0.03253 \end{pmatrix}$	$\begin{pmatrix} 17.3727\#\# \\ -1.0321 \\ 0.4795\#\# \\ 0.3048\# \end{pmatrix}$

Table 3.1 (cont.): Parameter estimates under the various covariance structures or reasons for exclusion from simulation study. Estimates from successfully fitted models were used to generate the simulated data.

$\boldsymbol{\omega_i}$	$\boldsymbol{\Sigma}$	Estimated $\boldsymbol{\omega_i}$	Estimated $\boldsymbol{\Sigma}$	$\boldsymbol{\beta}$
CS	TOEP	$\begin{pmatrix} 3.6510 & 1.9348 & 1.9348 & 1.9348 \\ & 3.6510 & 1.9348 & 1.9348 \\ & & 3.6510 & 1.9348 \\ & & & 3.6510 \end{pmatrix}$	$\begin{pmatrix} 0.03251 & -0.2895 \\ & 0.03251 \end{pmatrix}*$	$\begin{pmatrix} 17.3727\#\# \\ -1.0321 \\ 0.4795\#\# \\ 0.3048\# \end{pmatrix}$
CSH	None	$\begin{pmatrix} 5.6959 & 3.1205 & 3.7414 & 3.3304 \\ & 4.2362 & 3.2266 & 2.8721 \\ & & 6.0896 & 3.4436 \\ & & & 4.8253 \end{pmatrix}$		$\begin{pmatrix} 17.4046\#\# \\ -1.2737 \\ 0.4788\#\# \\ 0.3156\# \end{pmatrix}$
	CSH	$\begin{pmatrix} 3.9049 & 1.0782 & 1.2889 & 0.3484 \\ & 2.2372 & 0.9756 & 0.2637 \\ & & 3.1966 & 0.3152 \\ & & & 0.2335 \end{pmatrix}$	$\begin{pmatrix} 1.2423 & -0.1584 \\ & 0.04054 \end{pmatrix}$	$\begin{pmatrix} 17.4201\#\# \\ -1.5428 \\ 0.4767\#\# \\ 0.3507\# \end{pmatrix}$
	ARH(1)	$\begin{pmatrix} 3.9049 & 1.0782 & 1.2889 & 0.3484 \\ & 2.2372 & 0.9756 & 0.2637 \\ & & 3.1966 & 0.3152 \\ & & & 0.2335 \end{pmatrix}$	$\begin{pmatrix} 1.2423 & -0.1584 \\ & 0.04054 \end{pmatrix}$	$\begin{pmatrix} 17.4201\#\# \\ -1.5428 \\ 0.4767\#\# \\ 0.3507\# \end{pmatrix}$
	UN	$\begin{pmatrix} 3.9045 & 1.0779 & 1.2885 & 0.3480 \\ & 2.2369 & 0.9753 & 0.2634 \\ & & 3.1962 & 0.3148 \\ & & & 0.2331 \end{pmatrix}$	$\begin{pmatrix} 1.2429 & -0.1584 \\ & 0.04055 \end{pmatrix}$	$\begin{pmatrix} 17.4201\#\# \\ -1.5428 \\ 0.4767\#\# \\ 0.3507\# \end{pmatrix}$
AR(1)	None	$\begin{pmatrix} 5.2145 & 3.2564 & 2.0336 & 1.2700 \\ & 5.2145 & 3.2564 & 2.0336 \\ & & 5.2145 & 3.2564 \\ & & & 5.2145 \end{pmatrix}$		$\begin{pmatrix} 17.3206\#\# \\ -0.7215 \\ 0.4838\#\# \\ 0.2854 \end{pmatrix}$
	Intercept only	$\begin{pmatrix} 1.8854 & -0.07077 & 0.002656 & -0.00010 \\ & 1.8854 & -0.07077 & 0.002656 \\ & & 1.8854 & -0.07077 \\ & & & 1.8854 \end{pmatrix}$	3.3355	$\begin{pmatrix} 17.3762\#\# \\ -1.0510 \\ 0.4792\#\# \\ 0.3062\# \end{pmatrix}$
	VC	$\begin{pmatrix} 1.7583 & -0.1822 & 0.01888 & -0.00196 \\ & 1.7583 & -0.1822 & 0.01888 \\ & & 1.7583 & -0.1822 \\ & & & 1.7583 \end{pmatrix}$	$\begin{pmatrix} 2.3624 & \\ & 0.009372 \end{pmatrix}$	$\begin{pmatrix} 17.3824\#\# \\ -1.0842 \\ 0.4785\#\# \\ 0.3086\# \end{pmatrix}$
	CSH	$\begin{pmatrix} 1.1926 & -0.5644 & 0.2671 & -0.1264 \\ & 1.1926 & -0.5644 & 0.2671 \\ & & 1.1926 & -0.5644 \\ & & & 1.1926 \end{pmatrix}$	$\begin{pmatrix} 11.3738 & -0.8147 \\ & 0.08452 \end{pmatrix}$	$\begin{pmatrix} 17.4171\#\# \\ -1.2647 \\ 0.4757\#\# \\ 0.3222\# \end{pmatrix}$

