Dealing with Heteroskedasticity

R-side Variance Modeling

2024-03-13

Contents

```r
library(nlme)
library(spida2)
```

```
Attaching package: 'spida2'

The following object is masked from 'package:nlme':

    getData
```

```r
library(lattice)
library(latticeExtra)
library(latex2exp)
```

1 Introduction

This is an example to illustrate methods to deal with heteroskedasticity. The methods can be used for 'single-level' for which you would use 'lm' in the absence of heteroskedasticity, or for mixed models. To allow for heteroskedasticity, you can use the ‘correlation’ argument in functions in the ‘nlme’ package. For single-level models, use the ‘gls’ function, for linear mixed models, the ‘lme’ function, and for non-linear mixed models, the ‘nlme’ function.

All three functions use the ‘correlation’ argument in the same way.

2 Generating a data set

Pay equity data set for a hypothetical university with two faculties: Medicine and Arts with a higher level and variance in Medicine vs Arts and a different gender gap

```
# Starting data frame:
dd <- expand.grid(Faculty = c("Arts", "Med"), Sex = c("F","M"), n = 1:400)

# sample(1000000,1)
set.seed(576530)
dd <- within(
  dd,
  {
    Age <- 45 + 5 * (Faculty == "Arts") + 5 * (Sex == "M") + 15 * rnorm(n)
```

```r
    ..esal <- 100 + 20 * (Faculty == 'Med') +
      (4 + .3 *(Sex == "M") + .5 * (Faculty == "Med")) * (Age - 30)

    ..sdsal <- 10 + 10 * (Faculty == "Med") + (.2 + .2 * (Faculty == "Med"))* (Age - 30)

    Base <- ..esal + ..sdsal * rnorm(n)
    keep <- Age > 28 & Age < 80
    ..sdsal <- NULL
    ..esal <- NULL
  }
)
tab(dd, ~ Faculty + Sex +keep)
```

```
, , keep = FALSE

          Sex
Faculty     F     M Total
  Arts     42    37    79
  Med      45    46    91
```

```
  Total    87    83    170

, , keep = TRUE

         Sex
Faculty     F     M  Total
  Arts    358   363    721
  Med     355   354    709
  Total   713   717   1430

, , keep = Total

         Sex
Faculty     F     M  Total
  Arts    400   400    800
  Med     400   400    800
  Total   800   800   1600
```

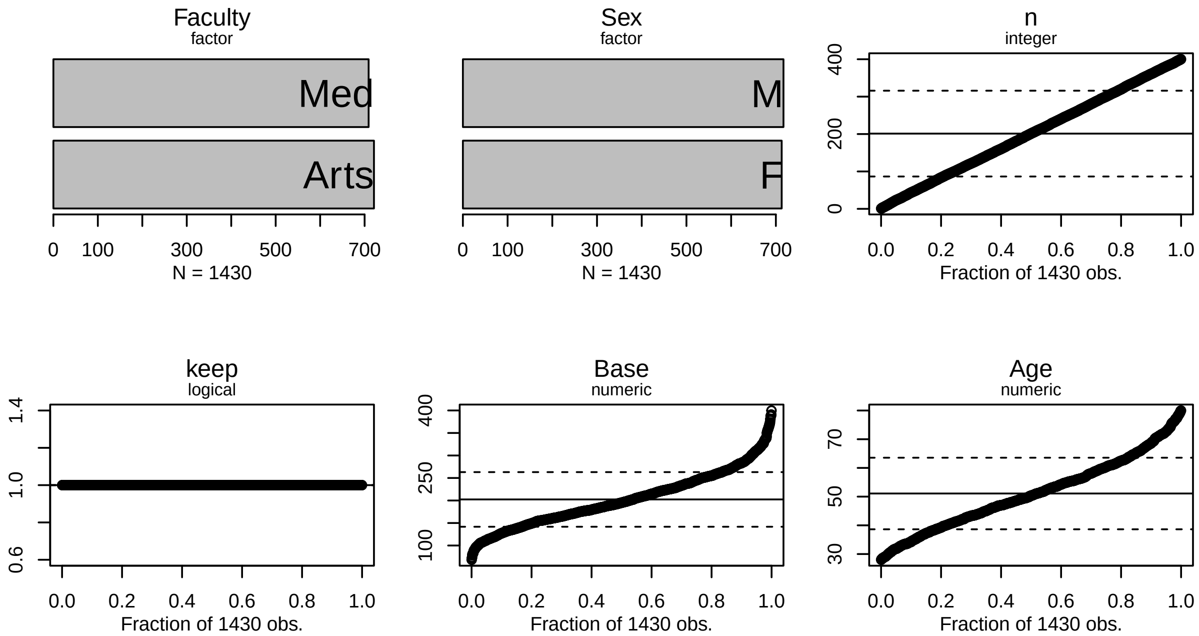
```
dd <- subset(dd, keep)
save(dd, file = 'salary.rda')
```

3 Analysis

```r
load('salary.rda', verbose = TRUE)
```

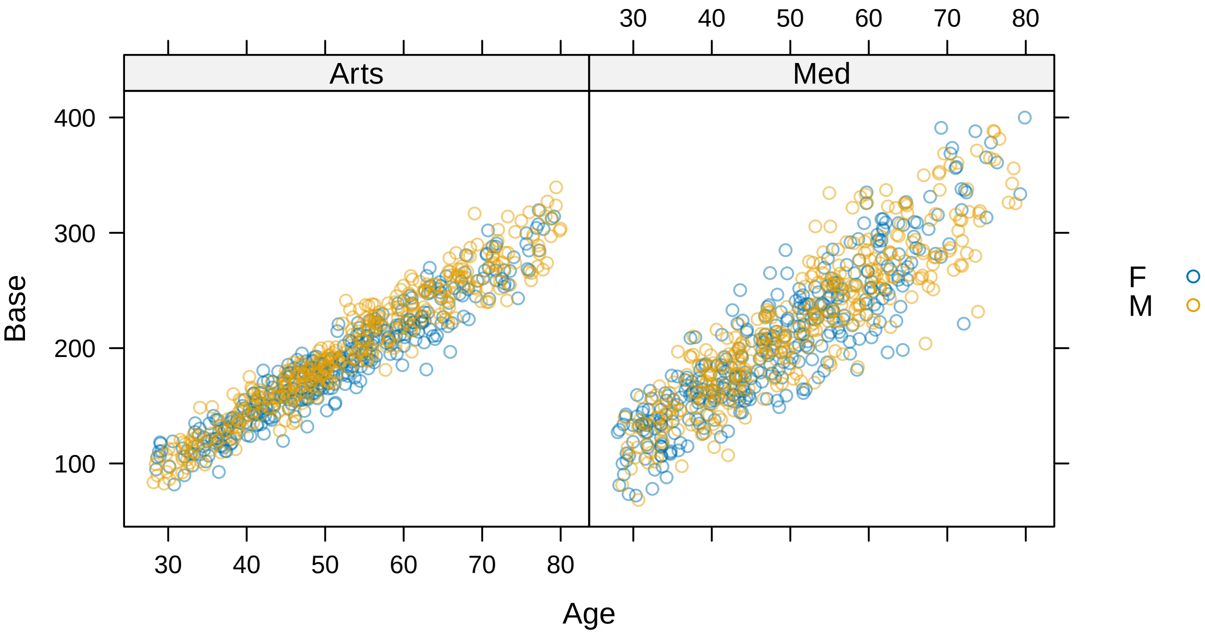
```
Loading objects:
  dd
```

```r
xqplot(dd)
```



Faculty
factor
Med
Arts
0
100
300
500
700
N = 1430
Sex
factor
M
F
0
100
300
500
700
N = 1430
n
integer
400
200
0
0.0
0.2
0.4
0.6
0.8
1.0
Fraction of 1430 obs.
keep
logical
1.4
1.0
0.6
0.0
0.2
0.4
0.6
0.8
1.0
Fraction of 1430 obs.
Base
numeric
400
250
100
0.0
0.2
0.4
0.6
0.8
1.0
Fraction of 1430 obs.
Age
numeric
70
50
30
0.0
0.2
0.4
0.6
0.8
1.0
Fraction of 1430 obs.

```r
xyplot(Base ~ Age | Faculty, dd, groups = Sex, alpha = .5,
       auto.key = T)
```

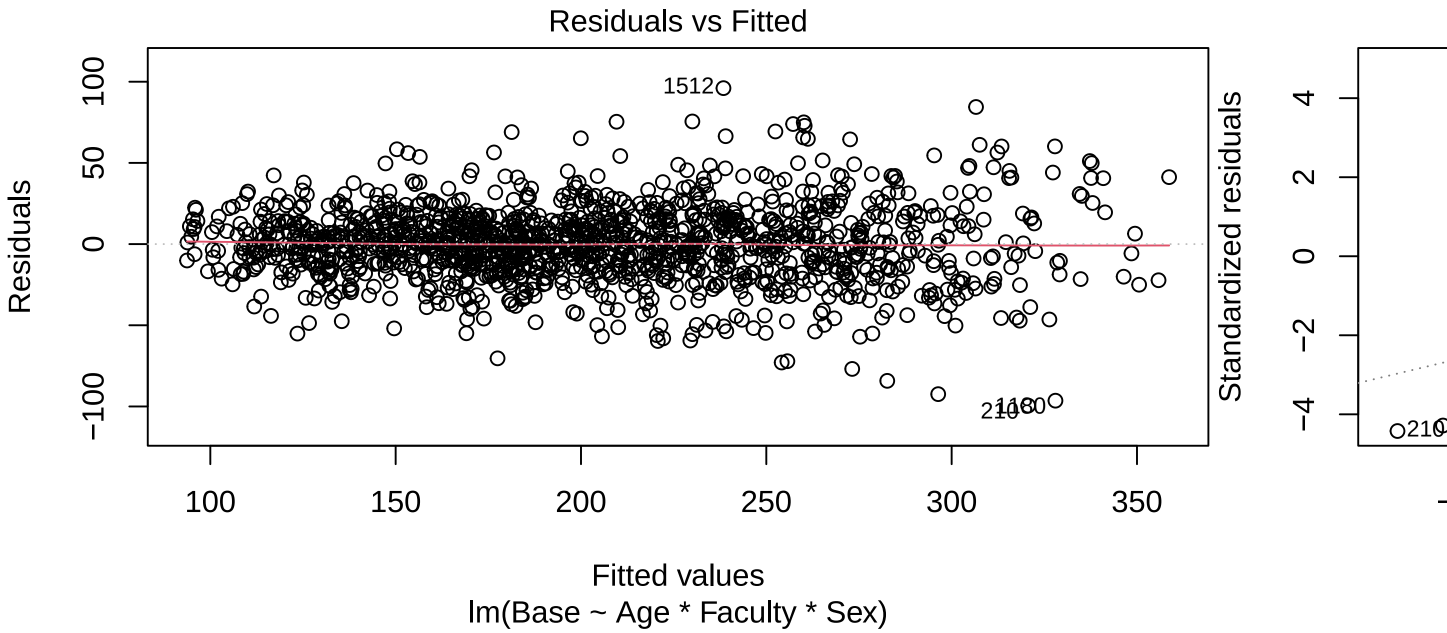

Arts
Med
Base
Age
F
M
400
300
200
100
30
40
50
60
70
80

```r
fit <- lm(Base ~ Age * Faculty * Sex, dd)
summary(fit)
```

