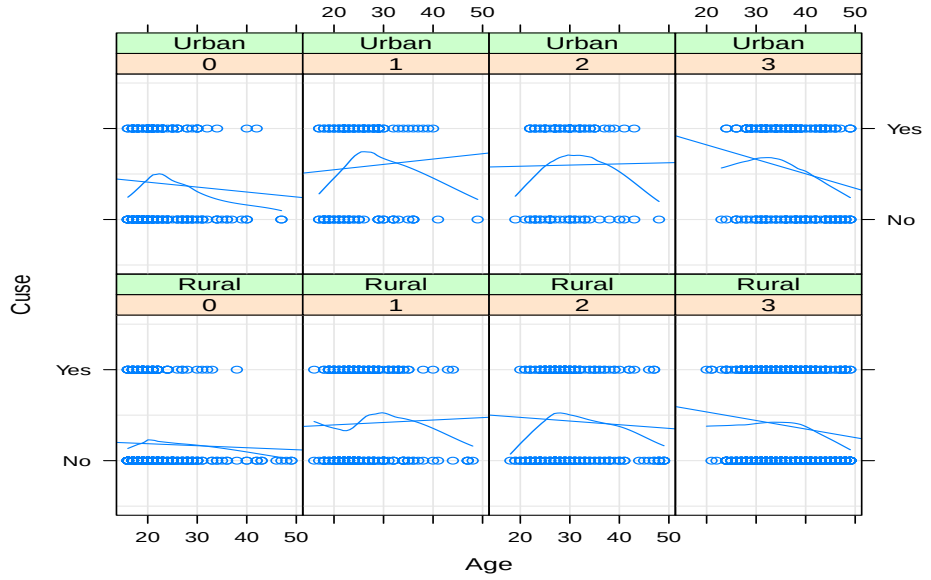
1 Stat 8053, Fall 2011: GLMMs

The data come from a 1988 fertility survey in Bangladesh. Data were collected on 1934 women grouped into 60 districts. The response of interest is whether or not the woman is using contraceptives at the time of the survey. Predictors include age, number of existing children, and whether the district is urban or rural.

```
> url <- "http://tinyurl.com/3dmnytg"
> data <- read.csv(url)
> data$Nchild <- factor(data$Nchild)
> data$Area <- factor(data$Area)
> data$Cage <- data$Age - mean(data$Age)
> require(car)
> library(effects)
> require(lme4)
> print(xyplot(Cuse ~ Age | Nchild * Urban, data, type = c("g",
+     "r", "smooth", "p")))
```



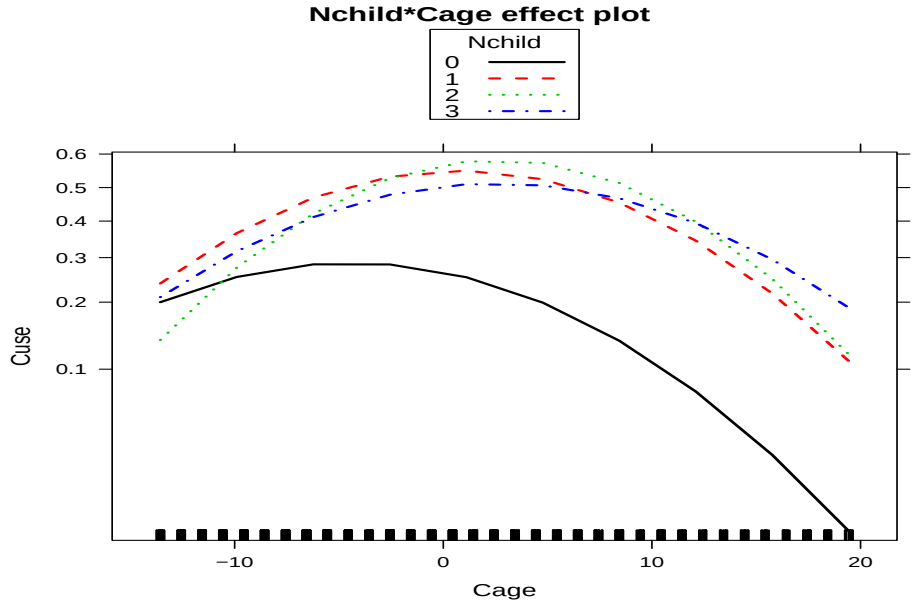
The package `effects` must be loaded before `lme4` because of a problem in `lme4`. The function `xyplot` is from the `lattice` package. When used with Sweave, as I have done here, you need to enclose `xyplot` by a `print` statement. This is also required on the `plot` method for effects, used below, because this function also uses the `lattice` package.

