

# Examples from Multilevel Software Comparative Reviews

Douglas Bates  
R Development Core Team  
`Douglas.Bates@R-project.org`

February 2005, with updates up to February 15, 2011

## **Abstract**

The Center for Multilevel Modelling at the Institute of Education, London maintains a web site of “Software reviews of multilevel modeling packages”. The data sets discussed in the reviews are available at this web site. We have incorporated these data sets in the `mlmRev` package for R and, in this vignette, provide the results of fitting several models to these data sets.

## **1 Introduction**

R is an Open Source implementation of John Chambers’ S language language for data analysis and graphics. R was initially developed by Ross Ihaka and Robert Gentleman of the University of Auckland and now is developed and maintained by an international group of statistical computing experts.

In addition to being Open Source software, which means that anyone can examine the source code to see exactly how the computations are being carried out, R is freely available from a network of archive sites on the Internet. There are precompiled versions for installation on the Windows operating system, Mac OS X and several distributions of the Linux operating system. Because the source code is available those interested in doing so can compile their own version if they wish.

R provides an environment for interactive computing with data and for graphical display of data. Users and developers can extend the capabilities of

R by writing their own functions in the language and by creating packages of functions and data sets. Many such packages are available on the archive network called CRAN (Comprehensive R Archive Network) for which the parent site is <http://cran.r-project.org>. One such package called `lme4` (along with a companion package called `Matrix`) provides functions to fit and display linear mixed models and generalized linear mixed models, which are the statisticians' names for the models called multilevel models or hierarchical linear models in other disciplines. The `lattice` package provides functions to generate several high level graphics plots that help with the visualization of the types of data to which such models are fit. Finally, the `mlmRev` package provides the data sets used in the "Software Reviews of Multilevel Modeling Packages" from the Multilevel Modeling Group at the Institute of Education in the UK. This package also contains several other data sets from the multilevel modeling literature.

The software reviews mentioned above were intended to provide comparison of the speed and accuracy of many different packages for fitting multilevel models. As such, there is a standard set of models that fit to each of the data sets in each of the packages that were capable of doing the fit. We will fit these models for comparative purposes but we will also do some graphical exploration of the data and, in some cases, discuss alternative models.

We follow the general outline of the previous reviews, beginning with simpler structures and moving to the more complex structures. Because the previous reviews were performed on older and considerably slower computers than the ones on which this vignette will be compiled, the timings produced by the `system.time` function and shown in the text should not be compared with previous timings given on the web site. They are an indication of the times required to fit such models to these data sets on recent computers with processors running at around 2 GHz or faster.

## 2 Two-level normal models

In the multilevel modeling literature a two-level model is one with two levels of random variation; the per-observation noise term and random effects which are grouped according to the levels of a factor. We call this factor a *grouping factor*. If the response is measured on a continuous scale (more or less) our initial models are based on a normal distribution for the per-observation noise and for the random effects. Thus such a model is called a "two-level normal

model” even though it has only one grouping factor for the random effects.

## 2.1 The Exam data

The data set called `Exam` provides the normalized exam scores attained by 4,059 students from 65 schools in inner London. Some of the covariates available with this exam score are the school the student attended, the sex of the student, the school gender (boys, girls, or mixed) and the student’s result on the Standardised London Reading test.

The R functions `str` and `summary` can be used to examine the structure of a data set (or, in general, any R object) and to provide a summary of an object.

```
> str(Exam)
'data.frame':      4059 obs. of  10 variables:
 $ school  : Factor w/ 65 levels "1","2","3","4",...: 1 1 1 1 1 1 1 1 1 1 ...
 $ normexam: num  0.261 0.134 -1.724 0.968 0.544 ...
 $ schgend : Factor w/ 3 levels "mixed","boys",...: 1 1 1 1 1 1 1 1 1 1 ...
 $ schavg  : num  0.166 0.166 0.166 0.166 0.166 ...
 $ vr      : Factor w/ 3 levels "bottom 25%","mid 50%",...: 2 2 2 2 2 2 2 2 2 2 ...
 $ intake  : Factor w/ 3 levels "bottom 25%","mid 50%",...: 1 2 3 2 2 1 3 2 2 3 ...
 $ standLRT: num  0.619 0.206 -1.365 0.206 0.371 ...
 $ sex     : Factor w/ 2 levels "F","M": 1 1 2 1 1 2 2 1 2 ...
 $ type    : Factor w/ 2 levels "Mxd","Sngl": 1 1 1 1 1 1 1 1 1 ...
 $ student : Factor w/ 650 levels "1","2","3","4",...: 143 145 142 141 138 155 158 115 117 113 ...

> summary(Exam)

      school      normexam      schgend      schavg
14      : 198   Min.      :-3.6660720   mixed:2169   Min.      :-0.755961
17      : 126   1st Qu.: -0.6995050   boys : 513   1st Qu.: -0.149341
18      : 120   Median :  0.0043222   girls:1377   Median : -0.020198
49      : 113   Mean    :-0.0001138                Mean    :  0.001810
8       : 102   3rd Qu.:  0.6787592                3rd Qu.:  0.210525
15      : 91    Max.     :  3.6660912                Max.     :  0.637656
(Other):3309

      vr      intake      standLRT      sex      type
bottom 25%: 640   bottom 25%:1176   Min.      :-2.934953   F:2436   Mxd :2169
mid 50%   :2263   mid 50%   :2344   1st Qu.: -0.620713   M:1623   Sngl:1890
top 25%   :1156   top 25%   : 539   Median :  0.040499
                                Mean    :  0.001810
                                3rd Qu.:  0.619059
                                Max.     :  3.015952

      student
20      : 34
54      : 34
15      : 33
22      : 33
31      : 33
59      : 33
(Other):3859
```

## 2.2 Model fits and timings

The first model to fit to the Exam data incorporates fixed-effects terms for the pretest score, the student's sex and the school gender. The only random-effects term is an additive shift associated with the school.

```
> (Em1 <- lmer(normexam ~ standLRT + sex + schgend + (1|school), Exam))

Linear mixed model fit by REML
Formula: normexam ~ standLRT + sex + schgend + (1 | school)
Data: Exam
   AIC   BIC logLik deviance REMLdev
9362 9406  -4674    9326    9348
Random effects:
Groups   Name              Variance Std.Dev.
school   (Intercept)  0.085829  0.29297
Residual                    0.562534  0.75002
Number of obs: 4059, groups: school, 65

Fixed effects:
              Estimate Std. Error t value
(Intercept)  -0.001049   0.055564  -0.02
standLRT      0.559755   0.012450   44.96
sexM          -0.167391   0.034100   -4.91
schgendboys   0.177690   0.113464    1.57
schgendgirls  0.158997   0.089394    1.78

Correlation of Fixed Effects:
          (Intr) stnLRT sexM   schgndb
standLRT  -0.014
sexM      -0.316  0.061
schgendboys -0.395 -0.003 -0.145
schgendgrls -0.622  0.009  0.197  0.245
```

The `system.time` function can be used to time the execution of an R expression. It returns a vector of five numbers giving the user time (time spend executing applications code), the system time (time spent executing system functions called by the applications code), the elapsed time, and the user and system time for any child processes. The first number is what is commonly viewed as the time required to do the model fit. (The elapsed time is unsuitable because it can be affected by other processes running on the computer.) These times are in seconds. On modern computers this fit takes only a fraction of a second.

```
> system.time(lmer(normexam ~ standLRT + sex + schgend + (1|school), Exam))
   user  system elapsed 
0.330   0.000   0.325
```

## 2.3 Interpreting the fit

As can be seen from the output, the default method of fitting a linear mixed model is restricted maximum likelihood (REML). The estimates of the variance components correspond to those reported by other packages as given on the Multilevel Modelling Group's web site. Note that the estimates of the variance components are given on the scale of the variance and on the scale of the standard deviation. That is, the values in the column headed **Std.Dev.** are simply the square roots of the corresponding entry in the **Variance** column. They are **not** standard errors of the estimate of the variance.

The estimates of the fixed-effects are different from those quoted on the web site because the terms for **sex** and **schgend** use a different parameterization than in the reviews. Here the reference level of **sex** is female (F) and the coefficient labelled **sexM** represents the difference for males compared to females. Similarly the reference level of **schgend** is **mixed** and the two coefficients represent the change from mixed to boys only and the change from mixed to girls only. The value of the coefficient labelled **Intercept** is affected by both these changes as is the value of the REML criterion.

To reproduce the results obtained from other packages, we must change the reference level for each of these factors.

```
> Exam$sex      <- relevel(Exam$sex, "M")
> Exam$schgend <- relevel(Exam$schgend, "girls")
> (Em2 <- lmer(normexam ~ standLRT + sex + schgend + (1|school), Exam))
Linear mixed model fit by REML
Formula: normexam ~ standLRT + sex + schgend + (1 | school)
Data: Exam
   AIC   BIC logLik deviance REMLdev
9362 9406 -4674    9326    9348
Random effects:
Groups   Name              Variance Std.Dev.
school  (Intercept)  0.085829  0.29297
Residual                    0.562534  0.75002
Number of obs: 4059, groups: school, 65

Fixed effects:
              Estimate Std. Error t value
(Intercept) -0.009443   0.077890  -0.12
standLRT     0.559755   0.012450  44.96
sexF         0.167391   0.034100   4.91
schgendmixed -0.158997   0.089394  -1.78
schgendboys  0.018694   0.126047   0.15

Correlation of Fixed Effects:
              (Intr) stnLRT sexF   schgndm
standLRT      0.027
sexF          -0.438 -0.061
schgendmixd -0.790 -0.009  0.197
schgendboys -0.618 -0.009  0.270  0.488
```

The coefficients now correspond to those in the tables on the web site. It happens that the REML criterion at the optimum in this fit is the same as in the previous fit, but you cannot depend on this occurring. In general the value of the REML criterion at the optimum depends on the parameterization used for the fixed effects.

## 2.4 Further exploration

### 2.4.1 Checking consistency of the data

It is important to check the consistency of data before trying to fit sophisticated models. One should plot the data in many different ways to see if it looks reasonable and also check relationships between variables.