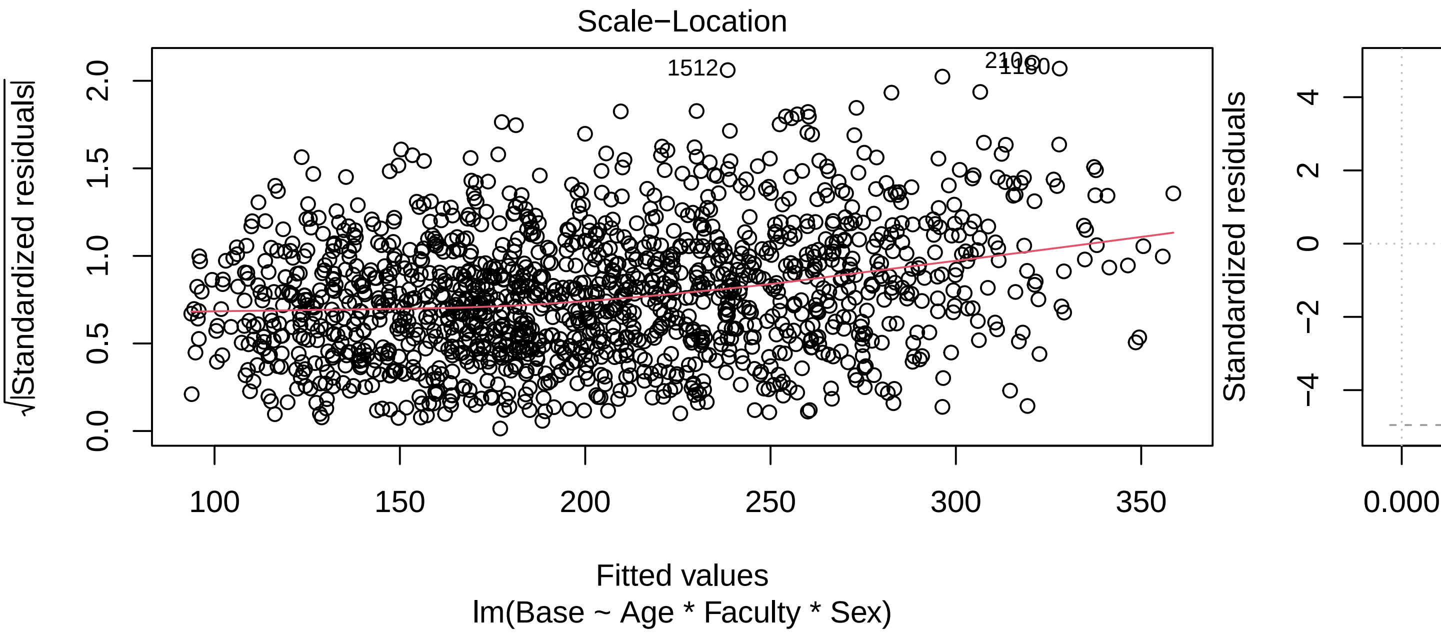
```

Call:
lm(formula = Base ~ Age * Faculty * Sex, data = dd)

Residuals:
   Min     1Q Median     3Q    Max 
-99.48 -13.17   0.04  13.59  96.04 

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)    
(Intercept)        -19.7709     5.3315  -3.708 0.000217 ***
Age                  3.9913     0.1008  39.581  < 2e-16 ***
FacultyMed         -12.2109     7.3111  -1.670 0.095104 .  
SexM                -4.4577     7.3055  -0.610 0.541835    
Age:FacultyMed       0.8994     0.1424   6.314 3.63e-10 ***
Age:SexM             0.2011     0.1358   1.480 0.139020    
```

```
FacultyMed:SexM       15.0498    10.2324   1.471 0.141566
Age:FacultyMed:SexM  -0.3651     0.1951  -1.872 0.061434 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 22.65 on 1422 degrees of freedom
Multiple R-squared:  0.8614,  Adjusted R-squared:  0.8608
F-statistic:  1263 on 7 and 1422 DF,  p-value: < 2.2e-16
```

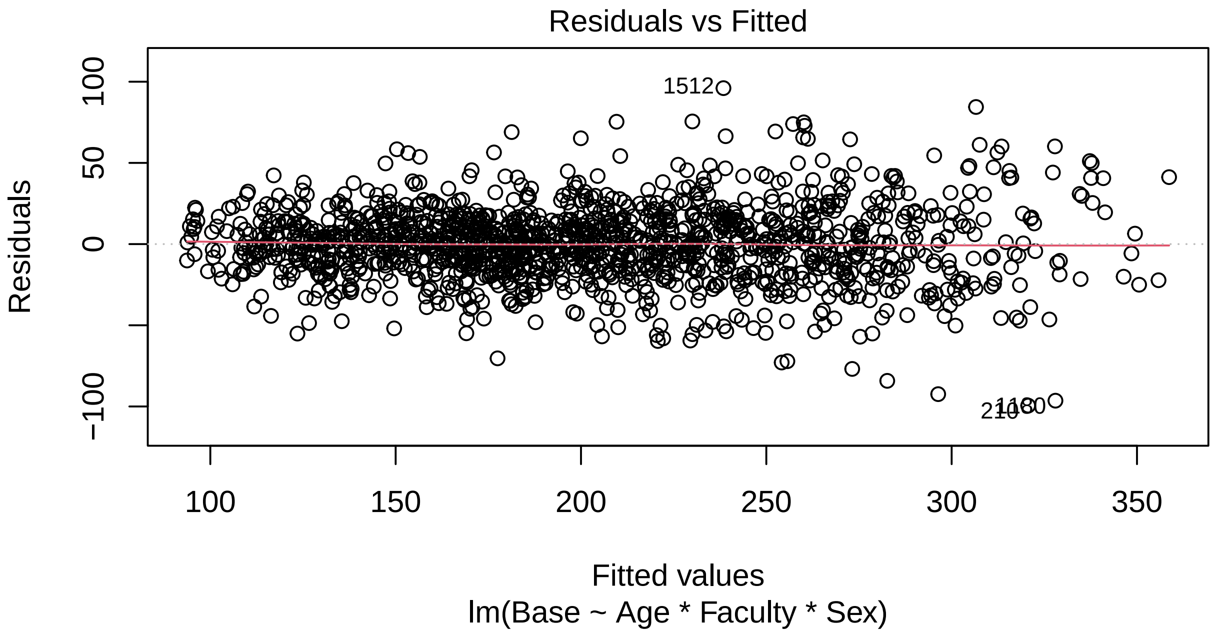
```r
for(i in c(1,2,3,5)) {
   plot(fit, which = i, add.smooth=T, mfcol = c(1,1))
}
```



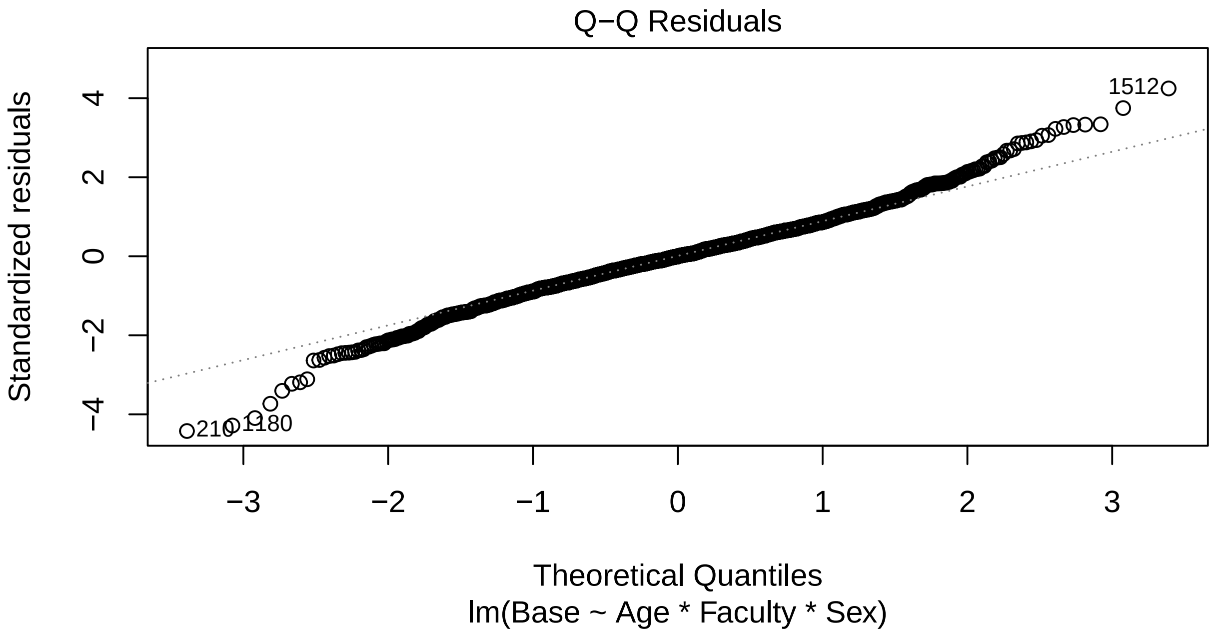
Residuals vs Fitted
Residuals
Fitted values
lm(Base ~ Age * Faculty * Sex)
1512
210
1180
Standardized residuals


