Dealing with Heteroskedasticity

R-side Variance Modeling

2022-04-03

Contents

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

```
Attaching package: 'spida2'
```

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

    getData
```

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

1 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

```r
dd <- expand.grid(Faculty = c("Arts", "Med"), Sex = c("F","M"), n = 1:400)
set.seed(1233)
dd <- within(
  dd,
  {
    Age <- 45 + 5 * (Faculty == "Arts") + 5 * (Sex == "M") + 15 * rnorm(n)
    ..esal <- 100 + 20 * (Faculty == 'Med') +
      (4 + .3 *(Sex == "M") + .5 * (Faculty == "Med")) * (Age - 30)
    ..sdsal <- 10 + 10 * (Faculty == "Med") + .2 * (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    35   32    67
  Med     50   32    82
  Total   85   64   149

, , keep = TRUE

       Sex
Faculty    F    M Total
  Arts   365  368   733
  Med    350  368   718
  Total  715  736  1451

, , keep = Total

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

```
        Med    400  400   800
        Total  800  800  1600
```

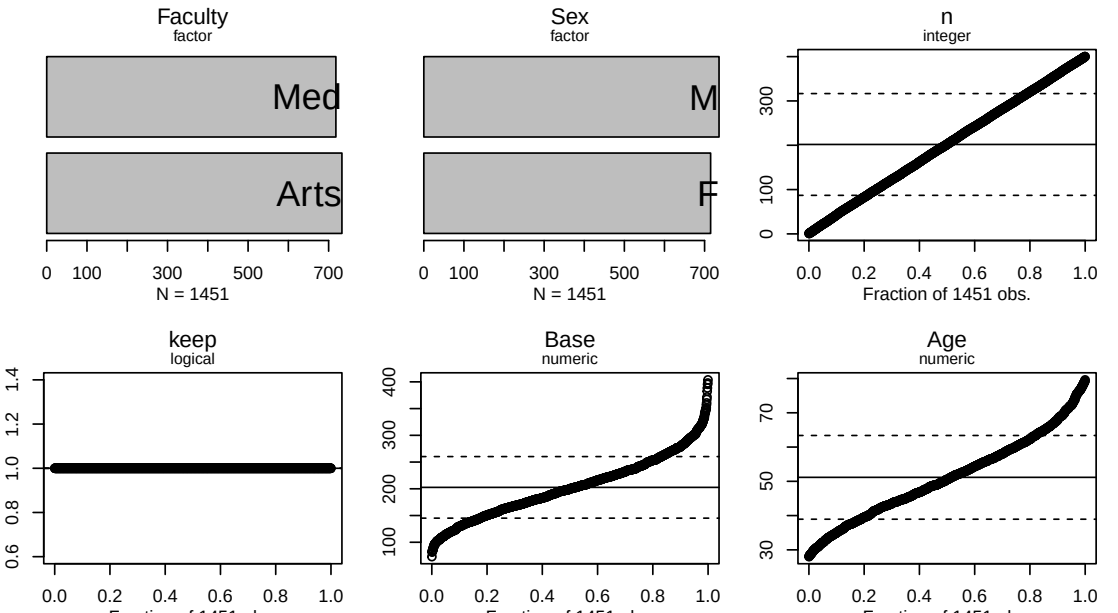
```
dd <- subset(dd, keep)
save(dd, file = 'salary.rda')
```