For example, each observation in these data is associated with a particular student. The variable `student` is not a unique identifier of the student as it only has 650 unique values. It is intended to be a unique identifier within a school but it is not. To show this we create a factor that is the interaction of school and student then drop unused levels.

```
> Exam <- within(Exam, ids <- factor(school:student))
> str(Exam)
'data.frame':      4059 obs. of  11 variables:
 $ school  : Factor w/ 65 levels "1","2","3","4",...: 1 1 1 1 1 1 1 1 1 1 ...
 $ normexam: num  0.261 0.134 -1.724 0.968 0.544 ...
 $ schgend : Factor w/ 3 levels "girls","mixed",...: 2 2 2 2 2 2 2 2 2 2 ...
 $ schavg  : num  0.166 0.166 0.166 0.166 0.166 ...
 $ vr      : Factor w/ 3 levels "bottom 25%","mid 50%",...: 2 2 2 2 2 2 2 2 2 2 ...
 $ intake  : Factor w/ 3 levels "bottom 25%","mid 50%",...: 1 2 3 2 2 1 3 2 2 3 ...
 $ standLRT: num  0.619 0.206 -1.365 0.206 0.371 ...
 $ sex     : Factor w/ 2 levels "M","F": 2 2 1 2 2 1 1 1 2 1 ...
 $ type    : Factor w/ 2 levels "Mxd","Sngl": 1 1 1 1 1 1 1 1 1 1 ...
 $ student : Factor w/ 650 levels "1","2","3","4",...: 143 145 142 141 138 155 158 115 117 113 ...
 $ ids     : Factor w/ 4055 levels "1:1","1:4","1:6",...: 48 49 47 46 45 50 51 39 40 38 ...
```

Notice that there are 4059 observations but only 4055 unique levels of student within school. We can check the ones that are duplicated

```
> as.character(Exam$ids[which(duplicated(Exam$ids))])
[1] "43:86" "50:39" "52:2"  "52:21"
```

One of these cases

```
> subset(Exam, ids == '43:86')
  school normexam schgend  schavg    vr intake  standLRT sex type
2758   43 -0.8526700  mixed 0.4334322 top 25% mid 50%  0.1231502  M  Mxd
2759   43  0.8219882  mixed 0.4334322 top 25% top 25% -0.0421520  F  Mxd
  student  ids
2758     86 43:86
2759     86 43:86
```

```
> xtabs(~ sex + school, Exam, subset = school %in% c(43, 50, 52), drop = TRUE)

      school
sex 43 50 52
M   1 35 61
F  60 38  0
```

is particularly interesting. Notice that one of the students numbered 86 in school 43 is the only male student out of 61 students from this school who took the exam. It is quite likely that this student's score was attributed to the wrong school and that the school is in fact a girls-only school, not a mixed-sex school.

The causes of the other three cases of duplicate student numbers within a school are not as clear. It would be necessary to go back to the original data records to check these.

The cross-tabulation of the students by sex and school for the mixed-sex schools

```
> xtabs(~ sex + school, Exam, subset = type == "Mxd", drop = TRUE)

      school
sex  1  3  4  5  9 10 12 13 14 15 17 19 20 22 23 26 28 32 33
M  45 29 45 16 21 31 23 26 92 47 31 33 21 48 10 44 31 27 44
F  28 23 34 19 13 19 24 38 106 44 95 22 18 42 18 31 26 15 33

      school
sex 34 38 42 43 45 46 47 50 51 54 55 56 59 61 62 63
M  18 31 35  1  5 47 81 35 26  4 26 16 30 35 43 13
F   8 23 23 60 48 36  1 38 32  4 25 22 17 29 28 17
```

shows another anomaly. School 47 is similar to school 43 in that, although it is classified as a mixed-sex school, 81 male students and only one female student took the exam. It is likely that the school was misrecorded for this one female student and the school is a male-only school.

Another school is worth noting. There were only eight students from school 54 who took the exam so any within-school estimates from this school will be unreliable.

A mosaic plot (Figure~1) produced with

```
> ExamMxd <- within(subset(Exam, type == "Mxd"), school <- factor(school))
> mosaicplot(~ school + sex, ExamMxd)
```

helps to detect mixed-sex schools with unusually large or unusually small ratios of females to males taking the exam.

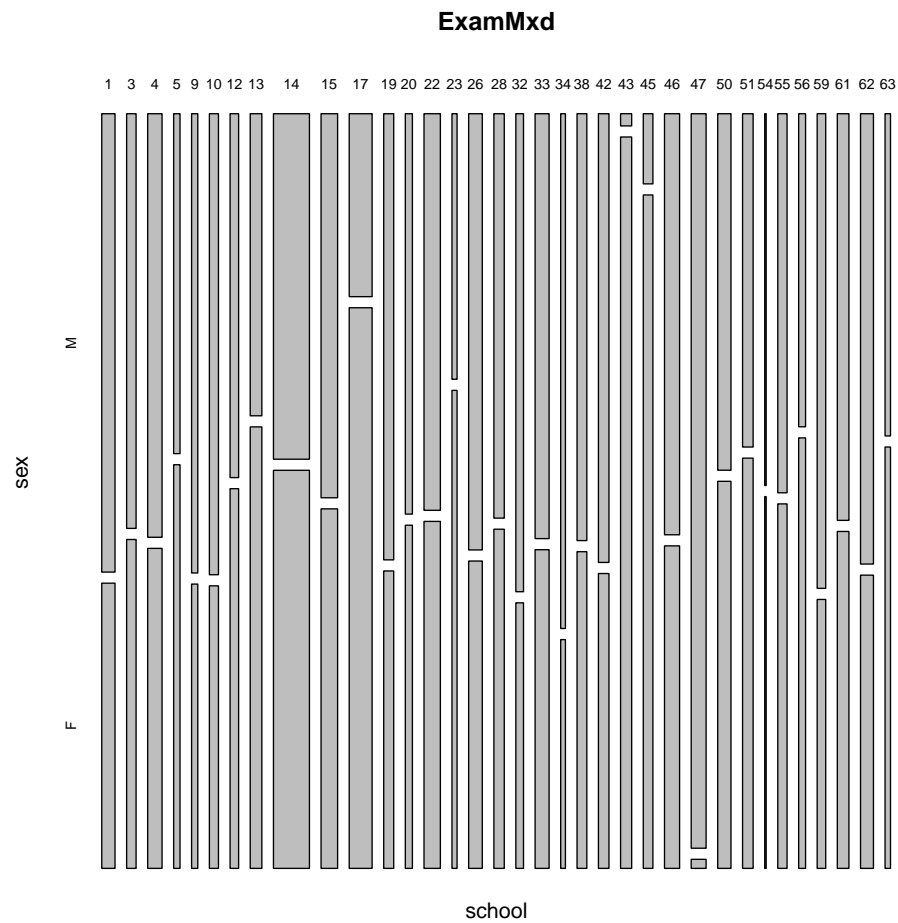


Figure 1: A mosaic plot of the sex distribution by school. The areas of the rectangles are proportional to the number of students of that sex from that school who took the exam. Schools with an unusually large or unusually small ratio of females to males are highlighted.



### 2.4.2 Preliminary graphical displays

In addition to the pretest score (`standLRT`), the predictor variables used in this model are the student's sex and the school gender, which is coded as having three levels. There is some redundancy in these two variables in that all the students in a boys-only school must be male. For graphical exploration we convert from `schgend` to `type`, an indicator of whether the school is a mixed-sex school or a single-sex school, and plot the response versus the pretest score for each combination of sex and school type.

This plot is created with the `xyplot` from the `lattice` package as (essentially)

```
> xyplot(normexam ~ standLRT | sex * type, Exam, type = c("g", "p", "smooth"))
```

The formula would be read as “plot `normexam` by `standLRT` given `sex` by (school) `type`”. A few other arguments were added in the actual call to make the axis annotations more readable.

Figure~2 shows the even after accounting for a student's sex, pretest score and school type, there is considerable variation in the response. We may attribute some of this variation to differences in schools but the fitted model indicates that most of the variation is unaccounted or “residual” variation.

In some ways the high level of residual variation obscures the pattern in the data. By removing the data points and overlaying the scatterplot smoothers we can concentrate on the relationships between the covariates. The call to `xyplot` is essentially

```
> xyplot(normexam ~ standLRT, Exam, groups = sex:type, type = c("g", "smooth"))
```

Figure~3 is a remarkable plot in that it shows nearly a perfect “main effects” relationship of the response with the three covariates and almost no interaction. It is rare to see real data follow a simple theoretical relationship so closely.

To elaborate, we can see that for each of the four groups the smoothed relationship between the exam score and the pretest score is close to a straight line and that the lines are close to being parallel. The only substantial deviation is in the smoothed relationship for the males in single-sex schools and this is the group with the fewest observations and hence the least precision in the estimated relationship. The lack of parallelism for this group is most apparent in the region of extremely low pretest scores where there are few observations and a single student who had a low pretest score and a moderate post-test score can substantially influence the curve. Five or six such points can be seen in the upper left panel of Figure~2.

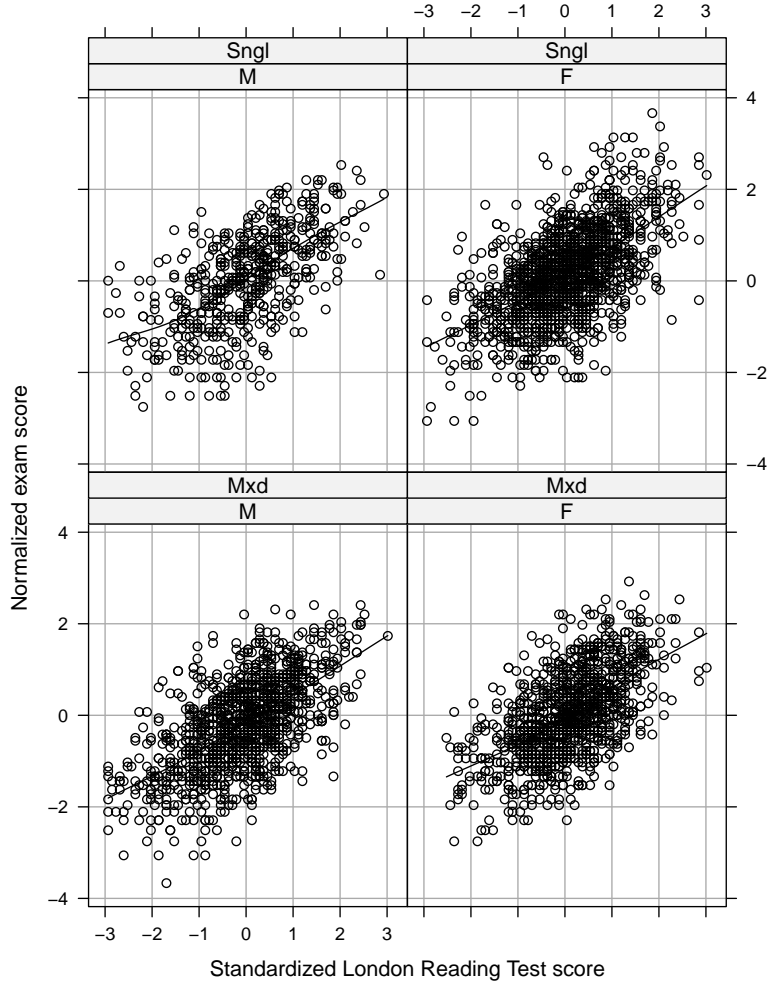


Figure 2: Normalized exam score versus pretest (Standardized London Reading Test) score for 4095 students from 65 schools in inner London. The panels on the left show the male students' scores; those on the right show the females' scores. The top row of panels shows the scores of students in single-sex schools and the bottom row shows the scores of students in mixed-sex schools. A scatterplot smoother line for each panel has been added to help visualize the trend.