Scale−Location
1512
210
1180
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
100
150
200
250
300
350
Fitted values
lm(Base ~ Age * Faculty * Sex)
Standardized residuals
−4
−2
0
2
4
0.000

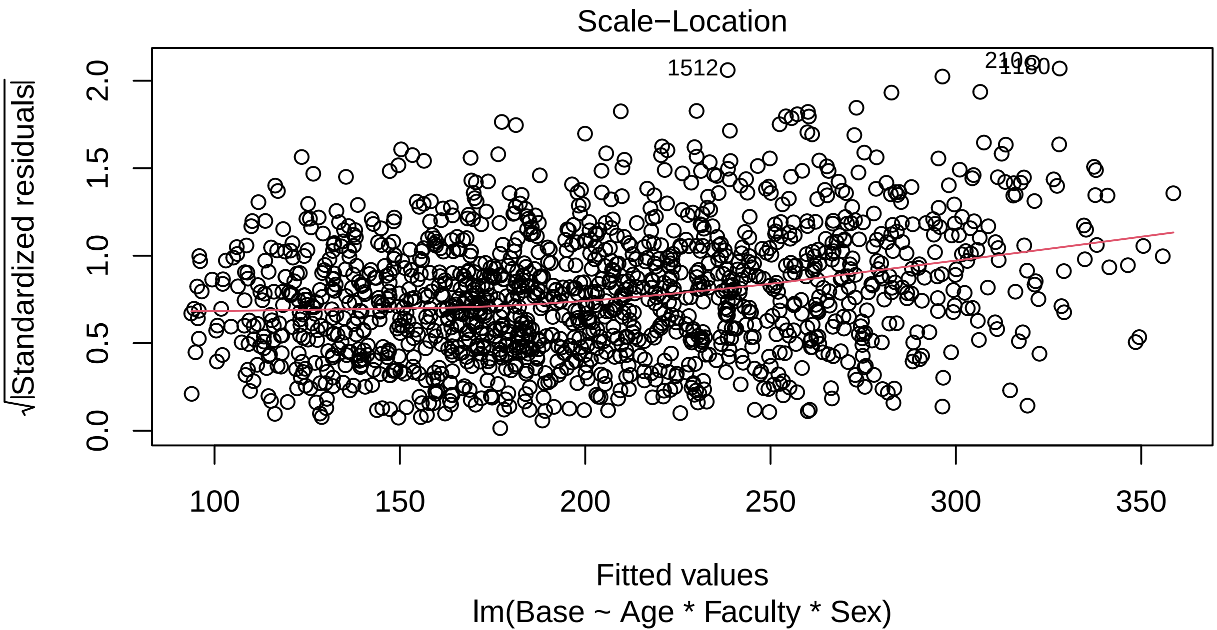
```r
plot(fit, 1, add.smooth = T)
```


Residuals vs Fitted
Residuals
100
50
0
−100
100
150
200
250
300
350
1512
Fitted values
lm(Base ~ Age * Faculty * Sex)

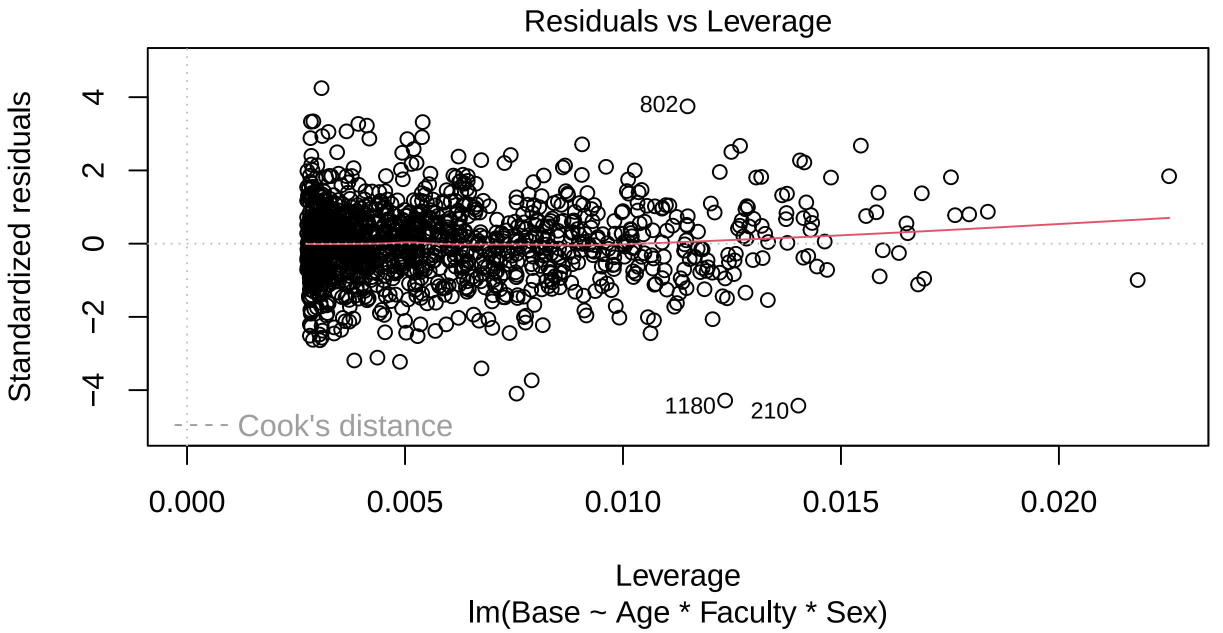
```r
plot(fit, 2, add.smooth = T)
```


Q−Q Residuals
Standardized residuals
Theoretical Quantiles
lm(Base ~ Age * Faculty * Sex)
1512
210
1180
4
2
0
−2
−4
−3
−2
−1
0
1
2
3

```r
plot(fit, 3, add.smooth = T)
```

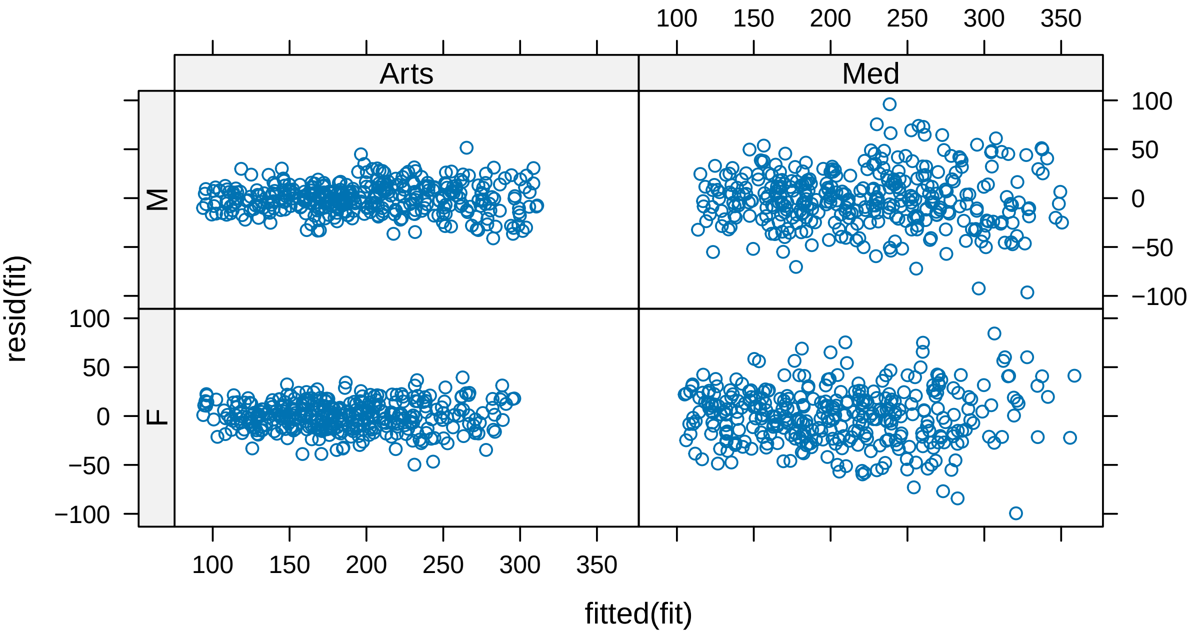

Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
100
150
200
250
300
350
1512
210
1180
Fitted values
lm(Base ~ Age * Faculty * Sex)

```r
plot(fit, 5, add.smooth = T)
```



Residuals vs Leverage
Standardized residuals
4
2
0
−2
−4
802
1180
210
Cook's distance
0.000
0.005
0.010
0.015
0.020
Leverage
lm(Base ~ Age * Faculty * Sex)

```r
xyplot(resid(fit) ~ fitted(fit)|Faculty * Sex, dd) %>%
  useOuterStrips
```


Arts
Med
M
F
resid(fit)
fitted(fit)
100
150
200
250
300
350
100
50
0
−50
−100

Functions in nlme to deal with heteroskedasticity:

Overview:

```
?varClasses
```

- varExp: exponential of a covariate or yhat
- varPower: power of a covariate or yhat
- varConstPower: constant + power of a covariate or yhat
- varConstProp: constant + proportion of a covariate or yhat
- varIdent: different variance in different subgroups
- varFixed: fixed weights given by a covariate
- varComb: combination of variance functions
- You can also build your own but count on spending a days figuring out how to do it

```
fit <- gls(Base ~ Age * Faculty * Sex, dd)  # re fit with gls model
```