2 Analysis

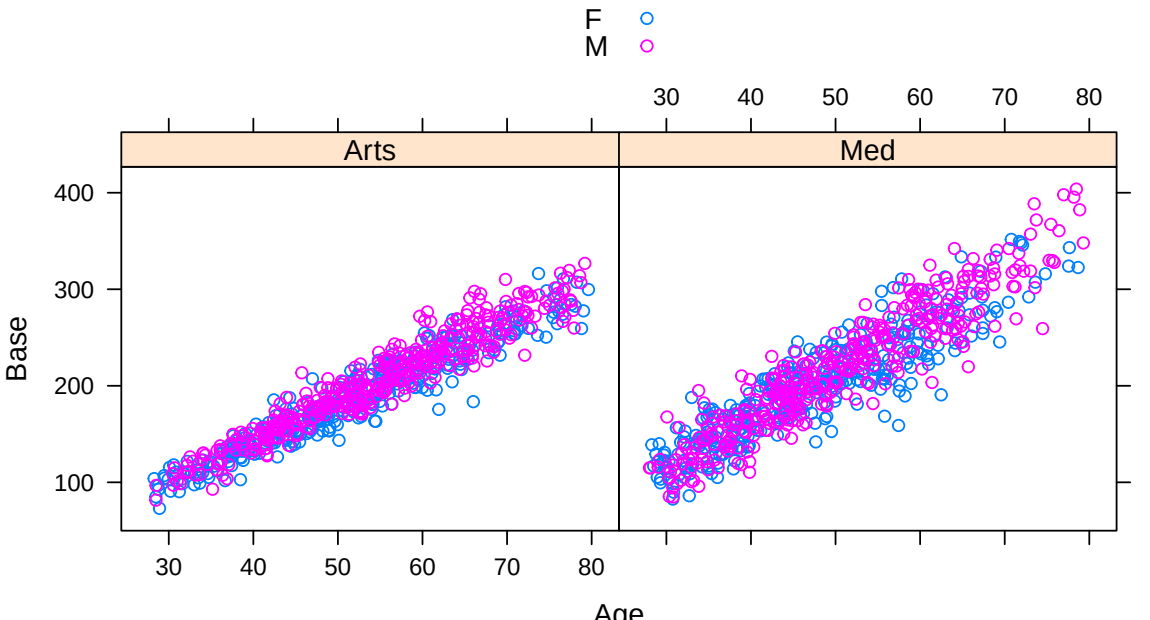
```
load('salary.rda', verbose = TRUE)
```

```
      Loading objects:
        dd
```

```
xqplot(dd)
```


Faculty
factor
Med
Arts
0 100 300 500 700
N = 1451
Sex
factor
M
F
0 100 300 500 700
N = 1451
n
integer
0 100 300
0.0 0.2 0.4 0.6 0.8 1.0
Fraction of 1451 obs.
keep
logical
0.6 0.8 1.0 1.2 1.4
0.0 0.2 0.4 0.6 0.8 1.0
Fraction of 1451 obs.
Base
numeric
100 200 300 400
0.0 0.2 0.4 0.6 0.8 1.0
Fraction of 1451 obs.
Age
numeric
30 50 70
0.0 0.2 0.4 0.6 0.8 1.0
Fraction of 1451 obs.