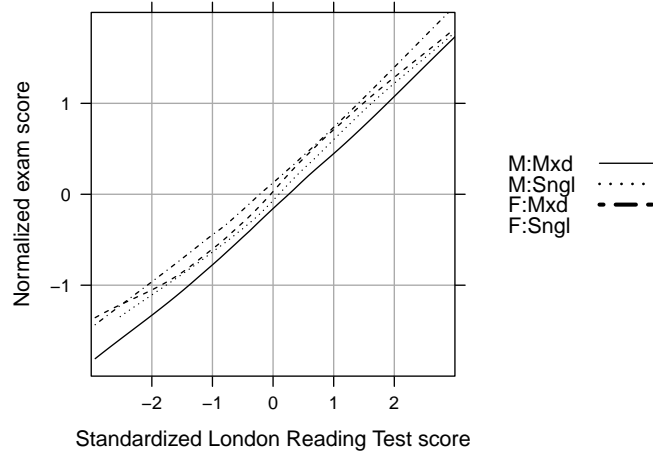


Figure 3: Overlaid scatterplot smoother lines of the normalized test scores versus the pretest (Standardized London Reading Test) score for female (F) and male (M) students in single-sex (Sngl) and mixed-sex (Mxd) schools.

We should also notice the ordering of the lines and the spacing between the lines. The smoothed relationships for students in single-sex schools are consistently above those in the mixed-sex schools and, except for the region of low pretest scores described above, the relationship for the females in a given type of school is consistently above that for the males. Furthermore the distance between the female and male lines in the single-sex schools is approximately the same as the corresponding distance in the mixed-sex schools. We would summarize this by saying that there is a positive effect for females versus males and a positive effect for single-sex versus mixed-sex and no indication of interaction between these factors.

### 2.4.3 The effect of schools

We can check for patterns within and between schools by plotting the response versus the pretest by school. Because there appear to be differences in this relationship for single-sex versus mixed-sex schools and for females versus males we consider these separately.

In Figure~4 we plot the normalized exam scores versus the pretest score by school for female students in single-sex schools. The plot is produced as

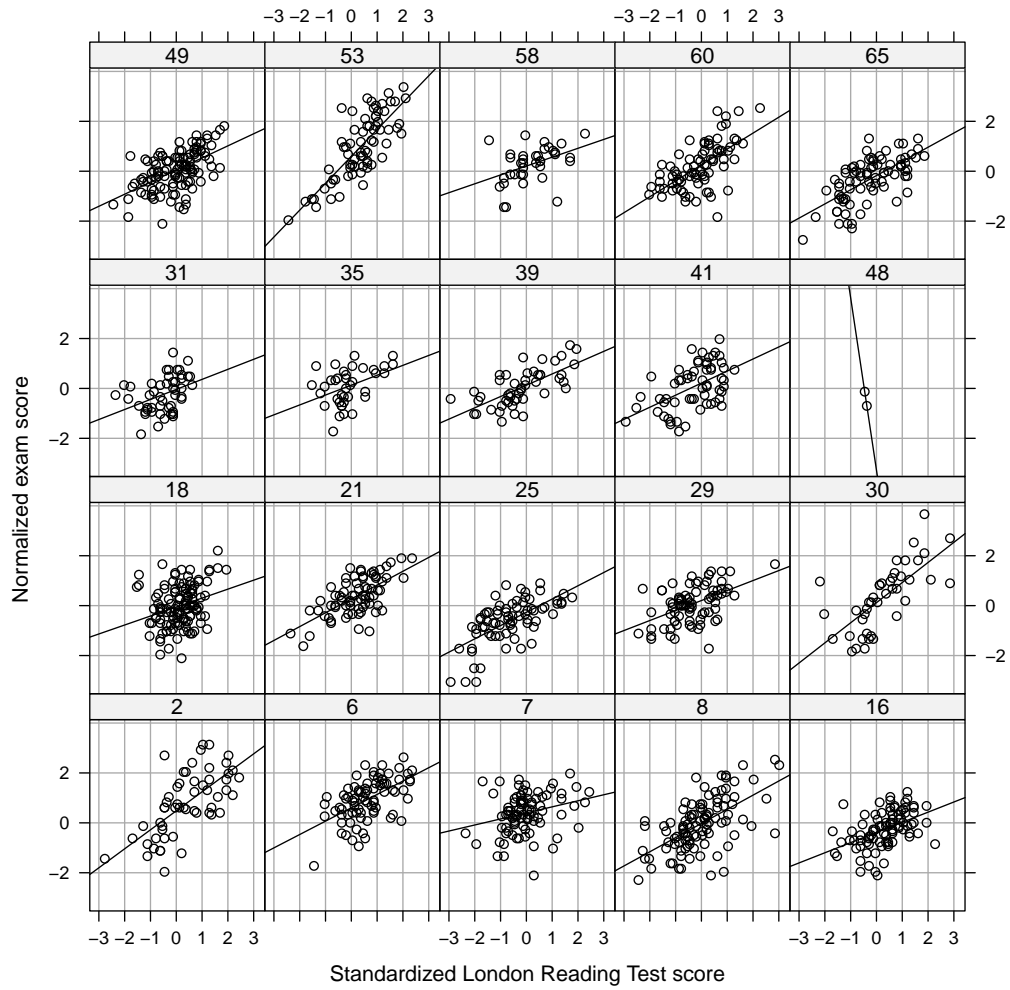


Figure 4: Normalized exam scores versus pretest (Standardized London Reading Test) score by school for female students in single-sex schools.

```
> xyplot(normexam ~ standLRT | school, Exam,
+         type = c("g", "p", "r"),
+         subset = sex == "F" & type == "Sngl")
```

The "r" in the `type` argument adds a simple linear regression line to each panel.

The first thing we notice in Figure~4 is that school 48 is an anomaly because only two students in this school took the exam. Because within-school results based on only two students are unreliable, we will exclude this school from further plots (but we do include these data when fitting comprehensive models).

Although the regression lines on the panels can help us to look for variation in the schools, the ordering of the panels is, for our purposes, random. We recreate this plot in Figure~5 using

```
> xyplot(normexam ~ standLRT | school, Exam, type = c("g", "p", "r"),
+         subset = sex == "F" & type == "Sngl" & school != 48,
+         index.cond = function(x, y) coef(lm(y ~ x))[1])
```

so that the panels are ordered (from left to right starting at the bottom row) by increasing intercept for the regression line (i.e. by increasing predicted exam score for a student with a pretest score of 0).

Alternatively, we could order the panels by increasing slope of the within-school regression lines, as in Figure~6.

Although it is informative to plot the within-school regression lines we need to assess the variability in the estimates of the coefficients before concluding if there is “significant” variability between schools. We can obtain the individual regression fits with the `lmList` function

```
> show(ExamFS <- lmList(normexam ~ standLRT | school, Exam,
+                        subset = sex == "F" & type == "Sngl" & school != 48))
```

```
Call: lmList(formula = normexam ~ standLRT | school, data = Exam, subset = sex ==
```

```
"F" & type == "Sngl")
```

```
Coefficients:
      (Intercept)  standLRT
2    0.48227991  0.7612884
6    0.60321439  0.5353444
7    0.39852689  0.2422785
8   -0.02519463  0.5674053
16  -0.38564292  0.4069399
18  -0.05733995  0.3593830
21   0.26872018  0.5544939
25  -0.26779146  0.5320575
29   0.20442314  0.4005158
30   0.11885028  0.8059021
31  -0.03922548  0.4022838
35   0.13173022  0.3966535
39   0.12754208  0.4525918
41   0.21249712  0.4834107
```

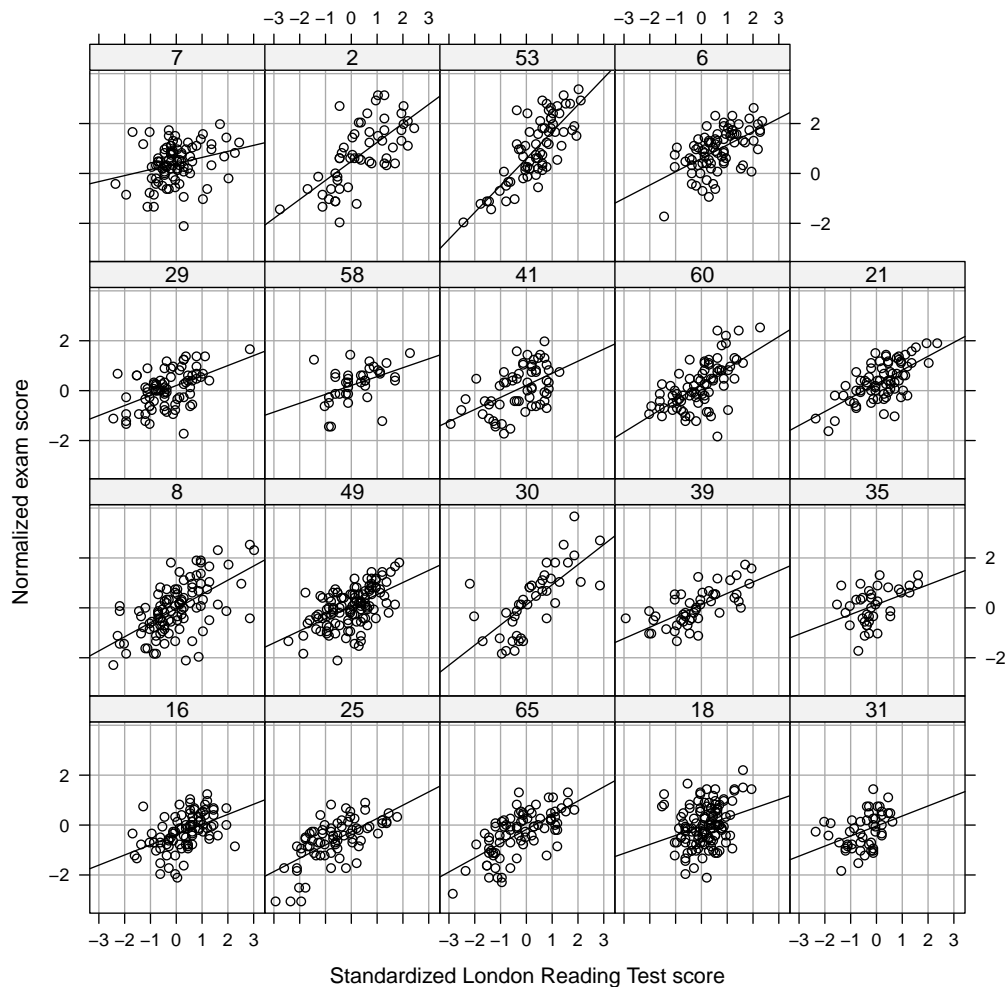


Figure 5: Normalized exam scores versus pretest (Standardized London Reading Test) score by school for female students in single-sex schools. School 48 where only two students took the exam has been eliminated and the panels have been ordered by increasing intercept (predicted normalized score for a pretest score of 0) of the regression line.