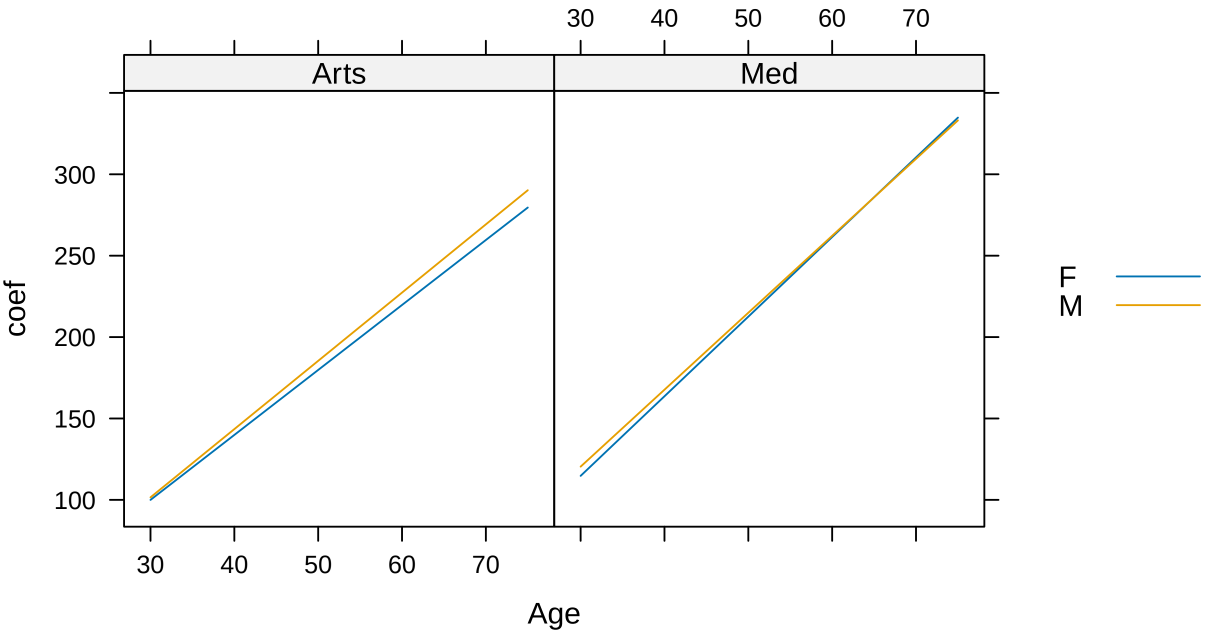
```
pred <- with(dd, pred.grid(Faculty, Sex, Age = seq(30,75,1)))
ww <- as.data.frame(wald(fit, pred = pred))
head(ww)
```

```
             coef         se          U2          L2        p-value  t-value    DF Faculty Sex
```

```
1  99.96824 2.478454 104.9252  95.01133 9.923735e-238 40.33491 1421    Arts   F
2 114.73922 2.196177 119.1316 110.34687  0.000000e+00 52.24498 1421     Med   F
3 101.54254 2.431412 106.4054  96.67972 2.210923e-249 41.76279 1421    Arts   M
4 120.40980 2.390463 125.1907 115.62888 2.068574e-318 50.37091 1421     Med   M
5 103.95955 2.390655 108.7409  99.17824 2.185769e-263 43.48580 1421    Arts   F
6 119.62992 2.112698 123.8553 115.40453  0.000000e+00 56.62425 1421     Med   F
  Age L.(Intercept) L.Age L.FacultyMed L.SexM L.Age:FacultyMed L.Age:SexM
1  30             1    30            0      0                0          0
2  30             1    30            1      0               30          0
3  30             1    30            0      1                0         30
4  30             1    30            1      1               30         30
5  31             1    31            0      0                0          0
6  31             1    31            1      0               31          0
  L.FacultyMed:SexM L.Age:FacultyMed:SexM
1                 0                     0
2                 0                     0
3                 0                     0
4                 1                    30
5                 0                     0
```

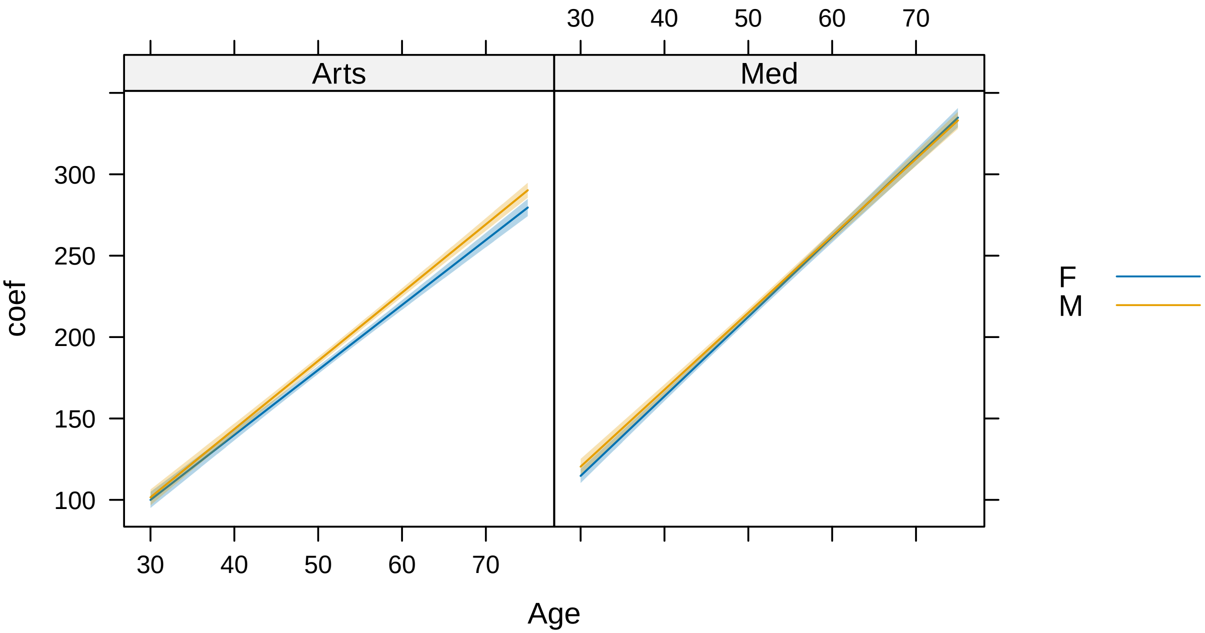
```
      6                   0                        0
```

```r
xyplot(coef ~ Age | Faculty, ww, groups = Sex,
       type = 'l', auto.key = list(space='right'))
```



Arts
Med
coef
Age
300
250
200
150
100
30
40
50
60
70
F
M

```r
xyplot(coef ~ Age | Faculty, ww, groups = Sex,
       type = 'l', auto.key = list(space='right'),
       fit = ww$coef,
       lower = ww$L2,
       upper = ww$U2,
       subscripts = TRUE) +
  glayer(panel.band(...))
```


Arts
Med
coef
Age
300
250
200
150
100
30
40
50
60
70
F
M

Analyzing the Gap

```r
Lmale <- subset(ww, Sex == 'M')$L
Lfemale <- subset(ww, Sex == 'F')$L
Lgap <- Lmale - Lfemale
wgap <- wald(fit,
             Lgap,
             data = subset(ww, Sex == 'F', select = c(Faculty,Age)))
wgap
```

```
   numDF denDF  F-value p-value
1      4  1421 4.609257 0.00106
     Estimate  Std.Error DF   t-value   p-value Lower 0.95 Upper 0.95
3     1.574296 3.471959  1421  0.453432 0.65031 -5.236419   8.385012
4     5.670578 3.246153  1421  1.746861 0.08088 -0.697188  12.038344
7     1.775363 3.353992  1421  0.529328 0.59666 -4.803945   8.354671
8     5.506529 3.128136  1421  1.760323 0.07857 -0.629732  11.642789
11    1.976430 3.237426  1421  0.610494 0.54163 -4.374218   8.327078
12    5.342479 3.012002  1421  1.773730 0.07632 -0.565970  11.250928
15    2.177497 3.122418  1421  0.697375 0.48568 -3.947546   8.302540
```

16	5.178429	2.897978	1421	1.786911	0.07416	-0.506345	10.863204
19	2.378564	3.009145	1421	0.790445	0.42940	-3.524280	8.281408
20	5.014380	2.786322	1421	1.799641	0.07213	-0.451366	10.480126
23	2.579631	2.897813	1421	0.890199	0.37351	-3.104819	8.264081
24	4.850330	2.677331	1421	1.811629	0.07025	-0.401615	10.102275
27	2.780698	2.788652	1421	0.997148	0.31886	-2.689619	8.251015
28	4.686281	2.571343	1421	1.822504	0.06859	-0.357754	9.730316
31	2.981765	2.681929	1421	1.111799	0.26641	-2.279201	8.242730
32	4.522231	2.468745	1421	1.831794	0.06719	-0.320545	9.365007
35	3.182832	2.577946	1421	1.234639	0.21717	-1.874157	8.239821
36	4.358182	2.369978	1421	1.838912	0.06614	-0.290849	9.007213
39	3.383899	2.477048	1421	1.366101	0.17212	-1.475166	8.242963
40	4.194132	2.275540	1421	1.843137	0.06552	-0.269647	8.657911
43	3.584966	2.379628	1421	1.506523	0.13216	-1.082996	8.252927
44	4.030083	2.185993	1421	1.843593	0.06545	-0.258038	8.318203
47	3.786033	2.286130	1421	1.656088	0.09792	-0.698520	8.270586
48	3.866033	2.101962	1421	1.839249	0.06609	-0.257249	7.989315
51	3.987100	2.197056	1421	1.814747	0.06977	-0.322721	8.296920
52	3.701983	2.024134	1421	1.828922	0.06762	-0.268629	7.672595

55	4.188166	2.112963	1421	1.982129	0.04766	0.043304	8.333029
56	3.537934	1.953251	1421	1.811306	0.07030	-0.293630	7.369498
59	4.389233	2.034472	1421	2.157431	0.03114	0.398343	8.380124
60	3.373884	1.890093	1421	1.785036	0.07447	-0.333788	7.081556
63	4.590300	1.962253	1421	2.339301	0.01946	0.741077	8.439524
64	3.209835	1.835459	1421	1.748791	0.08054	-0.390666	6.810335
67	4.791367	1.897023	1421	2.525730	0.01165	1.070101	8.512634
68	3.045785	1.790129	1421	1.701433	0.08908	-0.465795	6.557365
71	4.992434	1.839526	1421	2.713978	0.00673	1.383956	8.600913
72	2.881736	1.754825	1421	1.642178	0.10077	-0.560591	6.324062
75	5.193501	1.790507	1421	2.900575	0.00378	1.681180	8.705822
76	2.717686	1.730161	1421	1.570771	0.11646	-0.676257	6.111629
79	5.394568	1.750679	1421	3.081415	0.00210	1.960376	8.828760
80	2.553636	1.716594	1421	1.487618	0.13707	-0.813694	5.920967
83	5.595635	1.720679	1421	3.251993	0.00117	2.220292	8.970978
84	2.389587	1.714389	1421	1.393842	0.16358	-0.973418	5.752591
87	5.796702	1.701027	1421	3.407765	0.00067	2.459907	9.133496
88	2.225537	1.723588	1421	1.291223	0.19684	-1.155514	5.606588
91	5.997769	1.692085	1421	3.544602	0.00041	2.678515	9.317022