Table 3.1 (cont.): Parameter estimates under the various covariance structures or reasons for exclusion from simulation study. Estimates from successfully fitted models were used to generate the simulated data.

$\boldsymbol{\omega_i}$	$\boldsymbol{\Sigma}$	Estimated $\boldsymbol{\omega_i}$	Estimated $\boldsymbol{\Sigma}$	**B**
AR(1)	ARH(1)	$\begin{pmatrix} 1.1926 & -0.5644 & 0.2671 & -0.1264 \\ & 1.1926 & -0.5644 & 0.2671 \\ & & 1.1926 & -0.5644 \\ & & & 1.1926 \end{pmatrix}$	$\begin{pmatrix} 11.3738 & -0.8147 \\ & 0.08452 \end{pmatrix}$	$\begin{pmatrix} 17.4171\#\# \\ -1.2647 \\ 0.4757\#\# \\ 0.3222\# \end{pmatrix}$
	UN	$\begin{pmatrix} 1.1924 & -0.5643 & 0.2671 & -0.1264 \\ & 1.1924 & -0.5643 & 0.2671 \\ & & 1.1926 & -0.5644 \\ & & & 1.1924 \end{pmatrix}$	$\begin{pmatrix} 11.3775 & -0.8150 \\ & 0.08455 \end{pmatrix}$	$\begin{pmatrix} 17.4171\#\# \\ -1.2647 \\ 0.4757\#\# \\ 0.3222\# \end{pmatrix}$
ARH(1)	None	$\begin{pmatrix} 5.8149 & 3.2682 & 2.3721 & 1.3043 \\ & 4.5807 & 3.3246 & 1.8281 \\ & & 6.0172 & 3.3087 \\ & & & 4.5369 \end{pmatrix}$		$\begin{pmatrix} 17.3523\#\# \\ -0.9805 \\ 0.4827\#\# \\ 0.3009\# \end{pmatrix}$
	VC	$\begin{pmatrix} 2.7111 & -0.1975 & 0.02288 & -0.00161 \\ & 1.4452 & -0.16756 & 0.01176 \\ & & 1.9499 & -0.1369 \\ & & & 0.9661 \end{pmatrix}$	$\begin{pmatrix} 1.8325 & \\ & 0.01283 \end{pmatrix}$	$\begin{pmatrix} 17.4109\#\# \\ -1.5006 \\ 0.4768\#\# \\ 0.3431\# \end{pmatrix}$
	CSH	$\begin{pmatrix} 1.6176 & -0.5901 & 0.3553 & -0.05427 \\ & 1.0533 & -0.6342 & 0.09687 \\ & & 1.8685 & -0.2854 \\ & & & 0.2134 \end{pmatrix}$	$\begin{pmatrix} 11.9301 & -0.8616 \\ & 0.08717 \end{pmatrix}$	$\begin{pmatrix} 17.4203\#\# \\ -1.6346 \\ 0.4760\#\# \\ 0.3559\# \end{pmatrix}$
	ARH(1)	$\begin{pmatrix} 1.6176 & -0.5901 & 0.3553 & -0.05427 \\ & 1.0533 & -0.6342 & 0.09687 \\ & & 1.8685 & -0.2854 \\ & & & 0.2134 \end{pmatrix}$	$\begin{pmatrix} 11.9301 & -0.8616 \\ & 0.08717 \end{pmatrix}$	$\begin{pmatrix} 17.4203\#\# \\ -1.6346 \\ 0.4760\#\# \\ 0.3559\# \end{pmatrix}$
	UN	$\begin{pmatrix} 1.6176 & -0.5901 & 0.3553 & -0.05427 \\ & 1.0533 & -0.6342 & 0.09687 \\ & & 1.8685 & -0.2854 \\ & & & 0.2134 \end{pmatrix}$	$\begin{pmatrix} 11.9301 & -0.8616 \\ & 0.08717 \end{pmatrix}$	$\begin{pmatrix} 17.4203\#\# \\ -1.6346 \\ 0.4760\#\# \\ 0.3559\# \end{pmatrix}$
TOEP	None	$\begin{pmatrix} 5.2826 & 3.3659 & 3.6804 & 2.5285 \\ & 5.2826 & 3.3659 & 3.6804 \\ & & 5.2826 & 3.3659 \\ & & & 5.2826 \end{pmatrix}$		$\begin{pmatrix} 17.4089\#\# \\ -1.1385 \\ 0.4759\#\# \\ 0.3214\# \end{pmatrix}$
	CS	$\begin{pmatrix} 9.5596 & 7.7466 & 8.4209 & 7.7515 \\ & 9.5596 & 7.7466 & 8.4209 \\ & & 9.5596 & 7.7466 \\ & & & 9.5596 \end{pmatrix}$	$\begin{pmatrix} 0.03251 & -0.2895 \\ & 0.03251 \end{pmatrix}^{*}$	$\begin{pmatrix} 17.4152\#\# \\ -1.2252 \\ 0.4757\#\# \\ 0.3223\# \end{pmatrix}$