GLM fit(s)

```
> m1 <- glm(Cuse ~ (Nchild + poly(Cage, 2, raw = TRUE) + Urban)^2,
+     data = data, family = binomial)
> Anova(m1)

Analysis of Deviance Table (Type II tests)
```

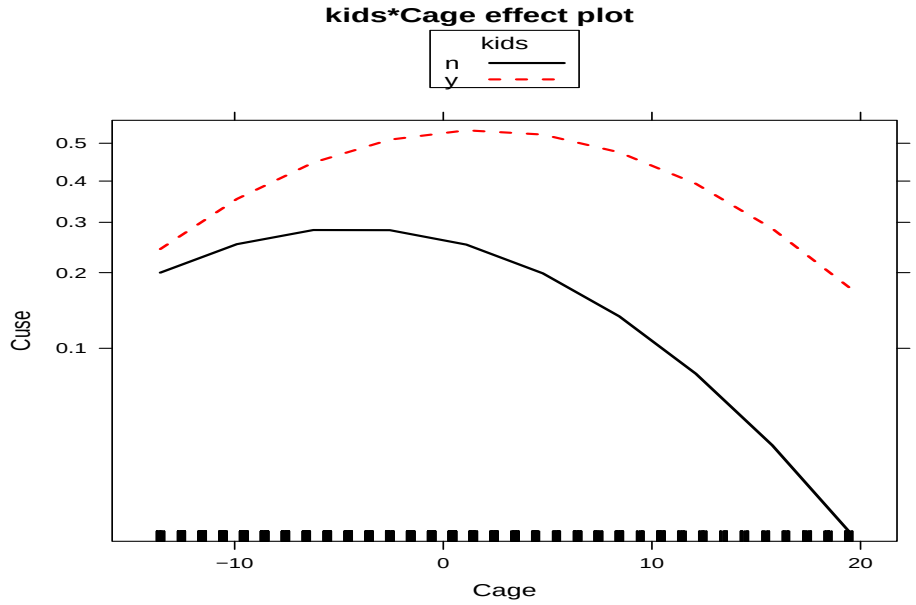
```
Response: Cuse
                                  LR Chisq Df Pr(>Chisq)
Nchild                              33.302  3  2.781e-07
poly(Cage, 2, raw = TRUE)           48.711  2  2.646e-11
Urban                               54.062  1  1.942e-13
Nchild:poly(Cage, 2, raw = TRUE)    12.747  6    0.04723
Nchild:Urban                         2.213  3    0.52932
poly(Cage, 2, raw = TRUE):Urban      1.188  2    0.55211
```

```
> Anova(m2 <- update(m1, ~Nchild * poly(Cage, 2, raw = TRUE) +
+     Urban))

Analysis of Deviance Table (Type II tests)

Response: Cuse
                                  LR Chisq Df Pr(>Chisq)
Nchild                              33.256  3  2.845e-07
poly(Cage, 2, raw = TRUE)           49.531  2  1.756e-11
Urban                               54.062  1  1.942e-13
Nchild:poly(Cage, 2, raw = TRUE)    11.891  6    0.06444

> print(plot(effect("Nchild*poly(Cage, 2, raw=TRUE)", m2), multiline = TRUE,
+     type = "link"))
```