92	2.061488	1.744013	1421	1.182037	0.23739	-1.359629	5.482604
95	6.198836	1.694022	1421	3.659241	0.00026	2.875783	9.521889
96	1.897438	1.775275	1421	1.068814	0.28534	-1.585002	5.379879
99	6.399903	1.706801	1421	3.749648	0.00018	3.051782	9.748023
100	1.733389	1.816814	1421	0.954081	0.34020	-1.830537	5.297314
103	6.600970	1.730181	1421	3.815189	0.00014	3.206986	9.994954
104	1.569339	1.867946	1421	0.840142	0.40097	-2.094889	5.233567
107	6.802036	1.763742	1421	3.856594	0.00012	3.342219	10.261854
108	1.405290	1.927907	1421	0.728920	0.46617	-2.376560	5.187139
111	7.003103	1.806915	1421	3.875723	0.00011	3.458596	10.547611
112	1.241240	1.995902	1421	0.621894	0.53411	-2.673991	5.156471
115	7.204170	1.859032	1421	3.875227	0.00011	3.557428	10.850912
116	1.077190	2.071139	1421	0.520096	0.60308	-2.985629	5.140010
119	7.405237	1.919364	1421	3.858173	0.00012	3.640147	11.170328
120	0.913141	2.152860	1421	0.424152	0.67152	-3.309985	5.136267
123	7.606304	1.987162	1421	3.827722	0.00013	3.708218	11.504391
124	0.749091	2.240355	1421	0.334363	0.73816	-3.645668	5.143850
127	7.807371	2.061691	1421	3.786877	0.00016	3.763086	11.851656
128	0.585042	2.332975	1421	0.250771	0.80203	-3.991402	5.161486

131	8.008438	2.142248	1421	3.738334	0.00019	3.806130	12.210746
132	0.420992	2.430133	1421	0.173238	0.86249	-4.346041	5.188025
135	8.209505	2.228179	1421	3.684401	0.00024	3.838631	12.580379
136	0.256943	2.531307	1421	0.101506	0.91916	-4.708557	5.222442
139	8.410572	2.318887	1421	3.626986	0.00030	3.861762	12.959382
140	0.092893	2.636034	1421	0.035240	0.97189	-5.078044	5.263830
143	8.611639	2.413834	1421	3.567619	0.00037	3.876579	13.346699
144	-0.071157	2.743909	1421	-0.025933	0.97931	-5.453704	5.311391
147	8.812706	2.512538	1421	3.507491	0.00047	3.884023	13.741388
148	-0.235206	2.854574	1421	-0.082396	0.93434	-5.834837	5.364425
151	9.013773	2.614575	1421	3.447509	0.00058	3.884931	14.142614
152	-0.399256	2.967716	1421	-0.134533	0.89300	-6.220831	5.422320
155	9.214840	2.719569	1421	3.388345	0.00072	3.880037	14.549642
156	-0.563305	3.083064	1421	-0.182710	0.85505	-6.611151	5.484540
159	9.415906	2.827192	1421	3.330481	0.00089	3.869989	14.961824
160	-0.727355	3.200378	1421	-0.227271	0.82025	-7.005329	5.550619
163	9.616973	2.937153	1421	3.274250	0.00108	3.855352	15.378595
164	-0.891404	3.319451	1421	-0.268540	0.78832	-7.402955	5.620147
167	9.818040	3.049200	1421	3.219874	0.00131	3.836624	15.799457

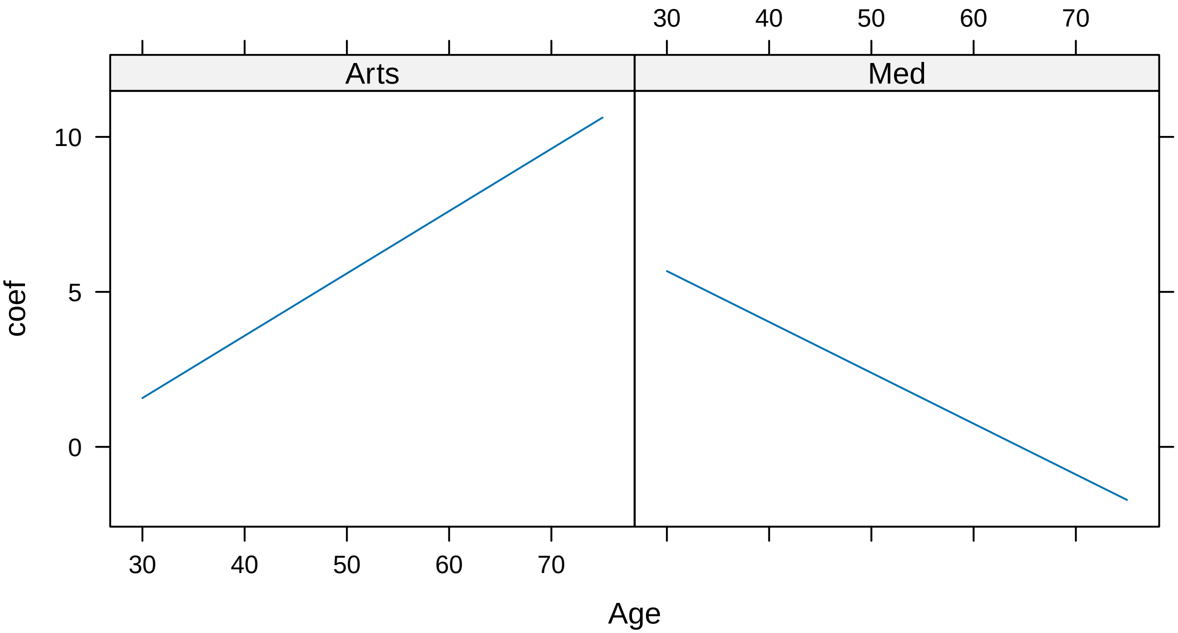
```
168 -1.055454 3.440100  1421 -0.306809 0.75903 -7.803673   5.692765
171 10.019107 3.163111  1421  3.167485 0.00157  3.814238  16.223977
172 -1.219503 3.562164  1421 -0.342349 0.73214 -8.207168   5.768161
175 10.220174 3.278693  1421  3.117149 0.00186  3.788576  16.651772
176 -1.383553 3.685502  1421 -0.375404 0.70742 -8.613163   5.846057
179 10.421241 3.395774  1421  3.068885 0.00219  3.759973  17.082509
180 -1.547603 3.809992  1421 -0.406196 0.68466 -9.021416   5.926211
183 10.622308 3.514204  1421  3.022678 0.00255  3.728722  17.515894
184 -1.711652 3.935524  1421 -0.434924 0.66368 -9.431714   6.008409
```

```
wgap <- as.data.frame(wgap)
head(wgap)
```

```
       coef       se        U2         L2    p-value   t-value   DF Faculty Age
3  1.574296 3.471959  8.518215 -5.3696218 0.65030713 0.4534317 1421    Arts  30
4  5.670578 3.246153 12.162883 -0.8217270 0.08087745 1.7468612 1421     Med  30
7  1.775363 3.353992  8.483348 -4.9326212 0.59666034 0.5293284 1421    Arts  31
8  5.506529 3.128136 11.762801 -0.7497435 0.07856823 1.7603226 1421     Med  31
11 1.976430 3.237426  8.451283 -4.4984220 0.54163203 0.6104943 1421    Arts  32
12 5.342479 3.012002 11.366484 -0.6815258 0.07632194 1.7737300 1421     Med  32
```

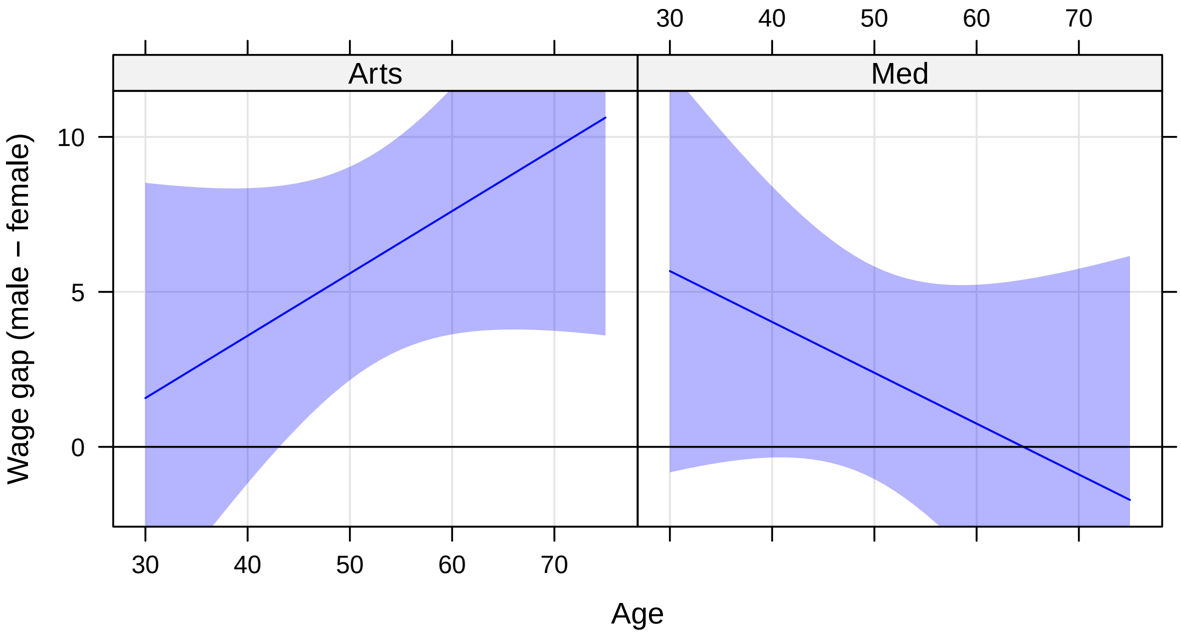
```
   L.(Intercept) L.Age L.FacultyMed L.SexM L.Age:FacultyMed L.Age:SexM
3              0     0            0      1                0         30
4              0     0            0      1                0         30
7              0     0            0      1                0         31
8              0     0            0      1                0         31
11             0     0            0      1                0         32
12             0     0            0      1                0         32
   L.FacultyMed:SexM L.Age:FacultyMed:SexM
3                  0                     0
4                  1                    30
7                  0                     0
8                  1                    31
11                 0                     0
12                 1                    32
```

```
xyplot(coef ~ Age | Faculty, wgap,
       type = 'l', auto.key = list(space='right'))
```