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

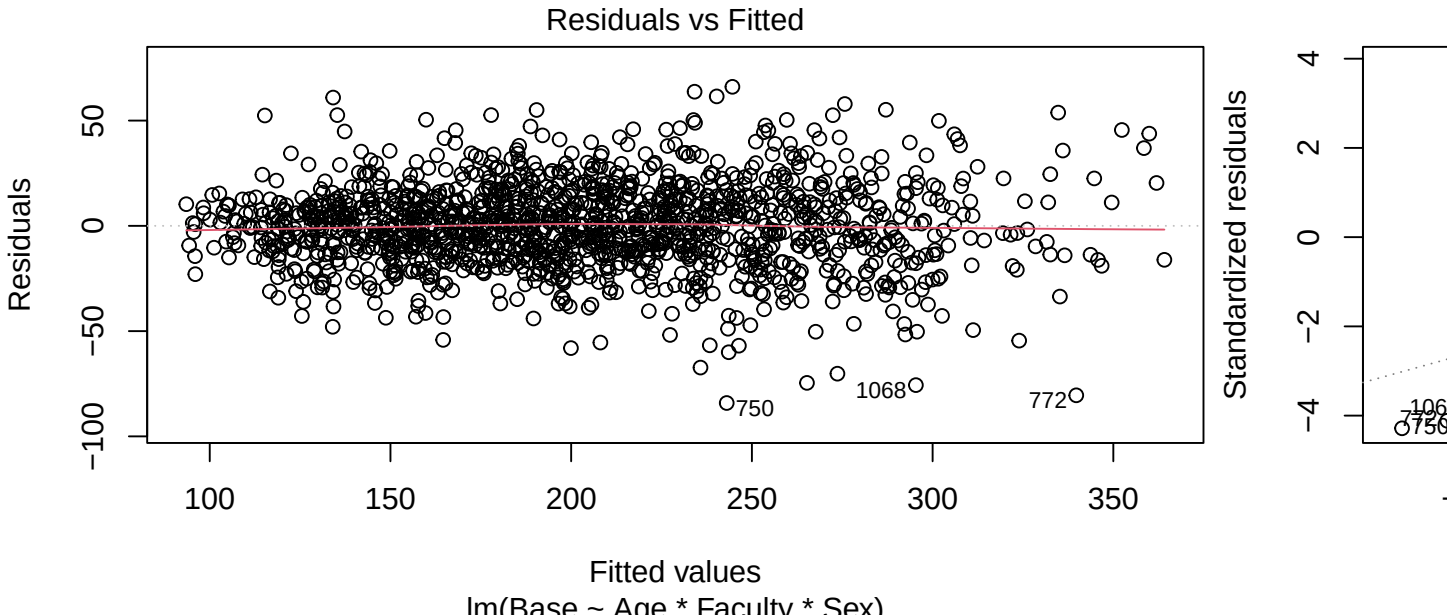

F
M
Arts
Med
Base
Age
100
200
300
400
30
40
50
60
70
80

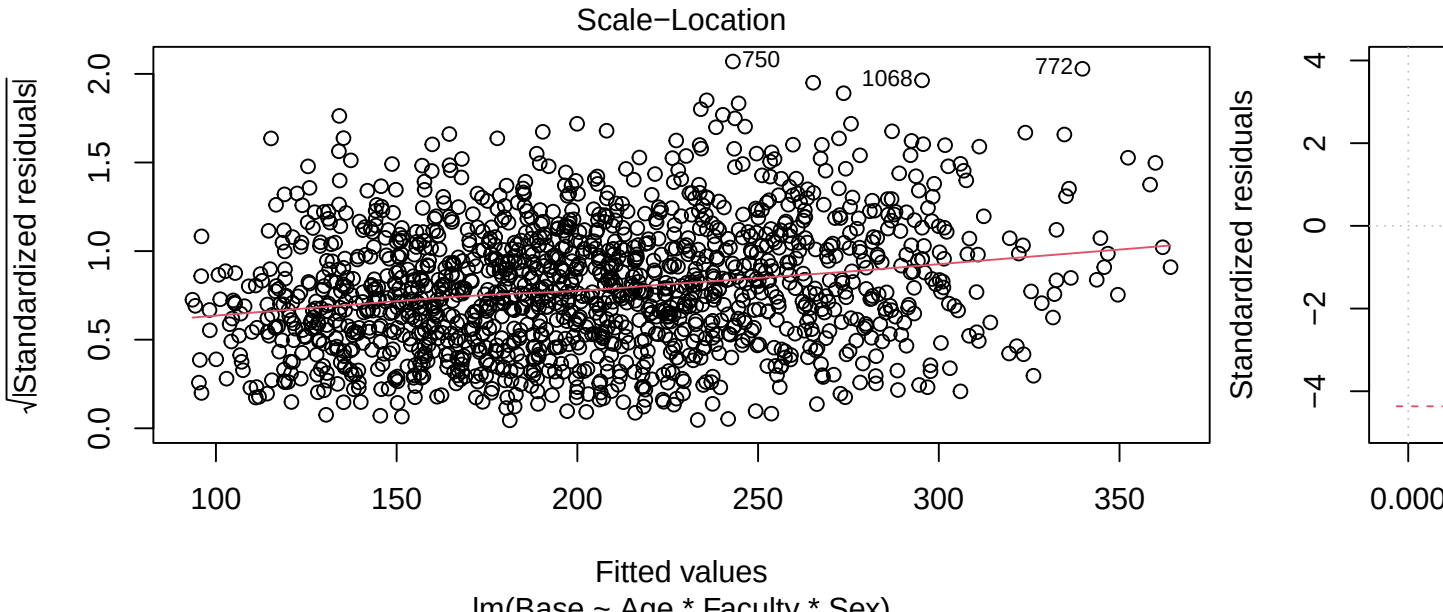
```r
fit <- lm(Base ~ Age * Faculty * Sex, dd)
summary(fit)
```