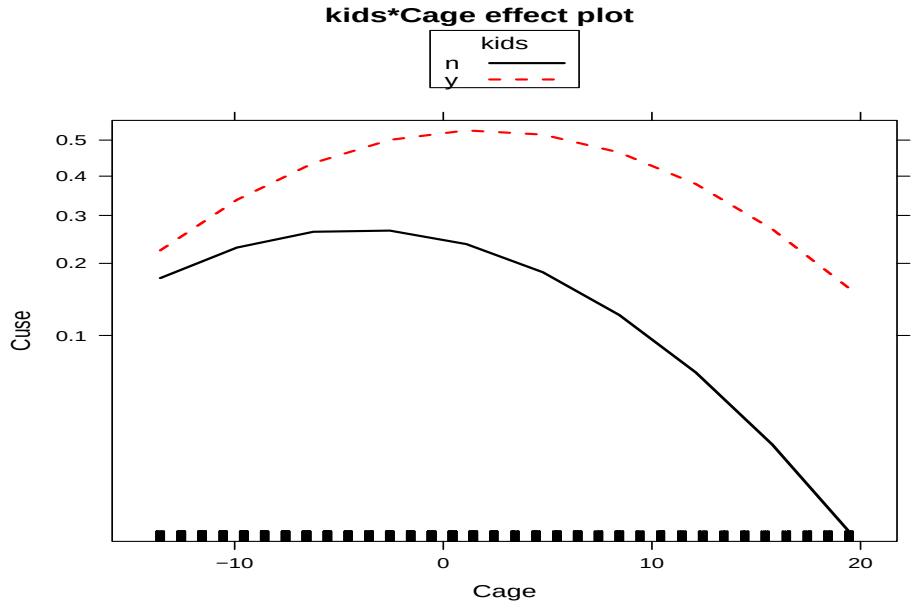

```
> data$kids <- factor(ifelse(data$Nchild == 0, "n", "y"))
> m3 <- update(m2, ~kids * poly(Cage, 2, raw = TRUE) + Urban)
> print(plot(effect("kids*poly(Cage, 2, raw=TRUE)", m3), multiline = TRUE,
+     type = "link"))
> Anova(m3)
```

```
Analysis of Deviance Table (Type II tests)

Response: Cuse
                                LR Chisq Df Pr(>Chisq)
kids                             33.128  1  8.629e-09
poly(Cage, 2, raw = TRUE)        52.231  2  4.551e-12
Urban                            55.326  1  1.021e-13
kids:poly(Cage, 2, raw = TRUE)    8.499  2    0.01427
```



GLMM fit

```
> g2 <- lmer(Cuse ~ kids * poly(Cage, 2, raw = TRUE) + Urban +
+     (1 | Area), data = data, family = binomial)
> Anova(g2)

Analysis of Deviance Table (Type II tests)

Response: Cuse
                                Chisq Df Pr(>Chisq)
kids                          29.8563  1  4.653e-08
poly(Cage, 2, raw = TRUE)     47.9889  2  3.796e-11
Urban                         35.1859  1  2.997e-09
kids:poly(Cage, 2, raw = TRUE)  7.0899  2    0.02887

> print(plot(effect("kids:poly(Cage, 2, raw=TRUE)", g2), multiline = TRUE,
+     type = "link"))
```