Arts
Med
coef
Age
10
5
0
30
40
50
60
70

```r
xyplot(coef ~ Age | Faculty, wgap,
       type = 'l', auto.key = list(space='right'),
       ylab = 'Wage gap (male - female)',
       fit = wgap$coef,
       lower = wgap$L2,
       upper = wgap$U2,
       subscripts = TRUE) +
  layer(panel.band(...)) +
  layer_(panel.grid(h=-1,v=-1)) +
  layer(panel.abline(h=0))
```



Arts
Med
Wage gap (male − female)
Age
10
5
0
30
40
50
60
70

4 Models with heterosckedasticity

```
fitconpower <- gls(Base ~ Age * Faculty * Sex, dd,
                   weights = varConstPower(form = ~fitted(.)|Faculty))
summary(fitconpower)
```

```
Generalized least squares fit by REML
  Model: Base ~ Age * Faculty * Sex
  Data: dd
       AIC      BIC    logLik
  12640.27 12708.65 -6307.134

Variance function:
 Structure: Constant plus power of variance covariate, different strata
 Formula: ~fitted(.) | Faculty
 Parameter estimates:
              Arts          Med
const 3.873281e+05 7.198900e+05
power 2.253126e+00 2.340018e+00
```

```
Coefficients:
                          Value Std.Error  t-value p-value
(Intercept)          -17.882990  3.271677 -5.46600  0.0000
Age                    3.952782  0.065876 60.00346  0.0000
FacultyMed            -9.512029  6.663510 -1.42748  0.1537
SexM                  -8.653431  4.516944 -1.91577  0.0556
Age:FacultyMed         0.837313  0.142579  5.87262  0.0000
Age:SexM               0.285478  0.090130  3.16739  0.0016
FacultyMed:SexM       12.571011  9.590554  1.31077  0.1901
Age:FacultyMed:SexM   -0.305644  0.200665 -1.52316  0.1279

 Correlation:
                     (Intr) Age    FcltyM SexM   Ag:FcM Ag:SxM FcM:SM
Age                  -0.974
FacultyMed           -0.491  0.478
SexM                 -0.724  0.706  0.356
Age:FacultyMed        0.450 -0.462 -0.971 -0.326
Age:SexM              0.712 -0.731 -0.350 -0.972  0.338
```

```
FacultyMed:SexM      0.341 -0.332 -0.695 -0.471  0.675  0.458
Age:FacultyMed:SexM -0.320  0.328  0.690  0.437 -0.711 -0.449 -0.971

Standardized residuals:
          Min            Q1           Med            Q3           Max
-3.0328923778 -0.6682578337 -0.0008458875  0.6744181513  3.2478019086

Residual standard error: 2.713027e-05
Degrees of freedom: 1430 total; 1422 residual
```

```r
fitpower <- update(fit, weights = varPower(form = ~fitted(.)|Faculty))
fitgroups <- update(fit, weights = varIdent(form = ~ 1 | Faculty))
anova(fit , fitgroups, fitpower, fitconpower)
```

```
           Model df      AIC      BIC    logLik   Test   L.Ratio p-value
fit            1  9 12994.50 13041.83 -6488.248
fitgroups      2 10 12717.15 12769.74 -6348.573 1 vs 2 279.35040  <.0001
fitpower       3 11 12642.22 12700.08 -6310.110 2 vs 3  76.92543  <.0001
fitconpower    4 13 12640.27 12708.65 -6307.134 3 vs 4   5.95226   0.051
```

```r
library(car)
```

```
Loading required package: carData
```

```r
compareCoefs(fit, fitgroups, fitpower, fitconpower)
```

```
Calls:
1: gls(model = Base ~ Age * Faculty * Sex, data = dd)
2: gls(model = Base ~ Age * Faculty * Sex, data = dd, weights =
  varIdent(form = ~1 | Faculty))
3: gls(model = Base ~ Age * Faculty * Sex, data = dd, weights =
  varPower(form = ~fitted(.) | Faculty))
4: gls(model = Base ~ Age * Faculty * Sex, data = dd, weights =
  varConstPower(form = ~fitted(.) | Faculty))

                    Model 1 Model 2 Model 3 Model 4
(Intercept)          -19.77  -19.77  -17.74  -17.88
SE                     5.33    3.51    3.18    3.27

Age                  3.9913  3.9913  3.9518  3.9528
```

SE	0.1008	0.0663	0.0640	0.0659
FacultyMed	-12.21	-12.21	-9.94	-9.51
SE	7.31	7.20	6.39	6.66
SexM	-4.46	-4.46	-8.40	-8.65
SE	7.31	4.81	4.37	4.52
Age:FacultyMed	0.899	0.899	0.849	0.837
SE	0.142	0.143	0.138	0.143
Age:SexM	0.2011	0.2011	0.2764	0.2855
SE	0.1358	0.0894	0.0868	0.0901
FacultyMed:SexM	15.05	15.05	12.72	12.57
SE	10.23	10.20	9.21	9.59
Age:FacultyMed:SexM	-0.365	-0.365	-0.312	-0.306
SE	0.195	0.197	0.193	0.201

5 Revisiting the gap

```
wgap2 <- wald(fitconpower,
             Lgap,
             data = subset(ww, Sex == 'F', select = c(Faculty,Age)))
wgap2
```

```
    numDF denDF  F-value p-value
1       4  1421 9.246016 <.00001
      Estimate  Std.Error DF   t-value   p-value Lower 0.95 Upper 0.95
3    -0.089104 1.991198  1421 -0.044749 0.96431 -3.995107   3.816899
4     3.312579 3.483557  1421  0.950919 0.34181 -3.520887  10.146046
7     0.196373 1.915402  1421  0.102523 0.91836 -3.560946   3.953693
8     3.292412 3.339197  1421  0.985989 0.32431 -3.257873   9.842698
11    0.481851 1.840899  1421  0.261748 0.79355 -3.129320   4.093022
12    3.272246 3.198373  1421  1.023097 0.30644 -3.001794   9.546286
```

15	0.767329	1.767852	1421	0.434046	0.66432	-2.700551	4.235208
16	3.252079	3.061573	1421	1.062225	0.28831	-2.753610	9.257768
19	1.052806	1.696448	1421	0.620594	0.53497	-2.275005	4.380618
20	3.231912	2.929361	1421	1.103282	0.27009	-2.514424	8.978249
23	1.338284	1.626905	1421	0.822595	0.41088	-1.853110	4.529677
24	3.211746	2.802386	1421	1.146075	0.25196	-2.285512	8.709003
27	1.623761	1.559471	1421	1.041225	0.29795	-1.435352	4.682875
28	3.191579	2.681392	1421	1.190269	0.23414	-2.068333	8.451491
31	1.909239	1.494433	1421	1.277568	0.20161	-1.022292	4.840770
32	3.171412	2.567225	1421	1.235346	0.21691	-1.864546	8.207371
35	2.194716	1.432115	1421	1.532500	0.12562	-0.614570	5.004003
36	3.151246	2.460836	1421	1.280559	0.20056	-1.676016	7.978507
39	2.480194	1.372889	1421	1.806550	0.07104	-0.212914	5.173302
40	3.131079	2.363275	1421	1.324890	0.18542	-1.504803	7.766961
43	2.765672	1.317173	1421	2.099703	0.03593	0.181860	5.349483
44	3.110912	2.275677	1421	1.367027	0.17183	-1.353135	7.574959
47	3.051149	1.265428	1421	2.411159	0.01603	0.568841	5.533458
48	3.090746	2.199234	1421	1.405374	0.16013	-1.223349	7.404840
51	3.336627	1.218163	1421	2.739064	0.00624	0.947035	5.726218

52	3.070579	2.135144	1421	1.438113	0.15062	-1.117795	7.258952
55	3.622104	1.175917	1421	3.080238	0.00211	1.315384	5.928824
56	3.050412	2.084547	1421	1.463345	0.14359	-1.038708	7.139533
59	3.907582	1.139248	1421	3.429965	0.00062	1.672792	6.142371
60	3.030245	2.048443	1421	1.479292	0.13928	-0.988052	7.048543
63	4.193059	1.108711	1421	3.781923	0.00016	2.018173	6.367945
64	3.010079	2.027606	1421	1.484548	0.13789	-0.967344	6.987501
67	4.478537	1.084823	1421	4.128359	0.00004	2.350511	6.606563
68	2.989912	2.022508	1421	1.478319	0.13954	-0.977510	6.957335
71	4.764015	1.068029	1421	4.460565	0.00001	2.668931	6.859098
72	2.969745	2.033268	1421	1.460578	0.14435	-1.018784	6.958274
75	5.049492	1.058669	1421	4.769660	<.00001	2.972770	7.126215
76	2.949579	2.059636	1421	1.432087	0.15234	-1.090676	6.989833
79	5.334970	1.056940	1421	5.047563	<.00001	3.261640	7.408299
80	2.929412	2.101027	1421	1.394277	0.16345	-1.192035	7.050859
83	5.620447	1.062878	1421	5.287953	<.00001	3.535469	7.705425
84	2.909245	2.156573	1421	1.349013	0.17755	-1.321164	7.139654
87	5.905925	1.076356	1421	5.486960	<.00001	3.794506	8.017343
88	2.889079	2.225217	1421	1.298336	0.19438	-1.475984	7.254141