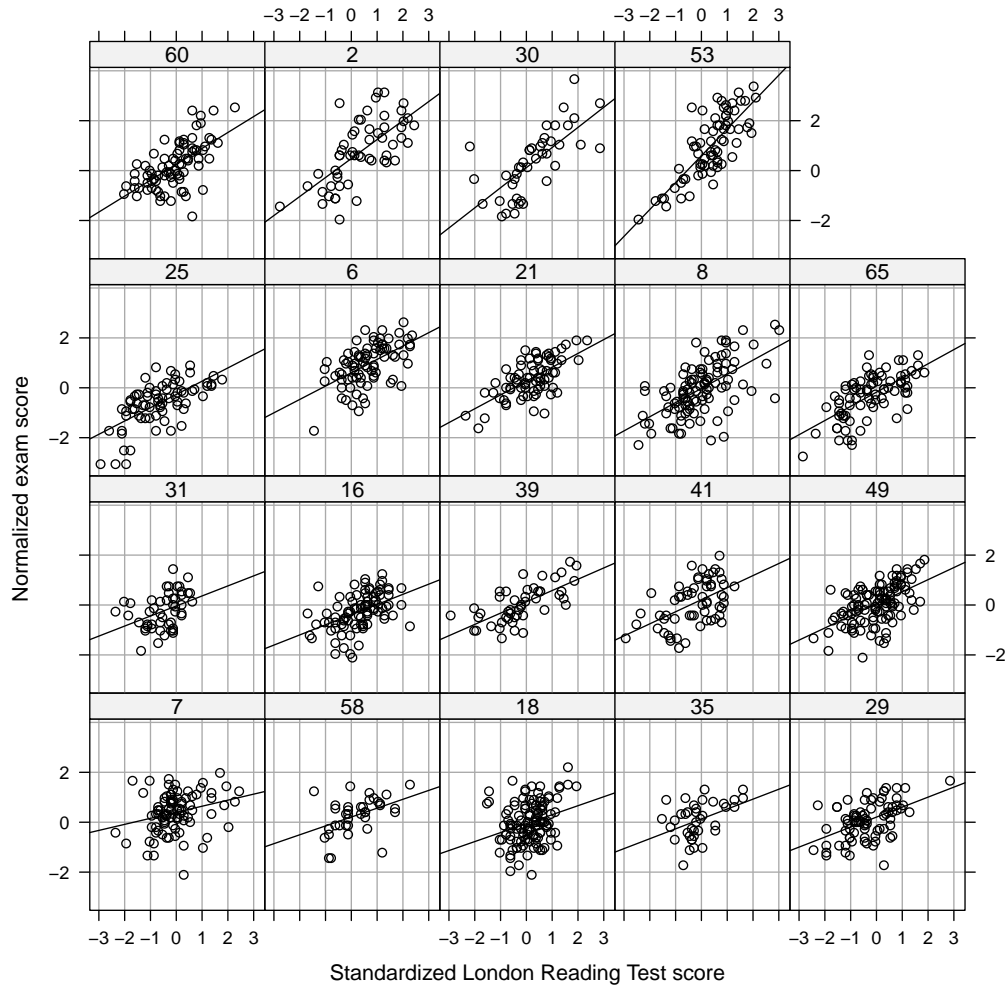


Figure 6: Normalized exam scores versus pretest (Standardized London Reading Test) score by school for female students in single-sex schools. School 48 has been eliminated and the panels have been ordered by increasing slope of the within-school regression lines.

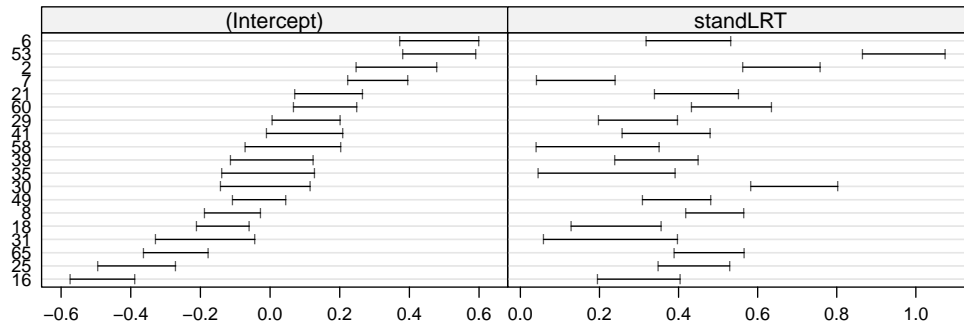


Figure 7: Confidence intervals on the coefficients of the within-school regression lines for female students in single-sex schools. School 48 has been eliminated and the schools have been ordered by increasing estimated intercept.

```
49  0.04747055  0.4845568
53  0.59370349  1.0769781
58  0.20707724  0.3557839
60  0.25196603  0.6378090
65 -0.17490019  0.5684592
```

```
Degrees of freedom: 1375 total; 1337 residual
Residual standard error: 0.7329521
```

and compare the confidence intervals on these coefficients.

```
> plot(confint(ExamFS, pool = TRUE), order = 1)
```

```
> show(ExamMS <- lmList(normexam ~ standLRT | school, Exam,
+                          subset = sex == "M" & type == "Sngl"))
```

```
Call: lmList(formula = normexam ~ standLRT | school, data = Exam, subset = sex ==
```

```
"M" & type == "Sngl")
```

```
Coefficients:
```

```
(Intercept)  standLRT
11  0.26596312  0.4586355
24  0.17773174  0.3976156
27  0.03518861  0.5728684
36 -0.20691842  0.4383453
37 -0.48522245  0.2382739
40 -0.25019842  0.7262845
44 -0.34440523  0.3696269
52  0.38803903  0.7402692
57  0.04055154  0.6123692
64  0.03749455  0.7055239
```

```
Degrees of freedom: 513 total; 493 residual
Residual standard error: 0.8082068
```



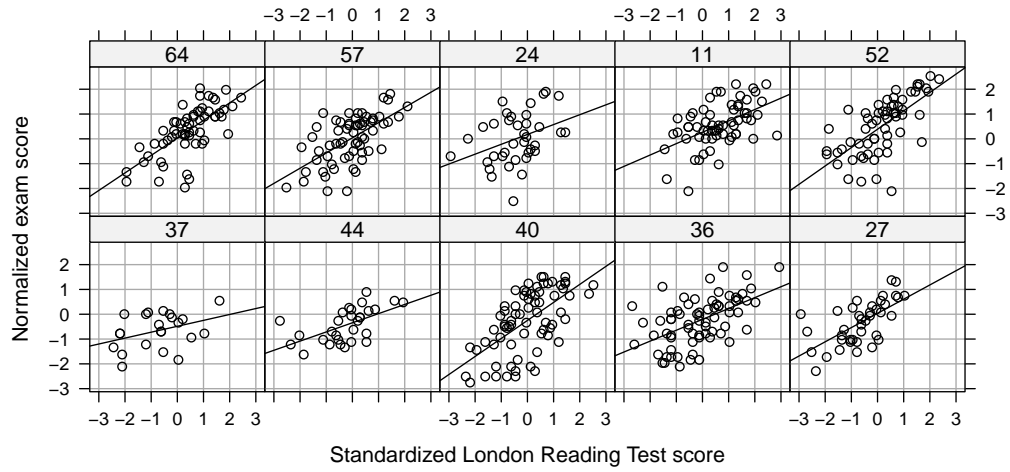


Figure 8: Normalized exam scores versus pretest (Standardized London Reading Test) score by school for male students in single-sex schools.

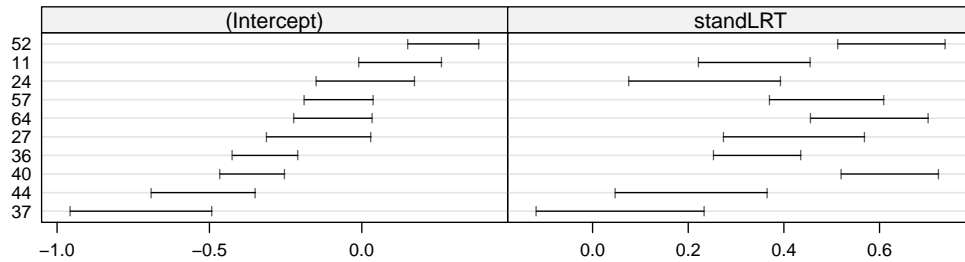


Figure 9: Confidence intervals on the coefficients of the within-school regression lines for female students in single-sex schools. School 48 has been eliminated and the schools have been ordered by increasing estimated intercept.

The corresponding plot of the confidence intervals is shown in Figure~9.

For the mixed-sex schools we can consider the effect of the pretest score and sex in the plot (Figure~10) and in the separate model fits for each school.

```
> show(ExamM <- lmList(normexam ~ standLRT + sex | school, Exam,
+                        subset = type == "Mxd" & !school %in% c(43,47,54)))

Call: lmList(formula = normexam ~ standLRT + sex | school, data = Exam, subset = type == "Mxd" & !school %in% c(43,47,54))

Coefficients:
      (Intercept)  standLRT      sexF
1    0.24770238  0.7044798  0.355719842
3    0.58030950  0.5843480 -0.057223931
4   -0.16739321  0.7372405  0.402829332
5    0.36213174  0.6695127 -0.183238345
9   -0.32487665  0.3961812  0.222382659
10  -0.45139239  0.2972074  0.330917735
12  -0.26201220  0.4200265  0.378359201
13  -0.27196976  0.6037392  0.196013604
14  -0.28741229  0.5966633  0.202122649
15  -0.30963145  0.7370363  0.144392527
17  -0.30553035  0.4905235  0.156646395
19  -0.22542808  0.7214611  0.385544198
20   0.25120209  0.5187894  0.041342515
22  -0.50744197  0.5206371  0.089368918
23  -0.51727825  0.3873051 -0.189332381
26  -0.11870804  0.5342909  0.185283146
28  -0.84451962  0.2583861  0.138672637
32   0.02596084  0.6560569  0.082029123
33  -0.02539396  0.5148927  0.147967544
34  -0.19582273  0.7662681  0.327656212
38  -0.19255275  0.6184554  0.084872081
42  -0.01913469  0.3827088  0.246533297
45  -0.21212351  0.5665400  0.102317128
46  -0.30554555  0.4491187 -0.201294993
50  -0.32718434  0.6752947  0.001906973
51  -0.40150931  0.3076539  0.445100548
55   0.35743002  0.6118447  0.400034477
56  -0.18744293  0.8558783  0.391178135
59  -0.97233088  0.3594417  0.329480168
61  -0.01215078  0.6428683 -0.060024560
62  -0.16445110  0.5411566  0.283642476
63   0.60216184  0.3091657  0.150390337

Degrees of freedom: 2018 total; 1922 residual
Residual standard error: 0.7273955
```

The confidence intervals for these separately fitted models, shown in Figure~11 indicate differences in the intercepts and possibly differences in the slopes with respect to the pretest scores. However, there is not a strong indication of variation by school in the effect of sex.