Table 3.1 (cont.): Parameter estimates under the various covariance structures or reasons for exclusion from simulation study. Estimates from successfully fitted models were used to generate the simulated data.

$\boldsymbol{\omega_i}$	$\boldsymbol{\Sigma}$	Estimated $\boldsymbol{\omega_i}$	Estimated $\boldsymbol{\Sigma}$	$\boldsymbol{\beta}$
TOEP	TOEP	$\begin{pmatrix} 9.5613 & 7.7483 & 8.4227 & 7.7534 \\ & 9.5613 & 7.7483 & 8.4227 \\ & & 9.5613 & 7.7483 \\ & & & 9.5613 \end{pmatrix}$	$\begin{pmatrix} 0.05769 & -0.5286 \\ & 0.05769 \end{pmatrix}*$	$\begin{pmatrix} 17.4152\#\# \\ -1.2252 \\ 0.4757\#\# \\ 0.3223\# \end{pmatrix}$
UN	None	$\begin{pmatrix} 5.4252 & 2.7092 & 3.8411 & 2.7151 \\ & 4.1906 & 2.9745 & 3.3137 \\ & & 6.2632 & 4.1332 \\ & & & 4.9862 \end{pmatrix}$		$\begin{pmatrix} 17.4254\#\# \\ -1.5831 \\ 0.4764\#\# \\ 0.3504\# \end{pmatrix}$