91	6.191402	1.097098	1421	5.643435	<.00001	4.039296	8.343509
92	2.868912	2.305788	1421	1.244222	0.21362	-1.654201	7.392025
95	6.476880	1.124701	1421	5.758756	<.00001	4.270627	8.683133
96	2.848745	2.397083	1421	1.188421	0.23487	-1.853457	7.550948
99	6.762357	1.158675	1421	5.836284	<.00001	4.489460	9.035255
100	2.828579	2.497929	1421	1.132370	0.25767	-2.071445	7.728602
103	7.047835	1.198478	1421	5.880653	<.00001	4.696858	9.398812
104	2.808412	2.607215	1421	1.077169	0.28159	-2.305992	7.922816
107	7.333313	1.243551	1421	5.897074	<.00001	4.893920	9.772706
108	2.788245	2.723927	1421	1.023612	0.30619	-2.555106	8.131596
111	7.618790	1.293343	1421	5.890775	<.00001	5.081724	10.155856
112	2.768079	2.847152	1421	0.972227	0.33110	-2.816994	8.353151
115	7.904268	1.347330	1421	5.866617	<.00001	5.261299	10.547237
116	2.747912	2.976081	1421	0.923332	0.35599	-3.090072	8.585895
119	8.189745	1.405029	1421	5.828879	<.00001	5.433591	10.945900
120	2.727745	3.110004	1421	0.877087	0.38059	-3.372946	8.828437
123	8.475223	1.466003	1421	5.781178	<.00001	5.599461	11.350985
124	2.707578	3.248303	1421	0.833536	0.40468	-3.664407	9.079563
127	8.760700	1.529859	1421	5.726477	<.00001	5.759676	11.761724

128	2.687412	3.390444	1421	0.792643	0.42812	-3.963402	9.338225
131	9.046178	1.596251	1421	5.667139	<.00001	5.914916	12.177440
132	2.667245	3.535963	1421	0.754319	0.45078	-4.269024	9.603514
135	9.331656	1.664877	1421	5.605012	<.00001	6.065775	12.597536
136	2.647078	3.684460	1421	0.718444	0.47260	-4.580487	9.874643
139	9.617133	1.735471	1421	5.541512	<.00001	6.212773	13.021493
140	2.626912	3.835589	1421	0.684878	0.49353	-4.897112	10.150936
143	9.902611	1.807803	1421	5.477706	<.00001	6.356362	13.448859
144	2.606745	3.989050	1421	0.653475	0.51356	-5.218314	10.431804
147	10.188088	1.881672	1421	5.414381	<.00001	6.496936	13.879241
148	2.586578	4.144585	1421	0.624086	0.53267	-5.543584	10.716741
151	10.473566	1.956904	1421	5.352110	<.00001	6.634835	14.312297
152	2.566412	4.301969	1421	0.596567	0.55089	-5.872481	11.005304
155	10.759043	2.033348	1421	5.291294	<.00001	6.770357	14.747730
156	2.546245	4.461006	1421	0.570778	0.56824	-6.204620	11.297110
159	11.044521	2.110873	1421	5.232206	<.00001	6.903760	15.185282
160	2.526078	4.621525	1421	0.546590	0.58475	-6.539667	11.591824
163	11.329999	2.189363	1421	5.175021	<.00001	7.035268	15.624729
164	2.505912	4.783378	1421	0.523879	0.60044	-6.877330	11.889153

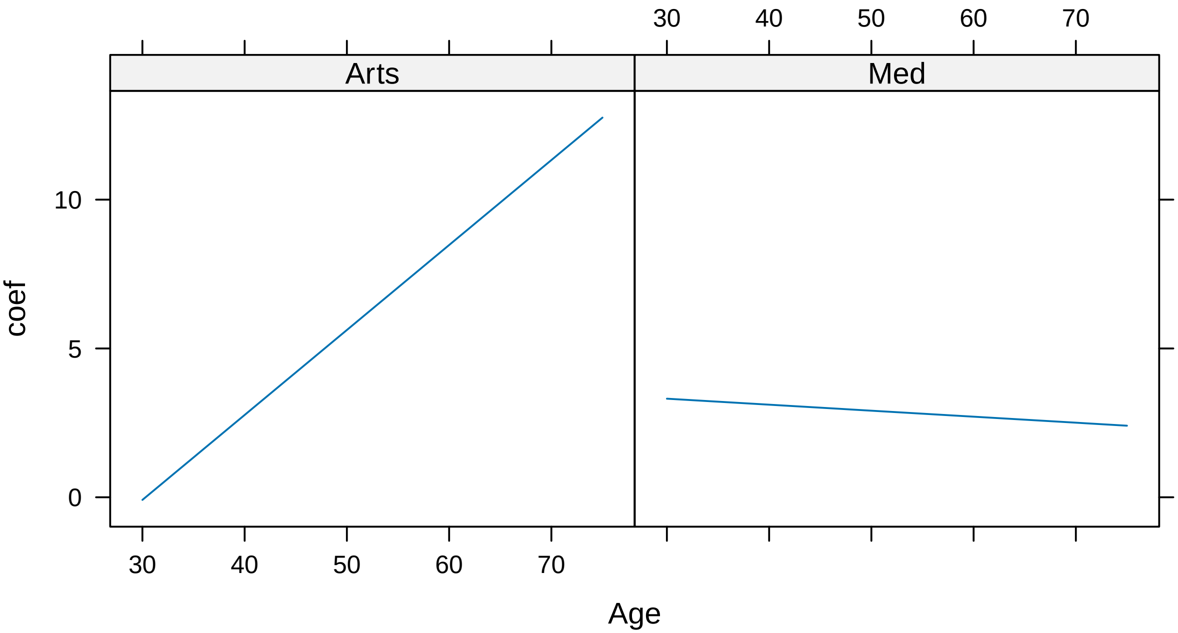
```
167 11.615476 2.268718  1421  5.119841 <.00001  7.165080  16.065873
168  2.485745 4.946433  1421  0.502533 0.61537 -7.217351  12.188840
171 11.900954 2.348851  1421  5.066713 <.00001  7.293366  16.508542
172  2.465578 5.110575  1421  0.482446 0.62956 -7.559504  12.490661
175 12.186431 2.429685  1421  5.015643 <.00001  7.420277  16.952585
176  2.445411 5.275703  1421  0.463523 0.64306 -7.903592  12.794415
179 12.471909 2.511151  1421  4.966610 <.00001  7.545947  17.397871
180  2.425245 5.441727  1421  0.445676 0.65590 -8.249437  13.099927
183 12.757386 2.593191  1421  4.919570 <.00001  7.670492  17.844281
184  2.405078 5.608568  1421  0.428822 0.66812 -8.596884  13.407040
```

```r
wgap2 <- as.data.frame(wgap2)
head(wgap2)
```

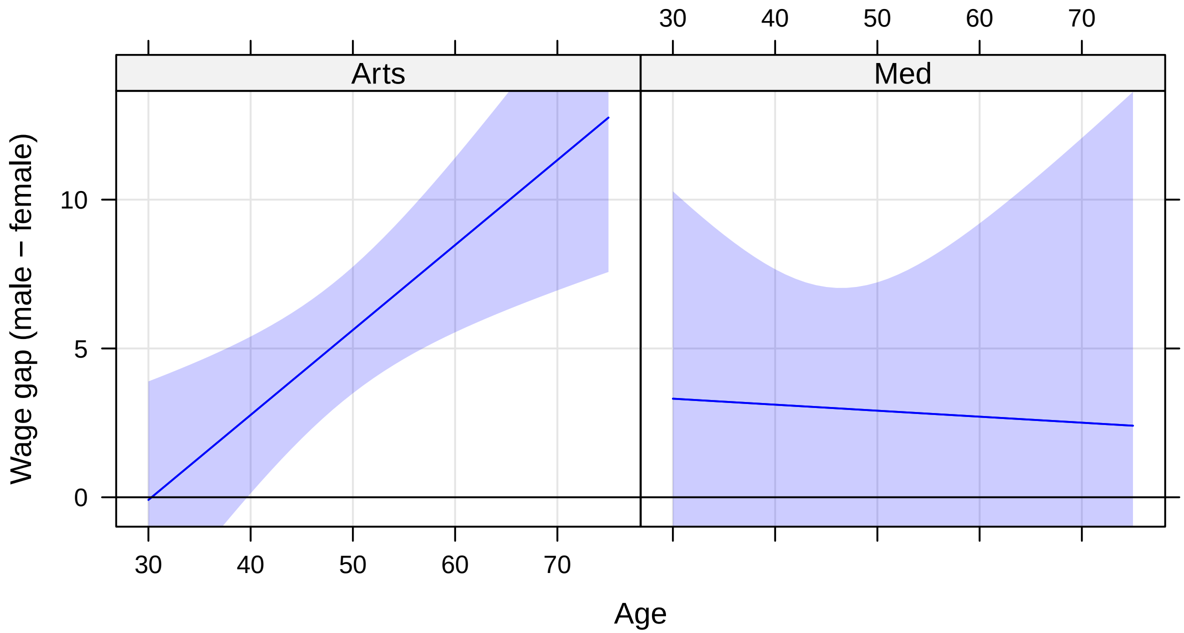
```
          coef       se        U2        L2   p-value     t-value   DF Faculty
3  -0.08910413 1.991198  3.893292 -4.071500 0.9643137 -0.04474901 1421    Arts
4   3.31257910 3.483557 10.279693 -3.654535 0.3418074  0.95091862 1421     Med
7   0.19637344 1.915402  4.027178 -3.634431 0.9183557  0.10252335 1421    Arts
8   3.29241241 3.339197  9.970806 -3.385982 0.3243062  0.98598926 1421     Med
11  0.48185101 1.840899  4.163649 -3.199947 0.7935539  0.26174768 1421    Arts
```

```
12  3.27224572 3.198373  9.668992 -3.124501 0.3064362  1.02309692 1421     Med
   Age L.(Intercept) L.Age L.FacultyMed L.SexM L.Age:FacultyMed L.Age:SexM
3   30             0     0            0      1                0         30
4   30             0     0            0      1                0         30
7   31             0     0            0      1                0         31
8   31             0     0            0      1                0         31
11  32             0     0            0      1                0         32
12  32             0     0            0      1                0         32
   L.FacultyMed:SexM L.Age:FacultyMed:SexM
3                  0                     0
4                  1                    30
7                  0                     0
8                  1                    31
11                 0                     0
12                 1                    32
```