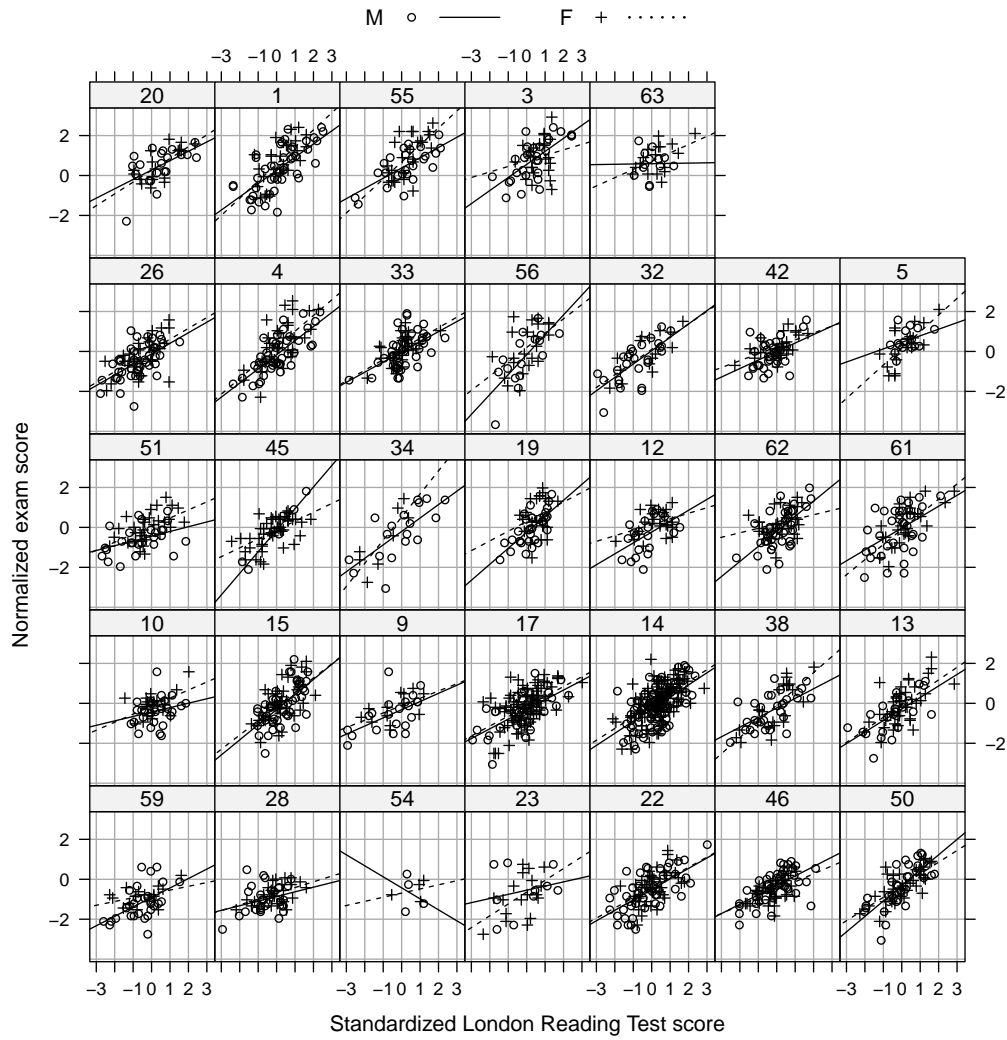


Figure 10: Normalized exam scores versus pretest score by school and sex for students in mixed-sex schools.

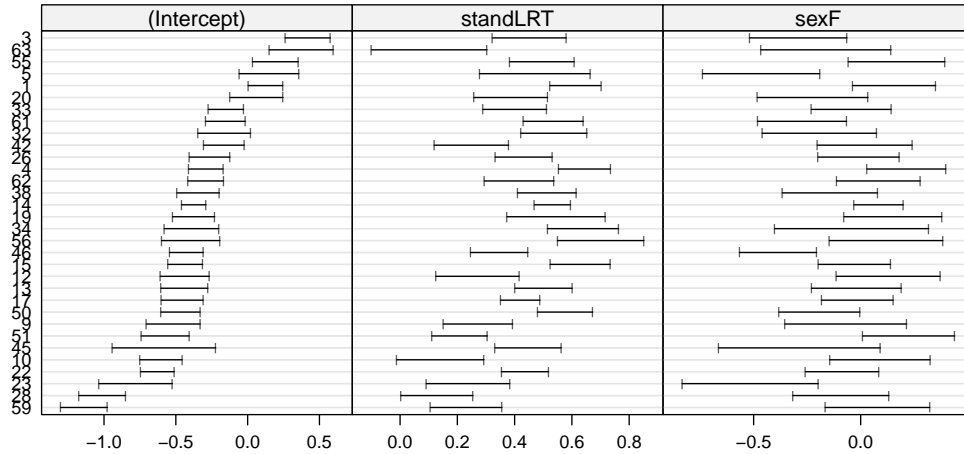


Figure 11: Confidence intervals on the coefficients of the within-school regression lines for female students in single-sex schools. School 48 has been eliminated and the schools have been ordered by increasing estimated intercept.

## 2.5 Multilevel models for the exam data

We begin with a model that has a random effects for the intercept by school plus additive fixed effects for the pretest score, the student's sex and the school type.

```
> (Em3 <- lmer(normexam ~ standLRT + sex + type + (1|school), Exam))
```

Linear mixed model fit by REML

Formula: normexam ~ standLRT + sex + type + (1 | school)

Data: Exam

AIC BIC logLik deviance REMLdev

9357 9395 -4673 9325 9345

Random effects:

Groups	Name	Variance	Std.Dev.
school	(Intercept)	0.084367	0.29046
Residual		0.562529	0.75002

Number of obs: 4059, groups: school, 65

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	-0.16768	0.05468	-3.07
standLRT	0.55984	0.01245	44.97
sexF	0.16596	0.03281	5.06
typeSngl	0.16546	0.07742	2.14

Correlation of Fixed Effects:

```

              (Intr) stnLRT sexF
standLRT    0.023
sexF        -0.291 -0.061
typeSngl    -0.624  0.005 -0.078

```

Our data exploration indicated that the slope with respect to the pretest score may vary by school. We can fit a model with random effects by school for both the slope and the intercept as

```

> (Em4 <- lmer(normexam ~ standLRT + sex + type + (standLRT|school), Exam))
Linear mixed model fit by REML
Formula: normexam ~ standLRT + sex + type + (standLRT | school)
Data: Exam
   AIC   BIC logLik deviance REMLdev
9317 9367 -4650   9281    9301
Random effects:
Groups      Name      Variance Std.Dev. Corr
school      (Intercept) 0.082477 0.28719
              standLRT  0.015081 0.12280  0.579
Residual                    0.550289 0.74181
Number of obs: 4059, groups: school, 65

Fixed effects:
              Estimate Std. Error t value
(Intercept) -0.18869    0.05194  -3.633
standLRT      0.55410    0.02012  27.546
sexF          0.16797    0.03228   5.203
typeSngl      0.17639    0.06959   2.535

Correlation of Fixed Effects:
              (Intr) stnLRT sexF
standLRT     0.317
sexF         -0.292 -0.037
typeSngl     -0.580  0.017 -0.101

```

and compare this fit to the previous fit with

```

> anova(Em3, Em4)
Data: Exam
Models:
Em3: normexam ~ standLRT + sex + type + (1 | school)
Em4: normexam ~ standLRT + sex + type + (standLRT | school)
   Df    AIC    BIC logLik Chisq Chi Df Pr(>Chisq)
Em3  6 9337.5 9375.3 -4662.7
Em4  8 9297.2 9347.6 -4640.6 44.316    2 2.382e-10

```

There is a strong evidence of a significant random effect for the slope by school, whether judged by AIC, BIC or the p-value for the likelihood ratio test.

The p-value for the likelihood ratio test is based on a  $\chi^2$  distribution with degrees of freedom calculated as the difference in the number of parameters in the two models. Because one of the parameters eliminated from the full

model in the submodel is at its boundary the usual asymptotics for the likelihood ratio test do not apply. However, it can be shown that the p-value quoted for the test is conservative in the sense that it is an upper bound on the p-value that would be calculated say from a parametric bootstrap.

Having an upper bound of  $1.9 \times 10^{-10}$  on the p-value can be regarded as “highly significant” evidence of the utility of the random effect for the slope by school.

We could also add a random effect for the student’s sex by school

```
> (Em5 <- lmer(normexam ~ standLRT + sex + type + (standLRT + sex|school), Exam))
Linear mixed model fit by REML
Formula: normexam ~ standLRT + sex + type + (standLRT + sex | school)
Data: Exam
   AIC   BIC logLik deviance REMLdev
9322 9391  -4650    9281    9300
Random effects:
Groups      Name          Variance Std.Dev. Corr
school      (Intercept)  0.07589442 0.275489
              standLRT    0.01508107 0.122805  0.616
              sexF         0.00084814 0.029123  0.696 -0.137
Residual                    0.55023004 0.741775
Number of obs: 4059, groups: school, 65

Fixed effects:
              Estimate Std. Error t value
(Intercept) -0.18952    0.05002   -3.789
standLRT      0.55407    0.02012   27.545
sexF          0.16983    0.03255    5.217
typeSngl      0.17617    0.06975    2.526

Correlation of Fixed Effects:
      (Intr) stnLRT sexF
standLRT  0.337
sexF     -0.202 -0.053
typeSngl -0.576  0.016 -0.161
```

Notice that the estimate of the variance of the `sexM` term is essentially zero so there is no need to test the significance of this variance component. We proceed with the analysis of `Em4`.

## 2.6 Markov Chain Monte Carlo methods for assessing parameter variation

An important part of the analysis of a statistical model is the assessment of the precision of parameter estimates, either through confidence intervals and confidence regions or through hypothesis tests on the parameters. Sometimes such information is condensed and given as the parameter estimate and a

Figure 12: Plots of the estimated posterior probability densities from a Markov Chain Monte Carlo sample of the parameters in the model `Em4`.

Figure 13: Normal probability plots of the Markov Chain Monte Carlo samples from the posterior densities of the parameters in the model `Em4`.

standard error of the estimate, with the assumption that a confidence interval will be formed as the estimate plus/minus some multiple of the standard error.

Summarizing the variability in parameter estimates with such symmetric intervals is appropriate if the distribution of the parameters can be assumed to be reasonably symmetric but this is not the case for variance components. In general we expect the distribution of a variance component to be approximately  $\chi^2$  and, depending upon the precision with which the component is estimated, such a distribution can be quite asymmetric.