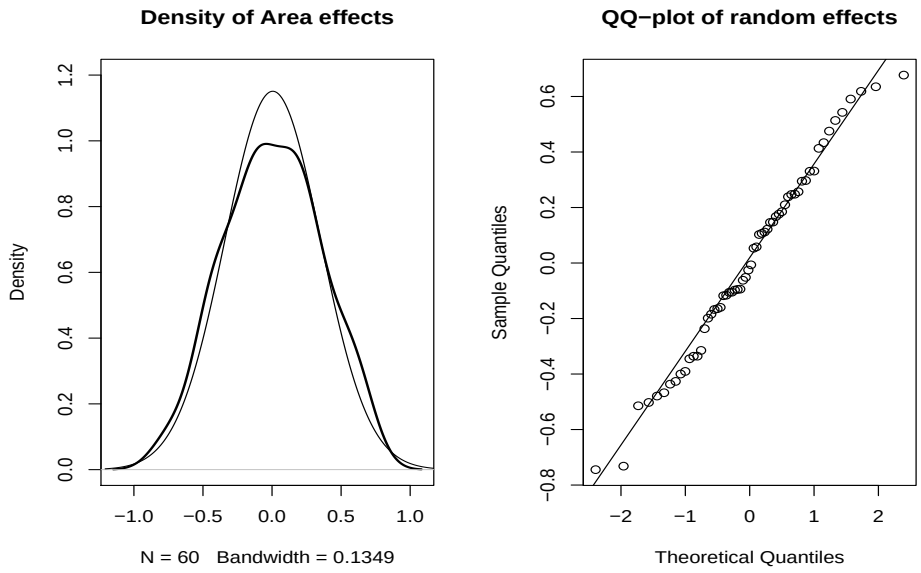


Coefficients

```
> compareCoefs(m3, g2)

Call:
1:glm(formula = Cuse ~ kids + poly(Cage, 2, raw = TRUE) + Urban + kids:poly(Cage,
    2, raw = TRUE), family = binomial, data = data)
2:glmer(formula = Cuse ~ kids * poly(Cage, 2, raw = TRUE) + Urban + (1 | Area),
    data = data, family = binomial)
                                    Est. 1      SE 1    Est. 2      SE 2
(Intercept)                       -1.252583  0.209190 -1.313796  0.224633
kidsy                              1.147070  0.215490  1.199571  0.221429
poly(Cage, 2, raw = TRUE)1        -0.050500  0.029055 -0.050301  0.030656
poly(Cage, 2, raw = TRUE)2        -0.005698  0.002595 -0.006130  0.002737
UrbanUrban                         0.788797  0.106694  0.713557  0.120294
kidsy:poly(Cage, 2, raw = TRUE)1   0.069894  0.030559  0.071048  0.032155
kidsy:poly(Cage, 2, raw = TRUE)2   0.000292  0.002731  0.000412  0.002873

> exp(fixef(g2))

                     (Intercept)                                    kidsy
                       0.2687978                                3.3186926
      poly(Cage, 2, raw = TRUE)1       poly(Cage, 2, raw = TRUE)2
                       0.9509428                                0.9938891
                      UrbanUrban kidsy:poly(Cage, 2, raw = TRUE)1
                       2.0412399                                1.0736324
kidsy:poly(Cage, 2, raw = TRUE)2
                       1.0004123
```

Random effects

```
> str(ranef(g2))

List of 1
 $ Area:'data.frame':        60 obs. of  1 variable:
  ..$ (Intercept): num [1:60] -0.7445 -0.0254 0.2099 0.1673 0.1069 ...
 - attr(*, "class")= chr "ranef.mer"

> par(mfrow = c(1, 2))
> plot(density(ff <- ranef(g2)$Area[, 1]), main = "Density of Area effects",
+     ylim = c(0, 1.2), lwd = 2)
> xx <- seq(-3.5, 3.5, length = 200)
> lines(mean(ff) + sd(ff) * xx, dnorm(xx)/sd(ff))
> qqnorm(ff, main = "QQ-plot of random effects")
> qqline(ff)
```




```
> print(g3 <- update(g2, ~. - (1 | Area) + (Urban - 1 | Area)),
+     corr = FALSE)

Generalized linear mixed model fit by the Laplace approximation
Formula: Cuse ~ kids + poly(Cage, 2, raw = TRUE) + Urban + (Urban - 1 |      Area) + (
   Data: data
  AIC  BIC logLik deviance
 2374 2429  -1177     2354
Random effects:
 Groups Name        Variance Std.Dev. Corr
 Area   UrbanRural 0.37832  0.61507
        UrbanUrban 0.19652  0.44330  0.091
Number of obs: 1934, groups: Area, 60
```

```
Fixed effects:
                                   Estimate Std. Error z value Pr(>|z|)
(Intercept)                       -1.3386370  0.2327849  -5.751 8.90e-09
kidsy                              1.2050303  0.2229396   5.405 6.47e-08
poly(Cage, 2, raw = TRUE)1        -0.0478833  0.0302846  -1.581   0.1139
poly(Cage, 2, raw = TRUE)2        -0.0058635  0.0027115  -2.162   0.0306
UrbanUrban                         0.7898778  0.1600359   4.936 7.99e-07
kidsy:poly(Cage, 2, raw = TRUE)1  0.0679756  0.0318206   2.136   0.0327
kidsy:poly(Cage, 2, raw = TRUE)2  0.0002353  0.0028504   0.083   0.9342
```

```
> anova(g2, g3, text = "Chisq")

Data: data
Models:
g2: Cuse ~ kids * poly(Cage, 2, raw = TRUE) + Urban + (1 | Area)
g3: Cuse ~ kids + poly(Cage, 2, raw = TRUE) + Urban + (Urban - 1 |
g3:     Area) + kids:poly(Cage, 2, raw = TRUE)
   Df    AIC    BIC  logLik  Chisq Chi Df Pr(>Chisq)
g2  8 2381.2 2425.7 -1182.6
g3 10 2373.5 2429.2 -1176.8 11.637      2   0.002972
```