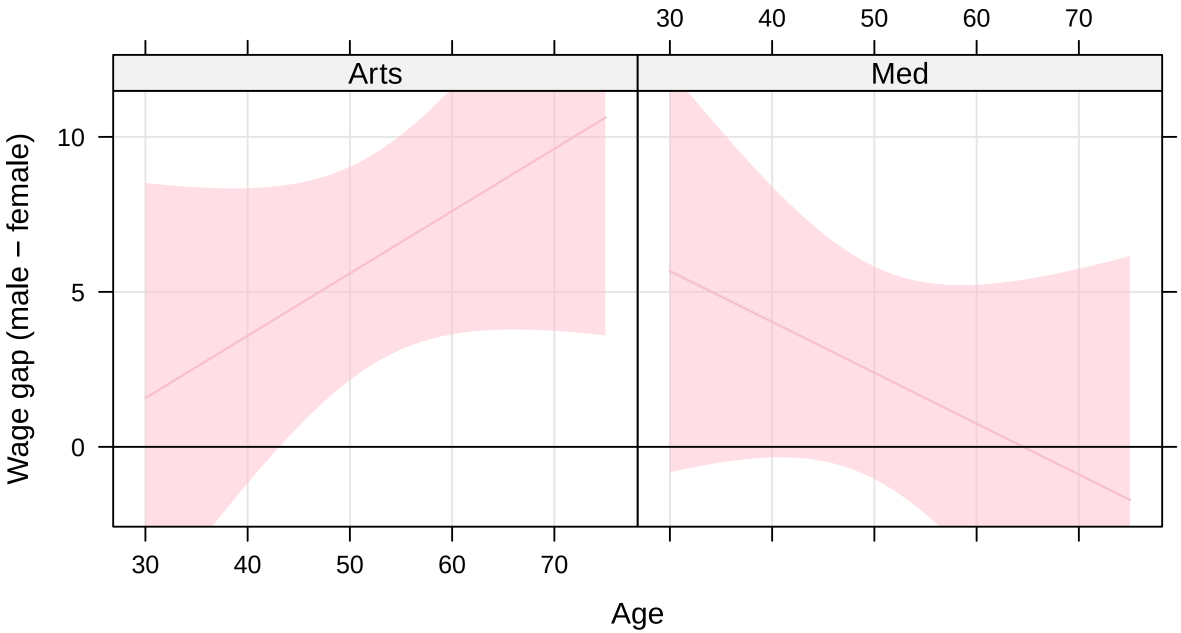
```r
xyplot(coef ~ Age | Faculty, wgap2,
       type = 'l', auto.key = list(space='right'))
```


Arts
Med
coef
Age
0
5
10
30
40
50
60
70

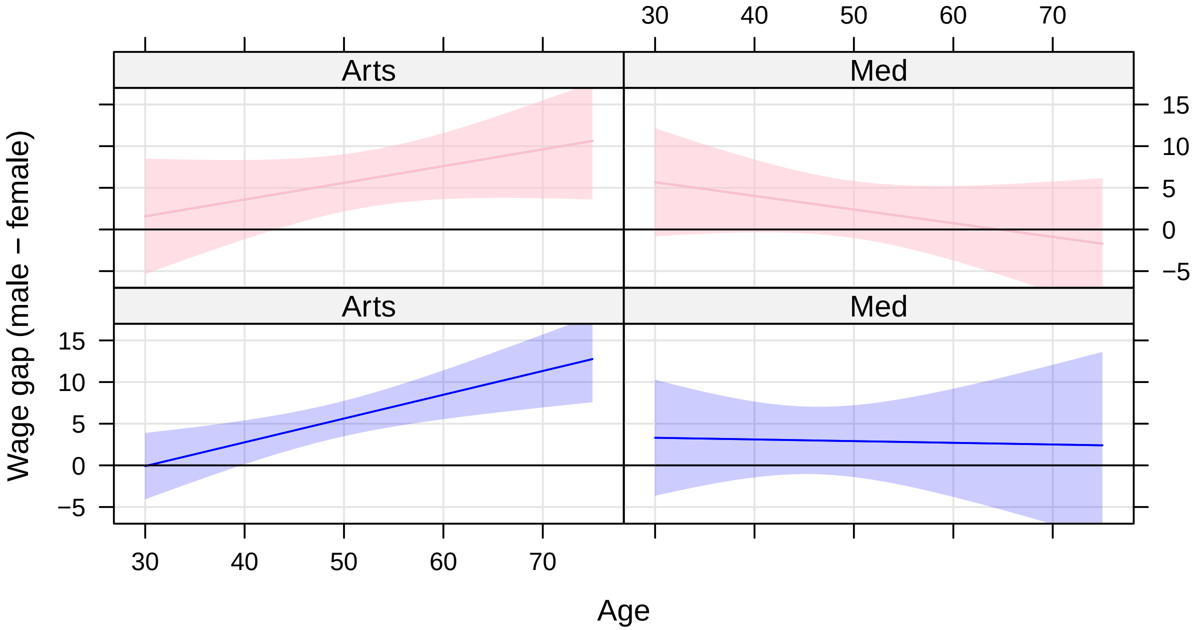
```r
xyplot(coef ~ Age | Faculty, wgap2,
       type = 'l', auto.key = list(space='right'),
       ylab = 'Wage gap (male - female)',
       fit = wgap2$coef,
       lower = wgap2$L2,
       upper = wgap2$U2,
       subscripts = TRUE) +
  layer(panel.fit(..., alpha = .2)) +
  layer(panel.abline(h=0)) +
  layer_(panel.grid(h=-1,v=-1)) -> plhet
plhet
```


Arts
Med
Wage gap (male – female)
0
5
10
30
40
50
60
70
Age

```r
xyplot(coef ~ Age | Faculty, wgap,
       type = 'l', auto.key = list(space='right'),
       ylab = 'Wage gap (male - female)',
       fit = wgap$coef,
       lower = wgap$L2,
       upper = wgap$U2,
       subscripts = TRUE) +
  layer(panel.fit(..., col = 'pink',alpha=.5)) +
  layer(panel.abline(h=0)) +
  layer_(panel.grid(h=-1,v=-1)) -> plnohet
plnohet
```


Arts
Med
Wage gap (male – female)
Age
0
5
10
30
40
50
60
70

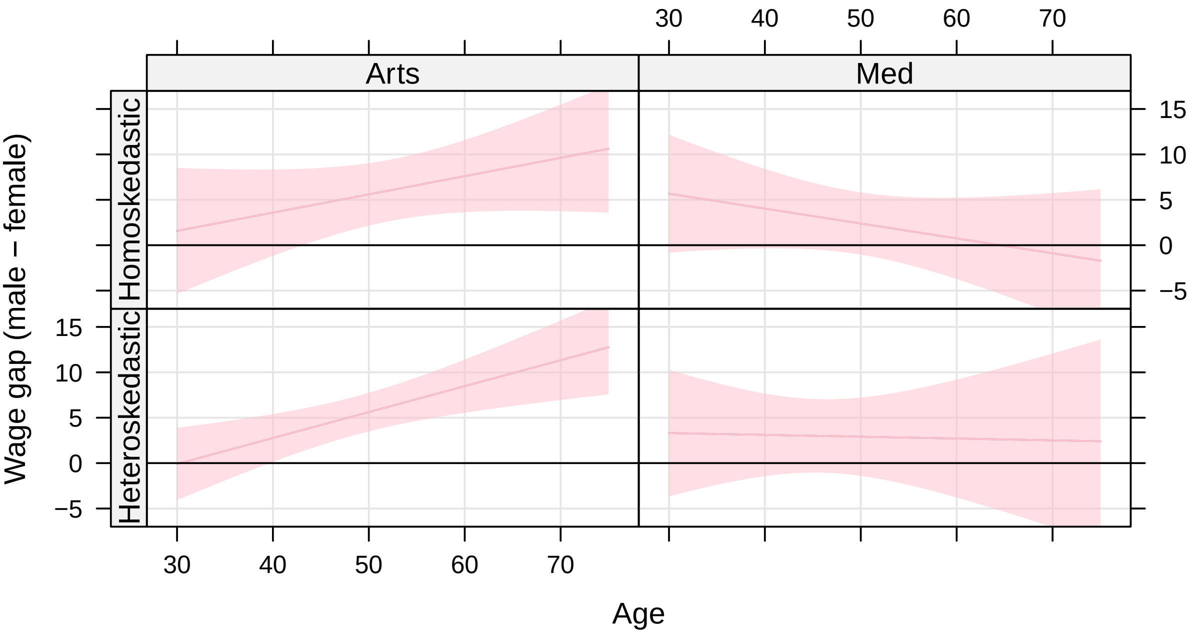
```r
ylim <- c(-7,17)
c(
  update(plhet, ylim = ylim), update(plnohet, ylim = ylim)
)
```



Arts
Med
Arts
Med
Wage gap (male – female)
Age
30
40
50
60
70
15
10
5
0
−5

Combining data frames

```
wgap$type <- 'Homoskedastic'
wgap2$type <- 'Heteroskedastic'

wgap_combined <- rbind(wgap, wgap2)

xyplot(coef ~ Age | Faculty * type, wgap_combined,
       type = 'l', auto.key = list(space='right'),
       ylab = 'Wage gap (male - female)',
       ylim = ylim,
       fit = wgap_combined$coef,
       lower = wgap_combined$L2,
       upper = wgap_combined$U2,
       subscripts = TRUE) +
  layer(panel.fit(..., col = 'pink',alpha=.5)) +
  layer(panel.abline(h=0)) +
  layer_(panel.grid(h=-1,v=-1)) -> cplot
useOuterStrips(cplot)
```



Arts
Med
Homoskedastic
Heteroskedastic
Wage gap (male – female)
Age
30
40
50
60
70
15
10
5
0
−5

Question: Where are the bands wider and where are they narrower when incorporating heteroskedasticity in the model? Do the patterns you see make sense? Note the blue bands use heteroskedasticity and the pink ones don't.