To examine the distribution of the parameter estimates we can use a parametric bootstrap or adopt a Bayesian approach and generate a sample from the posterior distribution of the parameters using a technique called Markov Chain Monte Carlo sampling. The `mcmcsm` function applied to a fitted `lmer` model generates such a sample.

**> NA**

The parameter estimates in this sample are the fixed effects, the variance of the random per-observation noise,  $\sigma^2$ , the variance of the intercept random effect, the variance of the slope random effect and their covariance. A plot of the estimated posterior densities of each of the parameters (Figure~12) shows that the posterior densities of the fixed effects are indeed quite symmetric and close to a normal (or Gaussian) distribution. The posterior density of  $\sigma^2$  and the covariance parameter are reasonably symmetric but the posterior densities of the variances of the random effects are not.

A better way of assessing possible asymmetry is to use normal probability plots of the samples from each of the parameters, as in Figure~13.

We can see from Figure~13 that the posterior distributions of the fixed effects are remarkably close to the normal (or Gaussian) distribution. Except for a few points in each panel, these panels show almost perfect linearity. The few points that do deviate from the straight line in each case represent fewer than a dozen samples out of a total of 10,000 and are not a cause for concern.

Figure 14: Estimated density plots of the Markov Chain Monte Carlo samples from the posterior densities of the variance and covariance parameters in model **Em4**. The variance parameters are on the log scale and the covariance is expressed as Fisher’s z transformation of the correlation.

Figure 15: Normal probability plots of the Markov Chain Monte Carlo samples from the posterior densities of the variance and covariance parameters in model **Em4**. The variance parameters are on the log scale and the covariance is expressed as Fisher’s z transformation of the correlation.

The posterior distribution of  $\sigma^2$  is satisfactorily close to a normal distribution. As we will see, it is not always the case that this parameter is close to normally distributed. It happens that in this example  $\sigma^2$  is very precisely determined and possible asymmetry in the distribution will not be noticeable over such a small range of values.

There is noticeable asymmetry in the distribution of the variances of the random effects and perhaps some asymmetry in the distribution of the covariance of the intercept and slope random effects.

Using this sample we can examine the distribution of functions of these parameters. For example, (Box and Tiao, 1973) advocate approximating the distribution of the logarithm of a variance rather than the variance itself. For a covariance we first convert it to a correlation then apply Fisher’s z transformation (the hyperbolic arc-tangent) to the correlation. The density plots and normal probability plots of these transformed parameters are shown in Figures~14 and 15.

### 3 Three-level Normal Models

These results are from the 1997 A-level Chemistry exam. The **school** is nested in **lea** (local education authority) and has unique levels for each of the 2410 schools. It is a good practice to make the nesting explicit by specifying the grouping factors as the ‘outer’ factor, **lea** in this case, and the interaction of the outer and inner factors, **lea:school** or **school:lea** in this case. This will ensure unique levels for each **school** within **lea** combination.

To fit the model **mC2** we increase the number of EM iterations from its default of 20 to 40. Without this change the current version of the **optim**



function in R will declare convergence to an incorrect optimum. By increasing the number of EM iterations we are able to get closer to the optimum before calling `optim` and converge to the correct value. The `optim` function will be patched so this change will not be needed in future versions of R.

Data from the 1997 A-level Chemistry exam are available as `Chem97`.

```
> str(Chem97)
'data.frame':      31022 obs. of  8 variables:
 $ lea      : Factor w/ 131 levels "1","2","3","4",...: 1 1 1 1 1 1 1 1 1 ...
 $ school   : Factor w/ 2410 levels "1","2","3","4",...: 1 1 1 1 1 1 1 1 1 ...
 $ student  : Factor w/ 31022 levels "1","2","3","4",...: 1 2 3 4 5 6 7 8 9 10 ...
 $ score    : num  4 10 10 10 8 10 6 8 4 10 ...
 $ gender   : Factor w/ 2 levels "M","F": 2 2 2 2 2 2 2 2 2 ...
 $ age      : num  3 -3 -4 -2 -1 4 1 4 3 0 ...
 $ gcsecore: num  6.62 7.62 7.25 7.5 6.44 ...
 $ gcsecnt  : num  0.339 1.339 0.964 1.214 0.158 ...

> system.time(mC1 <- lmer(score ~ 1+(1|lea:school) + (1|lea), Chem97))
   user  system elapsed 
 3.060   0.080   3.141 

> summary(mC1)
Linear mixed model fit by REML
Formula: score ~ 1 + (1 | lea:school) + (1 | lea)
Data: Chem97
      AIC      BIC logLik deviance REMLdev
157882 157915 -78937   157870   157874
Random effects:
Groups      Name      Variance Std.Dev.
lea:school (Intercept) 2.74860   1.65789
lea        (Intercept) 0.15349   0.39177
Residual                    8.51611   2.91824
Number of obs: 31022, groups: lea:school, 2410; lea, 131

Fixed effects:
              Estimate Std. Error t value
(Intercept)   5.31900    0.05809   91.57

> system.time(mC2 <- lmer(score ~ gcsecnt + (1|school) + (1|lea), Chem97))
   user  system elapsed 
 1.450   0.010   1.455 

> summary(mC2)
Linear mixed model fit by REML
Formula: score ~ gcsecnt + (1 | school) + (1 | lea)
Data: Chem97
      AIC      BIC logLik deviance REMLdev
141707 141749 -70848   141686   141697
Random effects:
Groups      Name      Variance Std.Dev.
school      (Intercept) 1.166205 1.07991
lea         (Intercept) 0.014774 0.12155
Residual                    5.154199 2.27029
Number of obs: 31022, groups: school, 2410; lea, 131
```

```
Fixed effects:
              Estimate Std. Error t value
(Intercept)  5.63544    0.03123   180.5
gcsecnt      2.47256    0.01690   146.3

Correlation of Fixed Effects:
      (Intr)
gcsecnt 0.058
```

## 4 Two-level models for binary data

When the response variable is binary or when it represents a count we frequently model the data with a *generalized linear model* (glm) or, if we use random effects in the model, a *generalized linear mixed model* (glmm). Determining maximum likelihood estimates of the parameters in such a model is not as easy as for the linear mixed model because the likelihood for a glmm is expressed as an integral that must be approximated.

The `lmer` function has provision for fitting such models using one of three approximations. The simplest approximation is called *penalized quasi-likelihood* (PQL). This method is generally quite fast but also the least accurate of the three. The Laplace approximation is more accurate than PQL and the most accurate approximation is called *adaptive Gauss-Hermite quadrature* (AGQ). At present AGQ can only be used on models that have one-dimensional random effects associated with a single grouping factor.

### 4.1 The contraception use data

A fertility survey of women in Bangladesh included as a response their use of artificial contraception. Some of the covariates included the woman's age (on a centered scale), the number of live children she had, whether she lived in an urban or rural setting, and the district in which she lived.

```
> str(Contraception)

'data.frame':      1934 obs. of  6 variables:
 $ woman   : Factor w/ 1934 levels "1","2","3","4",...: 1 2 3 4 5 6 7 8 9 10 ...
 $ district: Factor w/ 60 levels "1","2","3","4",...: 1 1 1 1 1 1 1 1 1 1 ...
 $ use     : Factor w/ 2 levels "N","Y": 1 1 1 1 1 1 1 1 1 1 ...
 $ livch   : Factor w/ 4 levels "0","1","2","3+": 4 1 3 4 1 1 4 4 2 4 ...
 $ age     : num  18.44 -5.56 1.44 8.44 -13.56 ...
 $ urban   : Factor w/ 2 levels "N","Y": 2 2 2 2 2 2 2 2 2 2 ...
```

## 4.2 Fitting the model used in the review

The “Software reviews of multilevel modeling packages” fit a simple model that includes fixed effects for urban, age, and livch and a simple additive random effect by district. We reproduce this fit as

```
> system.time(Cm1 <- lmer(use ~ age + urban + livch + (1 | district),
+   Contraception, binomial))
      user system elapsed
      3.25   0.00   3.25

> Cm1
Generalized linear mixed model fit by the Laplace approximation
Formula: use ~ age + urban + livch + (1 | district)
Data: Contraception
AIC BIC logLik deviance
2428 2467 -1207 2414
Random effects:
Groups Name Variance Std.Dev.
district (Intercept) 0.21239 0.46086
Number of obs: 1934, groups: district, 60

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.689710   0.145496 -11.613 < 2e-16
age          -0.026595   0.007828  -3.398 0.00068
urbanY        0.732992   0.118420   6.190 6.02e-10
livch1        1.109184   0.156825   7.073 1.52e-12
livch2        1.376396   0.173309   7.942 1.99e-15
livch3+       1.345181   0.177771   7.567 3.82e-14

Correlation of Fixed Effects:
      (Intr) age    urbanY livch1 livch2
age      0.447
urbanY  -0.301 -0.047
livch1  -0.591 -0.211 0.060
livch2  -0.633 -0.378 0.095 0.489
livch3+ -0.750 -0.673 0.100 0.540 0.620

> system.time(Cm3 <- lmer(use ~ age + urban + livch + (1 | district),
+   Contraception, binomial, nAGQ = 5))
      user system elapsed
      5.700   0.000   5.699

> Cm3
Generalized linear mixed model fit by the adaptive Gaussian Hermite approximation
Formula: use ~ age + urban + livch + (1 | district)
Data: Contraception
AIC BIC logLik deviance
2452 2491 -1219 2438
Random effects:
Groups Name Variance Std.Dev.
district (Intercept) 0.21268 0.46117
Number of obs: 1934, groups: district, 60

Fixed effects:
```

```

              Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.689739   0.145515 -11.612 < 2e-16
age          -0.026596   0.007828  -3.398 0.00068
urbanY        0.732933   0.118426   6.189 6.06e-10
livch1        1.109198   0.156828   7.073 1.52e-12
livch2        1.376410   0.173313   7.942 1.99e-15
livch3+       1.345214   0.177775   7.567 3.82e-14

Correlation of Fixed Effects:
      (Intr) age    urbanY livch1 livch2
age      0.447
urbanY   -0.301 -0.047
livch1   -0.591 -0.211  0.060
livch2   -0.633 -0.378  0.095  0.489
livch3+  -0.750 -0.673  0.100  0.540  0.620

> system.time(Cm4 <- lmer(use ~ age + urban + livch + (urban /
+   district), Contraception, binomial))

      user system elapsed
4.790    0.000    4.798

> Cm4

Generalized linear mixed model fit by the Laplace approximation
Formula: use ~ age + urban + livch + (urban | district)
Data: Contraception
      AIC   BIC logLik deviance
2417 2467  -1200    2399
Random effects:
Groups   Name              Variance Std.Dev. Corr
district (Intercept) 0.38125   0.61746
          urbanY      0.64194   0.80121  -0.798
Number of obs: 1934, groups: district, 60

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.711851   0.157433 -10.874 < 2e-16
age          -0.026520   0.007922  -3.348 0.000815
urbanY        0.815285   0.166405   4.899 9.61e-07
livch1        1.125698   0.158383   7.107 1.18e-12
livch2        1.368331   0.175012   7.819 5.35e-15
livch3+       1.354733   0.180069   7.523 5.34e-14

Correlation of Fixed Effects:
      (Intr) age    urbanY livch1 livch2
age      0.418
urbanY   -0.481 -0.036
livch1   -0.552 -0.211  0.039
livch2   -0.590 -0.378  0.070  0.490
livch3+  -0.698 -0.674  0.064  0.539  0.618

```

### 4.3 Data exploration

As with the Exam data, we examine several data plots when formulating a model for these data. Because the response is either "Y" or "N", plots of

the data points themselves tend to be uninformative because of overplotting. However, with a continuous covariate such as `age` it is useful to plot the scatterplot smoother line for the response versus the covariate to see the trend of the probability.