$\boldsymbol{\omega_i}$	$\boldsymbol{\Sigma}$	**Details**
Final Hessian matrix not positive definite		
CS	Intercept only	Final Hessian matrix not positive definite
	VC	Final Hessian matrix not positive definite
	UN	Final Hessian matrix not positive definite. Variance component estimated as zero
CSH	Intercept only	Final Hessian matrix not positive definite. Variance component estimated as zero
ARH(1)	Intercept only	Final Hessian matrix not positive definite
TOEP	Intercept only	Final Hessian matrix not positive definite
	VC	Final Hessian matrix not positive definite
	UN	Final Hessian matrix not positive definite
UN	Intercept only	Final Hessian matrix not positive definite
	VC	Final Hessian matrix not positive definite
	CS	Final Hessian matrix not positive definite
	CSH	Final Hessian matrix not positive definite. Variance estimated as zero
	AR(1)	Final Hessian matrix not positive definite
	ARH(1)	Final Hessian matrix not positive definite
	TOEP	Final Hessian matrix not positive definite
	UN	Final Hessian matrix not positive definite. Variance estimated close to zero
Covariance of random effects not positive definite		
CS	CS	Estimated $\boldsymbol{\Sigma}$ matrix not positive definite
CSH	CS	Estimated $\boldsymbol{\Sigma}$ matrix not positive definite
	TOEP	Estimated $\boldsymbol{\Sigma}$ matrix not positive definite
AR(1)	CS	Variance component estimated as zero
ARH(1)	TOEP	Estimated $\boldsymbol{\Sigma}$ matrix not positive definite

Table 3.1 (cont.): Parameter estimates under the various covariance structures or reasons for exclusion from simulation study. Estimates from successfully fitted models were used to generate the simulated data.

ω_i	**Σ**	**Details**
Parameter estimates out of bounds		
VC	TOEP	Variances estimated as zero
CS	AR(1)	Correlation parameter, ρ, estimated as -1
	ARH(1)	When this model was fitted to simulated data sets simulated from this model, the estimates for ρ were not valid
CSH	VC	Variances estimated as zero
AR(1)	AR(1)	Correlation parameter, ρ, estimated as 1
	TOEP	Variances estimated as zero
ARH(1)	CS	Variance component estimated as zero
	AR(1)	Correlation parameter, ρ, estimated as 1
TOEP	CSH	Correlation parameter, ρ, estimated as -1
	AR(1)	Correlation parameter, ρ, estimated as -1
	ARH(1)	Correlation parameter, ρ, estimated as -1
Nonconvergence		
VC	AR(1)	Likelihood tended to infinity
CSH	AR(1)	Failed to converge. On last iteration invalid estimate of -1 for ρ obtained

Note: Only the upper half the covariance matrices is shown. Matrices are symmetrical. Blank spaces in the upper half of the matrices indicate zero values.

* indicates non-positive (semi-)definite matrices. Model retained for comparative purposes.
** indicates that variance component estimated as zero. Model retained for comparative purposes.
# indicates a significant fixed effect parameter estimate at the 5% level of significance.
indicates a significant fixed effect parameter estimate at the 1% level of significance.

3.2 Parameter Estimation

The estimated fixed effects model is of the form:

$$\hat{L}ength = \hat{\beta}_0 + \hat{\beta}_1 Gender + \hat{\beta}_2 Age + \hat{\beta}_3 Gender \times Age$$

where *Gender* = 0 specifies a boy and *Gender* = 1 specifies a girl, and *Age* takes on values 8, 10, 12 and 14. To take the first model in Table 3.1, the model with the VC error structure and no random effects (i.e. the OLS model), as an example, the estimated model is

$$\hat{L}ength = 17.3727 - 1.0321 Gender + 0.4795 Age + 0.3048 Gender \times Age \ .$$

This model implies that the intercept of girls is lower compared to the boys, but that the girls have a steeper upward slope, shown by the positive interaction term of *Gender*×*Age*. The difference in intercepts between the boys and the girls is not significant, but the slopes are significantly (at the 5% level) different.

The conditional form of the model for the random intercept models is

$$\hat{L}ength_i = \hat{\beta}_0 + \hat{\beta}_1 Gender + \hat{\beta}_2 Age + \hat{\beta}_3 Gender \times Age + \hat{b}_i$$

and for the random intercept and slope models is

$$\hat{L}ength_i = \hat{\beta}_0 + \hat{\beta}_1 Gender + \hat{\beta}_2 Age + \hat{\beta}_3 Gender \times Age + \hat{b}_{1i} + \hat{b}_{2i} Age \ .$$

During the first stage of the simulation study many models were eliminated as possible robust models. Before a model could be considered as a robust model, valid estimates for the fixed effects and variance parameters needed to be obtained for the model fitted to the original data. Table 3.1 summarises the results of the initial parameter fitting process.

Firstly it should be noted that for all the models that were successfully fitted, the estimates of the fixed effects were very similar, if not exactly the same. The estimate which seemed to vary the most was that of the *Gender* estimate, which is the difference in intercepts between the line for the girls and the line for the boys. The p-value for this parameter estimate was non-significant for all models under consideration. This effect was kept in the model as the interaction between *Gender* and *Age* was significant for the majority of the models. Only in the cases of the OLS model and the no random effects AR(1) model was the interaction term not significant. Therefore, in most cases the researcher would have come to the same conclusion regardless of the covariance structure chosen. The OLS model obtained exactly the same parameter estimates compared to all other models with $\boldsymbol{\omega}_{\boldsymbol{i}}$ = VC. Since the interaction term of this model is non-significant, it indicates a large standard error estimate from the OLS model compared to other models which do not assume an independent model covariance structure.

In general, those models that had fewer parameters tended to be more likely to obtain valid estimates compared to more complicated models, although there were a large number of models with few parameters that were also not fitted successfully (Table 3.2). All models not containing random effects were successfully fitted, demonstrating that in some circumstances including random effects can come at the cost of poor estimates or no estimates at all. To support these observations, the percentage convergence of different categories of models was calculated. There were 42 models fitted with random effects covariance structures that did not force the variance of the intercept and the slope to be equal (referred to as models with appropriate covariance structures), and 61.90% of these were successfully fitted to the data. There were 21

Table 3.2: Number of covariance parameter estimates for each model arranged in ascending order, with AIC for fitted models or reason for failure to obtain parameter estimates.

ω_i	Σ	Number ω_i parameters	Number Σ parameters	Total number covariance parameters	Reason for failure to fit model	AIC
VC	None	1	0	1		485.6
AR(1)	None	2	0	2		448.6
CS	None	2	0	2		437.8
VC	Intercept only	1	1	2		437.8
AR(1)	Intercept only	2	1	3		439.7
CS	Intercept only	2	1	3	Final Hessian not positive definite	
VC	AR(1)	1	2	3	Non-convergence	
VC	CS	1	2	3		440.1
VC	TOEP	1	2	3	Parameter estimate out of bounds	
VC	VC	1	2	3		439.2
AR(1)	AR(1)	2	2	4	Parameter estimate out of bounds	
AR(1)	CS	2	2	4	$\hat{\Sigma}$ not positive definite	
AR(1)	TOEP	2	2	4	Parameter estimate out of bounds	
AR(1)	VC	2	2	4		440.8
CS	AR(1)	2	2	4	Parameter estimate out of bounds	
CS	CS	2	2	4	$\hat{\Sigma}$ not positive definite	
CS	TOEP	2	2	4		440.6
CS	VC	2	2	4	Final Hessian not positive definite	
TOEP	None	4	0	4		437.4
VC	ARH(1)	1	3	4		440.6
VC	CSH	1	3	4		440.6
VC	UN	1	3	4		440.6
AR(1)	ARH(1)	2	3	5		438.8
AR(1)	CSH	2	3	5		438.8
AR(1)	UN	2	3	5		438.8
CS	ARH(1)	2	3	5	Parameter estimate out of bounds	
CS	CSH	2	3	5		442.6
CS	UN	2	3	5	Final Hessian not positive definite	
TOEP	Intercept only	4	1	5	Final Hessian not positive definite	
TOEP	AR(1)	4	2	6	Parameter estimate out of bounds	
TOEP	CS	4	2	6		440.7
TOEP	TOEP	4	2	6		440.7

Table 3.2 (cont.): Number of covariance parameter estimates for each model arranged in ascending order, with AIC for fitted models or reason for failure to obtain parameter estimates.