The model being fit assumes that the logistic transformation of the probability of a woman using artificial contraception is approximately linear in `age` given her urban/rural status and the number of live children she has. Figure~16, produced by shows that this is not the case. Indeed the proportion of women using artificial contraception is more like a quadratic in age. Very young women and older women are less likely to use artificial contraception.

We also see that the differences according to the number of live children are predominantly differences between those women who have live children and those who don't.

## 4.4 Reformulated models

We include a quadratic term in `age`

```
> (Cm6 <- lmer(use ~ age + I(age^2) + urban + livch + (1/district),
+             Contraception, binomial))

Generalized linear mixed model fit by the Laplace approximation
Formula: use ~ age + I(age^2) + urban + livch + (1 | district)
Data: Contraception
    AIC   BIC logLik deviance
2389 2433  -1186     2373
Random effects:
    Groups   Name      Variance Std.Dev.
district (Intercept) 0.22586  0.47524
Number of obs: 1934, groups: district, 60

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.0350725  0.1743606  -5.936 2.91e-09
age           0.0035327  0.0092311   0.383  0.702
I(age^2)     -0.0045623  0.0007252  -6.291 3.15e-10
urbanY        0.6972694  0.1198788   5.816 6.01e-09
livch1        0.8150442  0.1621898   5.025 5.03e-07
livch2        0.9165119  0.1850995   4.951 7.37e-07
livch3+       0.9150212  0.1857689   4.926 8.41e-07

Correlation of Fixed Effects:
      (Intr) age      I(g^2) urbanY livch1 livch2
age      0.572
I(age^2) -0.547 -0.478
urbanY    -0.270 -0.054  0.022
livch1    -0.610 -0.274  0.253  0.068
livch2    -0.681 -0.456  0.357  0.100  0.525
livch3+   -0.763 -0.698  0.325  0.108  0.565  0.652

> anova(Cm1, Cm6)
```

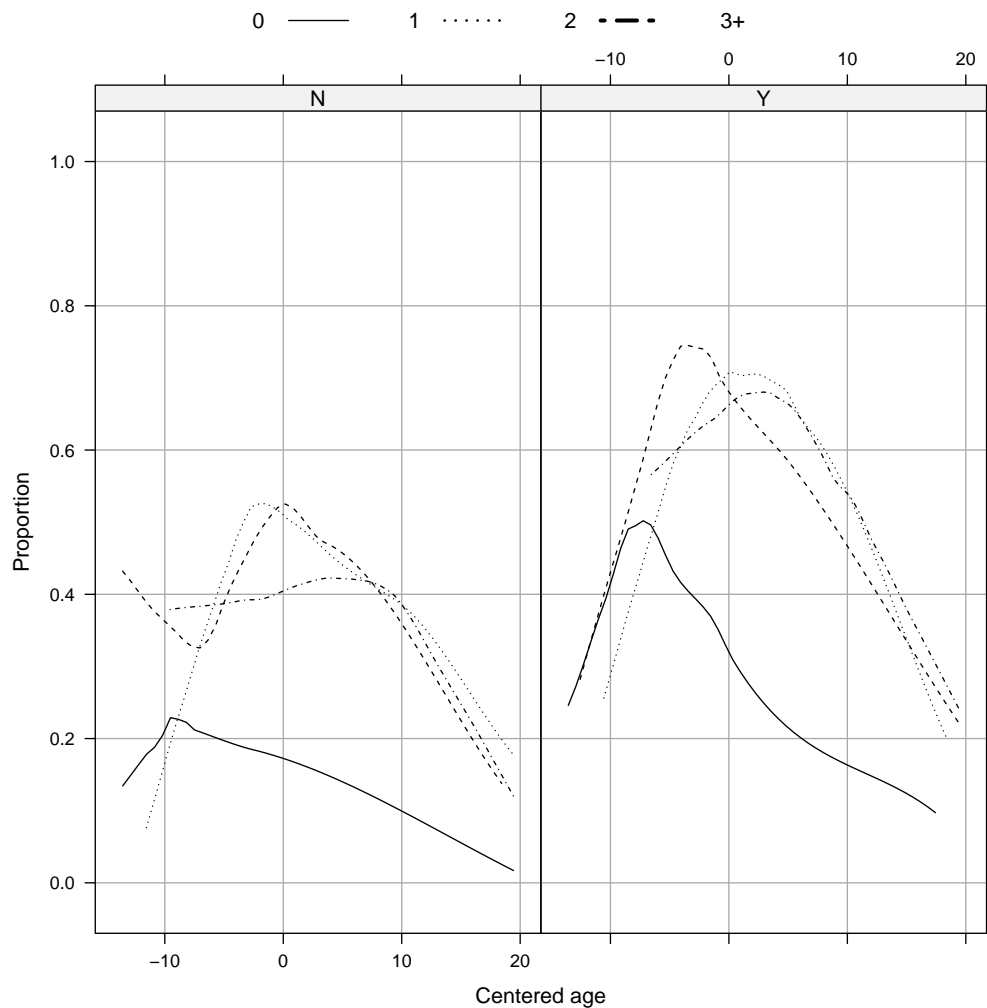


Figure 16: Scatterplot smoother curves of the use of artificial contraception versus age for women in the Bangladesh fertility survey. The panel on the left depicts the proportion for rural women and the panel on the right depicts the proportion for urban women.

```
Data: Contraception
Models:
Cm1: use ~ age + urban + livch + (1 | district)
Cm6: use ~ age + I(age^2) + urban + livch + (1 | district)
      Df      AIC      BIC logLik Chisq Chi Df Pr(>Chisq)
Cm1   7 2427.6 2466.6 -1206.8
Cm6   8 2388.7 2433.3 -1186.4 40.887      1 1.613e-10
```

and we see that, as indicated by the plots, the quadratic term is highly significant.

We can check if the differences in the number of live children is primarily the difference between no children and any children by fitting a model with the indicator of any children as one of the terms in the fixed effects.

```
> Contraception <- within(Contraception, ch <- factor(ifelse(livch == 0, "N", "Y")))
> (Cm7 <- lmer(use ~ age + I(age^2) + urban + ch + (1|district),
+             Contraception, binomial))
```

Generalized linear mixed model fit by the Laplace approximation

Formula: use ~ age + I(age^2) + urban + ch + (1 | district)

```
Data: Contraception
      AIC      BIC logLik deviance
2385 2419 -1187      2373
```

Random effects:

```
Groups      Name      Variance Std.Dev.
district (Intercept) 0.22470 0.47402
```

Number of obs: 1934, groups: district, 60

Fixed effects:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-1.0064262	0.1678949	-5.994	2.04e-09
age	0.0062563	0.0078404	0.798	0.425
I(age^2)	-0.0046354	0.0007163	-6.471	9.73e-11
urbanY	0.6929504	0.1196687	5.791	7.01e-09
chY	0.8603757	0.1473539	5.839	5.26e-09

Correlation of Fixed Effects:

	(Intr)	age	I(g^2)	urbanY
age		0.530		
I(age^2)	-0.539	-0.504		
urbanY	-0.265	-0.031	0.014	
chY	-0.801	-0.565	0.349	0.100

```
> anova(Cm6, Cm7)
```

Data: Contraception

Models:

```
Cm7: use ~ age + I(age^2) + urban + ch + (1 | district)
Cm6: use ~ age + I(age^2) + urban + livch + (1 | district)
      Df      AIC      BIC logLik Chisq Chi Df Pr(>Chisq)
Cm7   6 2385.2 2418.6 -1186.6
Cm6   8 2388.7 2433.3 -1186.4 0.4571      2 0.7957
```

Based on the change in the log-likelihood the simpler model is adequate. We check for an interaction with the urban factor.

```
> (Cm8 <- lmer(use ~ age + I(age^2) + urban * ch + (1|district),
+             Contraception, binomial))
Generalized linear mixed model fit by the Laplace approximation
Formula: use ~ age + I(age^2) + urban * ch + (1 | district)
Data: Contraception
AIC BIC logLik deviance
2386 2425 -1186 2372
Random effects:
Groups Name Variance Std.Dev.
district (Intercept) 0.22433 0.47364
Number of obs: 1934, groups: district, 60

Fixed effects:
Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.0966297 0.1908554 -5.746 9.15e-09
age 0.0056149 0.0078667 0.714 0.475
I(age^2) -0.0045760 0.0007185 -6.369 1.91e-10
urbanY 0.8883821 0.2229111 3.985 6.74e-05
chY 0.9673774 0.1814047 5.333 9.68e-08
urbanY:chY -0.2610103 0.2506490 -1.041 0.298

Correlation of Fixed Effects:
(Intr) age I(g^2) urbanY chY
age 0.501
I(age^2) -0.509 -0.510
urbanY -0.522 -0.080 0.073
chY -0.849 -0.501 0.328 0.536
urbanY:chY 0.465 0.076 -0.076 -0.844 -0.578
```

The interaction is not significant.