ω_i	Σ	Number ω_i parameters	Number Σ parameters	Total number covariance parameters	Reason for failure to fit model	AIC
TOEP	VC	4	2	6	Final Hessian not positive definite	
TOEP	ARH(1)	4	3	7	Parameter estimate out of bounds	
TOEP	CSH	4	3	7	Parameter estimate out of bounds	
TOEP	UN	4	3	7	Final Hessian not positive definite	
ARH(1)	None	10	0	10		452.8
CSH	None	10	0	10		442.0
UN	None	10	0	10		444.5
ARH(1)	Intercept only	10	1	11	Final Hessian not positive definite	
CSH	Intercept only	10	1	11	Final Hessian not positive definite	
UN	Intercept only	10	1	11	Final Hessian not positive definite	
ARH(1)	AR(1)	10	2	12	Parameter estimate out of bounds	
ARH(1)	CS	10	2	12	Parameter estimate out of bounds	
ARH(1)	TOEP	10	2	12	$\hat{\Sigma}$ not positive definite	
ARH(1)	VC	10	2	12		443.9
CSH	AR(1)	10	2	12	Non-convergence	
CSH	CS	10	2	12	$\hat{\Sigma}$ not positive definite	
CSH	TOEP	10	2	12	$\hat{\Sigma}$ not positive definite	
CSH	VC	10	2	12	Parameter estimate out of bounds	
UN	AR(1)	10	2	12	Final Hessian not positive definite	
UN	CS	10	2	12	Final Hessian not positive definite	
UN	TOEP	10	2	12	Final Hessian not positive definite	
UN	VC	10	2	12	Final Hessian not positive definite	
ARH(1)	ARH(1)	10	3	13		441.0
ARH(1)	CSH	10	3	13		441.0
ARH(1)	UN	10	3	13		441.0
CSH	ARH(1)	10	3	13		443.9

Table 3.2 (cont.): Number of covariance parameter estimates for each model arranged in ascending order, with AIC for fitted models or reason for failure to obtain parameter estimates.

ω_i	Σ	Number ω_i parameters	Number Σ parameters	Total number covariance parameters	Reason for failure to fit model	AIC
CSH	CSH	10	3	13		443.9
CSH	UN	10	3	13		443.9
UN	ARH(1)	10	3	13	Final Hessian not positive definite	
UN	CSH	10	3	13	Final Hessian not positive definite	
UN	UN	10	3	13	Final Hessian not positive definite	

models fitted with appropriate covariance structures with six or less covariance parameters, and of these 71.43% were successfully fitted to the data. Twelve models with six or less covariance parameters were fitted with inappropriate random effects covariance structures, and of these only 33.33% converged. Twenty-one models were fitted with appropriate covariance structures and with more than six covariance parameters. Of these models, 47.62% were successfully fitted to the data. Eighteen random effects models were fitted to the data with appropriate random effects covariance structures and with more than six parameters, and of these models only 38.89% were successfully fitted. Three no random effects models were fitted with more than six parameters, and 100.00% of these models were successfully fitted to the data.

The PR data set has a total of 27 subjects with four observation each, resulting in a total number of observations of 108. A variety of models, of varying complexity (i.e. having large differences in the number of parameters to be estimated), were considered. The OLS, with four fixed effects parameters and one covariance parameter, results in an observation to parameter ratio of 21.6:1. The no random effects models have a maximum number of covariance parameters of ten, leading to fourteen parameters in total and an observation to parameter ratio of 7.7:1. The random intercept models require an additional 27 estimates for the random effects, and one additional covariance parameter, leading to a maximum total number of parameters of 42 and an observation to parameter ratio of 2.6:1. The random intercept and slope models require 54 random effects parameter estimates and up to three random effects covariance parameters, leading to a total number of parameters of 71, resulting in an observation to parameter ratio of 1.5:1. The fewer observations there are per parameter, the more sensitive the parameter estimates will be to individual data points. Generally, it is accepted that for regression models an observation to parameter ratio of at least 4:1 is required (J Galpin, personal communication from DM Hawkins, 1985). Hocking (2005, p. 42) states that “in general, we would like to have at least six to ten observations per predictor” in the case of a multiple linear regression model. Therefore it is expected that the high parameter models will show high sensitivity to individual data points.

From the AIC and AICc values obtained for the fitted models, it seems that for the original data, simpler models were preferred, with the no random effects model with $\boldsymbol{\omega}_i$ = TOEP obtaining the lowest AIC of 437.4, followed by the CS models (no random effects model with $\boldsymbol{\omega}_i$ = CS and the random intercept model with $\boldsymbol{\omega}_i$ = VC) obtaining the minimum AIC value (Table 3.2). The OLS model (no random effects model with $\boldsymbol{\omega}_i$ = VC) obtained the highest AIC value. Of the random effects models, the CSH random error models with heterogeneous random effects models or US random effects models obtained the highest AIC value, and the no random effects model (other than the OLS model) which had the highest AIC value was the ARH(1) no random effects model. These results agree with those obtained by Jennrich and Schluchter (1986), Verbeke and Molenberghs (2000), and Davis (2002), but many more covariance structures are compared in addition to those used in these studies, as discussed in Chapter two.

Three main problems occurred during the fitting procedure. The most common problem was indicated by the warning message "Convergence criteria met but final Hessian not positive definite". This means that a point on the cost function curve has been reached where the gradient is equal to zero, but the curvature of the cost function is not downwards in all directions (Weiss, 2005). Basically, this means the optimisation procedure was unable to reach a local maximum, and the details of SAS PROC MIXED (2003) state that in general this indicates that the model has been over parameterised and that duplicate parameters may have been fitted. Collinearity in the predictor variables can also result in destabilising the optimisation procedure, and if present may result in this error message (Verbeke & Molenberghs, 2000). In this case collinearity cannot be the cause as there was only one continuous predictor variable, and so the problem is related to inappropriate covariance structure specification, or inability of the optimisation procedure to locate a local maximum.

The second type of problem that occurred was when fitted models produced estimates that were outside of the parameter boundaries. The details of the SAS PROC MIXED (2003) indicate that the estimate of the correlation parameter ρ needs to be such that -1< $\hat{\rho}$ <1. Therefore fitted models with $\hat{\rho}$ = 1 or -1 have invalid estimates, which would indicate that the model has been over parameterised and that a model with fewer parameters would be more appropriate. In other cases, variance components were estimated as zero, again indicating that a simpler model would be a better choice. In more serious cases, variances of the random effects were estimated as zero. Since it is assumed that the random effects are normally distributed, a zero estimate for the variance is invalid. Sometimes it was not obvious that the estimates of the parameters were invalid. The covariance structure of the random effects was sometimes found not to be positive (semi-)definite. This resulted in a failure to simulate data from these parameters as it is assumed that the covariance structures of the random effects and errors are positive (semi-)definite as discussed in Section 2.2.3. These problems of invalid parameter estimates were only found through investigation of the model estimates; no warning message was produced in the SAS PROC MIXED (ver. 9.1) output. Some of these models were used after the data had been simulated for comparative purposes. These include models with non-positive (semi-)definite random effects covariance structures, which were the models with $\boldsymbol{\omega}_{\boldsymbol{i}}$ = TOEP and $\boldsymbol{\Sigma}$ = CS, $\boldsymbol{\omega}_{\boldsymbol{i}}$ = CS and $\boldsymbol{\Sigma}$ = TOEP, and $\boldsymbol{\omega}_{\boldsymbol{i}}$ = TOEP and $\boldsymbol{\Sigma}$ = TOEP, as well as the model with $\boldsymbol{\omega}_{\boldsymbol{i}}$ = VC and $\boldsymbol{\Sigma}$ = CS, which had a variance component equal to zero, resulting in the variances and covariances all having the same value.

The three models included with estimated $\boldsymbol{\Sigma}$ non-positive definite were included as the non-positive definiteness was not evident without calculating the eigenvalues of these matrices, and therefore the estimates for the covariance structure could easily have been accepted if further investigation was not carried out on $\boldsymbol{\Sigma}$. The AIC values obtained for these models were not unreasonable, therefore these models may potentially have been considered as good models (Table 3.2). One model with a covariance parameter estimated as zero was also included in order to compare how well this model, with an unnecessary covariance parameter, performed compared to other models. This model (model with $\boldsymbol{\omega_i}$ = VC and Σ = CS) obtained a reasonable AIC value so may also have been considered as a good model if the covariance parameters were not investigated.