We can check if there is significant variation by district in the effect of urban versus rural or in the effect of any children versus no children by fitting further models

```
> (Cm9 <- lmer(use ~ age + I(age^2) + urban + ch + (urban|district),
+             Contraception, binomial))
Generalized linear mixed model fit by the Laplace approximation
Formula: use ~ age + I(age^2) + urban + ch + (urban | district)
Data: Contraception
AIC BIC logLik deviance
2377 2422 -1180 2361
Random effects:
Groups Name Variance Std.Dev. Corr
district (Intercept) 0.38255 0.61851
urbanY 0.54549 0.73857 -0.794
Number of obs: 1934, groups: district, 60

Fixed effects:
Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.0373488 0.1783436 -5.817 6.01e-09
age 0.0057913 0.0079170 0.732 0.464
I(age^2) -0.0045541 0.0007235 -6.294 3.09e-10
urbanY 0.7702697 0.1608410 4.789 1.68e-06
chY 0.8723357 0.1487625 5.864 4.52e-09
```



```

Correlation of Fixed Effects:
              (Intr) age      I(g^2) urbanY
age              0.507
I(age^2)        -0.515 -0.505
urbanY          -0.428 -0.031  0.012
chY             -0.761 -0.566  0.351  0.073

> anova(Cm7, Cm9)

Data: Contraception
Models:
Cm7: use ~ age + I(age^2) + urban + ch + (1 | district)
Cm9: use ~ age + I(age^2) + urban + ch + (urban | district)
      Df    AIC    BIC logLik Chisq Chi Df Pr(>Chisq)
Cm7   6 2385.2 2418.6 -1186.6
Cm9   8 2377.0 2421.5 -1180.5 12.210      2   0.002232

> (Cm10 <- lmer(use ~ age + I(age^2) + urban + ch + (ch|district),
+               Contraception, binomial))

Generalized linear mixed model fit by the Laplace approximation
Formula: use ~ age + I(age^2) + urban + ch + (ch | district)
      Data: Contraception
      AIC   BIC logLik deviance
2386 2431 -1185      2370
Random effects:
Groups   Name              Variance Std.Dev. Corr
district (Intercept)  0.45734   0.67627
          chY          0.29002   0.53854  -0.722
Number of obs: 1934, groups: district, 60

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.0653157   0.1859469  -5.729 1.01e-08
age           0.0069940   0.0079092   0.884  0.377
I(age^2)     -0.0046503   0.0007208  -6.451 1.11e-10
urbanY        0.6964071   0.1205490   5.777 7.61e-09
chY           0.9187713   0.1717604   5.349 8.84e-08

Correlation of Fixed Effects:
              (Intr) age      I(g^2) urbanY
age              0.485
I(age^2)        -0.494 -0.512
urbanY          -0.236 -0.027  0.011
chY             -0.842 -0.491  0.308  0.080

> anova(Cm7, Cm10)

Data: Contraception
Models:
Cm7: use ~ age + I(age^2) + urban + ch + (1 | district)
Cm10: use ~ age + I(age^2) + urban + ch + (ch | district)
      Df    AIC    BIC logLik Chisq Chi Df Pr(>Chisq)
Cm7   6 2385.2 2418.6 -1186.6
Cm10  8 2386.0 2430.5 -1185.0  3.222      2   0.1997

```

Variation due to district in the effect of any children versus no children does not appear to be significant but there is a significant variation in the

Figure 17: Plots of the estimated posterior probability densities from a Markov Chain Monte Carlo sample of the parameters in the model Cm11 fit to the Contraception data.

effect of urban versus rural. It is interesting that AIC (Akaike's Information Criterion) and BIC (Schwartz's Bayesian Information Criterion) lead to different conclusions in the comparison of Cm7 versus Cm9. For both these criteria the preferred model is the one with the lower value of the criterion. Hence, AIC prefers Cm9 and BIC prefers Cm7. The BIC criterion puts a heavier penalty on having a greater number of parameters.

At present the code for creating a Markov Chain Monte Carlo sample from the distribution of the parameters in a generalized linear mixed model does not allow for random effects of dimension greater than one so we produce a sample from the posterior distribution of the parameters in model Cm11

```
> (Cm11 <- lmer(use ~ age + I(age^2) + urban + ch + (1 | district),
+ Contraception, binomial))
```

Generalized linear mixed model fit by the Laplace approximation

Formula: use ~ age + I(age^2) + urban + ch + (1 | district)

Data: Contraception

AIC BIC logLik deviance

2385 2419 -1187 2373

Random effects:

Groups	Name	Variance	Std.Dev.
district	(Intercept)	0.22470	0.47402

Number of obs: 1934, groups: district, 60

Fixed effects:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-1.0064262	0.1678949	-5.994	2.04e-09
age	0.0062563	0.0078404	0.798	0.425
I(age^2)	-0.0046354	0.0007163	-6.471	9.73e-11
urbanY	0.6929504	0.1196687	5.791	7.01e-09
chY	0.8603757	0.1473539	5.839	5.26e-09

Correlation of Fixed Effects:

	(Intr)	age	I(g^2)	urbanY
age		0.530		
I(age^2)	-0.539	-0.504		
urbanY	-0.265	-0.031	0.014	
chY	-0.801	-0.565	0.349	0.100

Density plots (Figure~17) and normal probability plots (Figure~18)

for this sample show that the fixed-effects parameters are very well approximated by a normal distribution but the variance of the random effects shows noticeable skewness.

Figure 18: Normal probability plots of the Markov Chain Monte Carlo samples from the posterior densities of the parameters in the model Cm11 fit to the Contraception data.

## 5 Growth curve model for repeated measures data

```
> str(Oxboys)

'data.frame':      234 obs. of  4 variables:
 $ Subject : Factor w/ 26 levels "1","10","11",...: 1 1 1 1 1 1 1 1 1 12 ...
 $ age      : num  -1 -0.7479 -0.463 -0.1643 -0.0027 ...
 $ height   : num  140 143 145 147 148 ...
 $ Occasion: Factor w/ 9 levels "1","2","3","4",...: 1 2 3 4 5 6 7 8 9 1 ...
- attr(*, "ginfo")=List of 7
 ..$ formula      :Class 'formula' length 3 height ~ age | Subject
 .. ..- attr(*, ".Environment")=<environment: R_GlobalEnv>
 ..$ order.groups: logi TRUE
 ..$ FUN           :function (x)
 .. ..- attr(*, "source")= chr "function (x) max(x, na.rm = TRUE)"
 ..$ outer         : NULL
 ..$ inner         : NULL
 ..$ labels        :List of 2
 .. ..$ age        : chr "Centered age"
 .. ..$ height      : chr "Height"
 ..$ units         :List of 1
 .. ..$ height      : chr "(cm)"

> system.time(mX1 <- lmer(height ~ age + I(age^2) + I(age^3) + I(age^4) + (age + I(age^2)|Subject),
+               Oxboys))

   user  system elapsed 
 1.190   0.000   1.186 

> summary(mX1)

Linear mixed model fit by REML
Formula: height ~ age + I(age^2) + I(age^3) + I(age^4) + (age + I(age^2) |      Subject)
Data: Oxboys
    AIC   BIC logLik deviance REMLdev
 651.9 693.4 -314.0   625.4   627.9
Random effects:
Groups   Name              Variance Std.Dev. Corr
Subject (Intercept) 64.03456 8.00216
          age         2.86417 1.69239 0.614
          I(age^2)    0.67429 0.82115 0.215 0.658
Residual              0.21737 0.46623
Number of obs: 234, groups: Subject, 26

Fixed effects:
              Estimate Std. Error t value
(Intercept) 149.0189      1.5704   94.89
age           6.1742      0.3565   17.32
I(age^2)      1.1282      0.3514    3.21
I(age^3)      0.4539      0.1625    2.79
```

```

I(age^4)      -0.3769      0.3002      -1.26

Correlation of Fixed Effects:
              (Intr) age      I(g^2) I(g^3)
age           0.572
I(age^2)      0.076  0.264
I(age^3)     -0.001 -0.340  0.025
I(age^4)      0.021  0.016 -0.857 -0.021

> system.time(mX2 <- lmer(height ~ poly(age,4) + (age + I(age^2)|Subject), Oxboys))

      user system elapsed
      1.180   0.000   1.181

> summary(mX2)

Linear mixed model fit by REML
Formula: height ~ poly(age, 4) + (age + I(age^2) | Subject)
Data: Oxboys
AIC   BIC logLik deviance REMLdev
640.9 682.3 -308.4   625.4   616.9
Random effects:
Groups   Name             Variance Std.Dev. Corr
Subject  (Intercept)  64.03426  8.00214
          age          2.86414  1.69238  0.614
          I(age^2)     0.67429  0.82115  0.215 0.658
Residual                    0.21737  0.46623
Number of obs: 234, groups: Subject, 26

Fixed effects:
              Estimate Std. Error t value
(Intercept)   149.5198     1.5903   94.02
poly(age, 4)1    64.5409     3.3278   19.39
poly(age, 4)2     4.2032     1.0236    4.11
poly(age, 4)3     1.2908     0.4663    2.77
poly(age, 4)4    -0.5855     0.4663   -1.26

Correlation of Fixed Effects:
              (Intr) p(,4)1 p(,4)2 p(,4)3
poly(ag,4)1  0.631
poly(ag,4)2  0.230  0.583
poly(ag,4)3  0.000  0.000  0.000
poly(ag,4)4  0.000  0.000  0.000  0.000

```

## 6 Cross-classification model

```

> str(ScotsSec)

'data.frame':   3435 obs. of  6 variables:
 $ verbal : num  11 0 -14 -6 -30 -17 -17 -11 -9 -19 ...
 $ attain : num  10 3 2 3 2 2 4 6 4 2 ...
 $ primary: Factor w/ 148 levels "1","2","3","4",...: 1 1 1 1 1 1 1 1 1 1 ...
 $ sex    : Factor w/ 2 levels "M","F": 1 2 1 1 2 2 2 1 1 1 ...
 $ social : num   0 0 0 20 0 0 0 0 0 0 ...
 $ second : Factor w/ 19 levels "1","2","3","4",...: 9 9 9 9 9 9 1 1 9 9 ...

> system.time(mS1 <- lmer(attain ~ sex + (1|primary) + (1|second), ScotsSec))

```

```

      user system elapsed
0.340    0.000    0.339

> summary(mS1)
Linear mixed model fit by REML
Formula: attain ~ sex + (1 | primary) + (1 | second)
Data: ScotsSec
   AIC   BIC logLik deviance REMLdev
17138 17169  -8564   17123   17128
Random effects:
Groups   Name             Variance Std.Dev.
primary (Intercept)  1.10962   1.0534
second  (Intercept)  0.36966   0.6080
Residual                    8.05511   2.8382
Number of obs: 3435, groups: primary, 148; second, 19

Fixed effects:
              Estimate Std. Error t value
(Intercept)   5.25515     0.18432  28.511
sexF           0.49852     0.09825   5.074

Correlation of Fixed Effects:
      (Intr)
sexF -0.264

```

## 7 Session Info

```
> toLatex(sessionInfo())
```

- R version 2.13.0 Under development (unstable) (2011-02-15 r54430), x86\_64-unknown-linux-gnu
- Locale: LC\_CTYPE=en\_US.utf8, LC\_NUMERIC=C, LC\_TIME=en\_US.utf8, LC\_COLLATE=C, LC\_MONETARY=C, LC\_MESSAGES=en\_US.utf8, LC\_PAPER=en\_US.utf8, LC\_NAME=C, LC\_ADDRESS=C, LC\_TELEPHONE=C, LC\_MEASUREMENT=en\_US.utf8, LC\_IDENTIFICATION=C
- Base packages: base, datasets, grDevices, graphics, methods, stats, utils
- Other packages: Matrix~0.999375-47, lattice~0.19-17, lme4~0.999375-37, mlmRev~1.0-0
- Loaded via a namespace (and not attached): grid~2.13.0, nlme~3.1-98, stats4~2.13.0, tools~2.13.0

## References

G.E.P. Box and G.C. Tiao. *Bayesian Inference in Statistical Analysis*.  
Addison-Wesley, Reading, MA, 1973.