Failure of the optimisation algorithm to converge was the third problem that occurred in the fitting procedure, but it only occurred for the model with $\boldsymbol{\omega_i}$ = CSH and Σ = AR(1). In this particular case there were too many likelihood evaluations before convergence of the optimisation algorithm occurred. The number of likelihood evaluation was increased to 1000, but still no convergence occurred. The convergence criterion cycled between very small values (<0.0006) to extremely large values ($>10^{12}$), but always above the convergence level.

Errors or invalid parameter estimates occurred for all models where the specified covariance matrix of the random effects forced the variance of the intercept and the slope to be equal, such as in the cases of CS, TOEP and AR(1) options (Table 3.1). SAS PROC MIXED (ver. 9.1) seems to have dealt with this problem in the VC case, as variance estimates for the intercept and slope are different when VC is specified for $\boldsymbol{\Sigma}$. When VC is specified for $\boldsymbol{\omega_i}$, all of the diagonal elements are equal. All models with four or less covariance parameters which failed to obtain valid estimates, except in the case of the random intercept model with $\boldsymbol{\omega_i}$ = CS and the random intercept and slope model with $\boldsymbol{\omega_i}$ = CS and $\boldsymbol{\Sigma}$ = VC, had specifications for $\boldsymbol{\omega_i}$ that forced the diagonal elements of the covariance matrix to be equal (Table 3.2).

Since the covariance matrix of the random effects in this case is 2×2, it is easy to state the condition of positive definiteness when the diagonal of the matrix has equal elements. A random intercept and slope model, with equal variances for the intercept and slope, would have a 2×2 symmetric matrix of the following form $\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{11} \end{pmatrix}$, where a_{11}>0. For this matrix to be positive definite, the eigenvalues would need to be positive (Johnson, 1970), therefore solving for the determinant $\begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{12} & a_{11}-\lambda \end{vmatrix}$ results in eigenvalues equal to $a_{12} + a_{11}$ and $a_{11} – a_{12}$. Since the eigenvalues need to be positive, this implies that $-a_{11} < a_{12} < a_{11}$. The problem with random effects covariance structures that weren't positive definite in Table 3.1 is that the absolute covariance is too large for the small variance estimates that were obtained. If the variances of the intercept and slope are allowed to differ, the estimates are found to be very far apart from each other; therefore the covariance would need to be large as well. To take the cases of the random effects models with $\boldsymbol{\omega}_i$ = CS and $\boldsymbol{\Sigma}$ = CS and with $\boldsymbol{\omega}_i$ = CS and $\boldsymbol{\Sigma}$ = CSH, the estimated $\boldsymbol{\Sigma}$ matrix of the first model, which assumes equal diagonal elements, was non-positive definite, whereas the $\boldsymbol{\Sigma}$ matrix estimated for the second model, which allows the diagonal elements to differ, is positive definite. The diagonal elements estimated for the second model are 3.8517 and 0.0325 for the random intercept and slope respectively, which are very different from each other, and therefore indicate that assuming these two values are equal would not be appropriate. Therefore it is unreasonable to assume that the random intercept and slope have the same variance, and therefore covariance structures which assume equal variance parameters along the diagonal are poor choices.

These problems of parameter estimates that are out of their bounds, non-positive definite random effects covariance matrices, and nonconvergence are due to the marginal modelling approach used to obtain the parameter estimates (Verbeke &

Molenberghs, 2000). The parameter estimation methods are designed to ensure positive definiteness of the covariance structure of $\mathbf{y}_i$ (Lindstrom & Bates, 1988; Wolfinger, 1993; Pourahmadi, 2000), but this does not imply that the individual covariances of the hierarchical model, $\mathbf{\Sigma}$ and $\boldsymbol{\omega}_i$, will be positive definite. These problems can in some cases be avoided by specifying better starting values for the parameters or by specifying a different fitting procedure (Verbeke & Molenberghs, 2000).