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

Residuals:
    Min      1Q  Median      3Q     Max
-84.147 -11.606   0.375  12.001  66.045

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)        -18.96213    4.43844  -4.272 2.06e-05 ***
Age                  3.97750    0.08367  47.537  < 2e-16 ***
FacultyMed           8.94979    6.35381   1.409 0.159178
SexM                -4.18547    6.52812  -0.641 0.521531
Age:FacultyMed       0.42701    0.12518   3.411 0.000665 ***
Age:SexM             0.20294    0.11953   1.698 0.089762 .
```

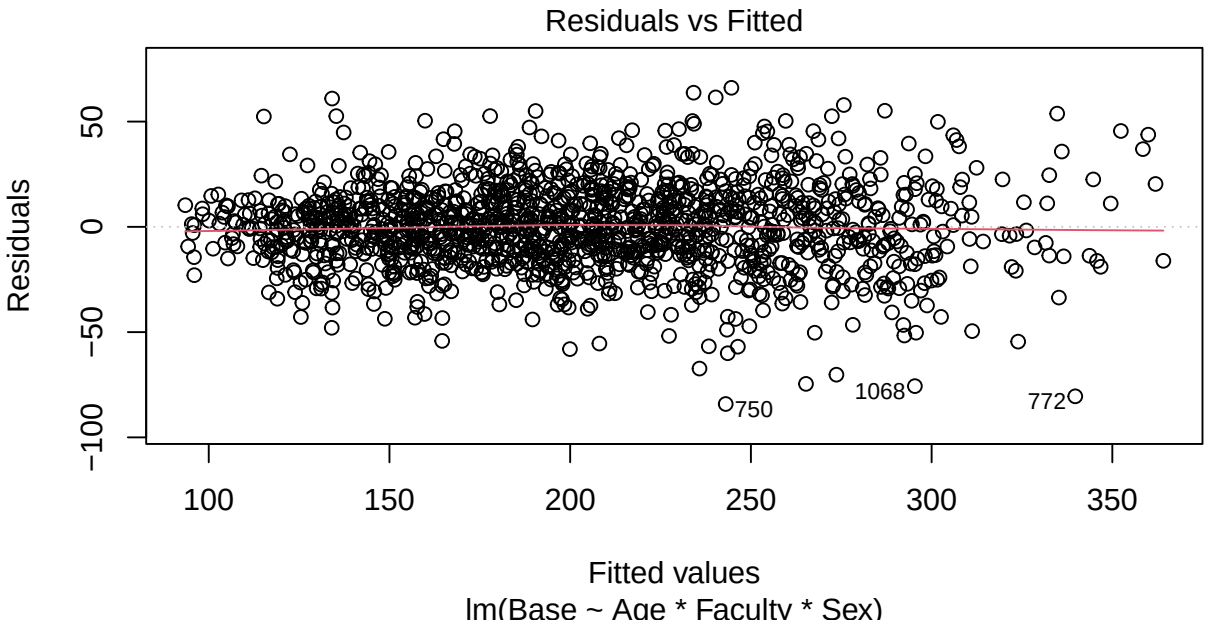
```
FacultyMed:SexM     -22.58948    9.10350  -2.481 0.013200 *
Age:FacultyMed:SexM   0.44828    0.17377   2.580 0.009986 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 19.69 on 1443 degrees of freedom
Multiple R-squared:  0.8836,  Adjusted R-squared:  0.883
F-statistic:  1565 on 7 and 1443 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))
}
```

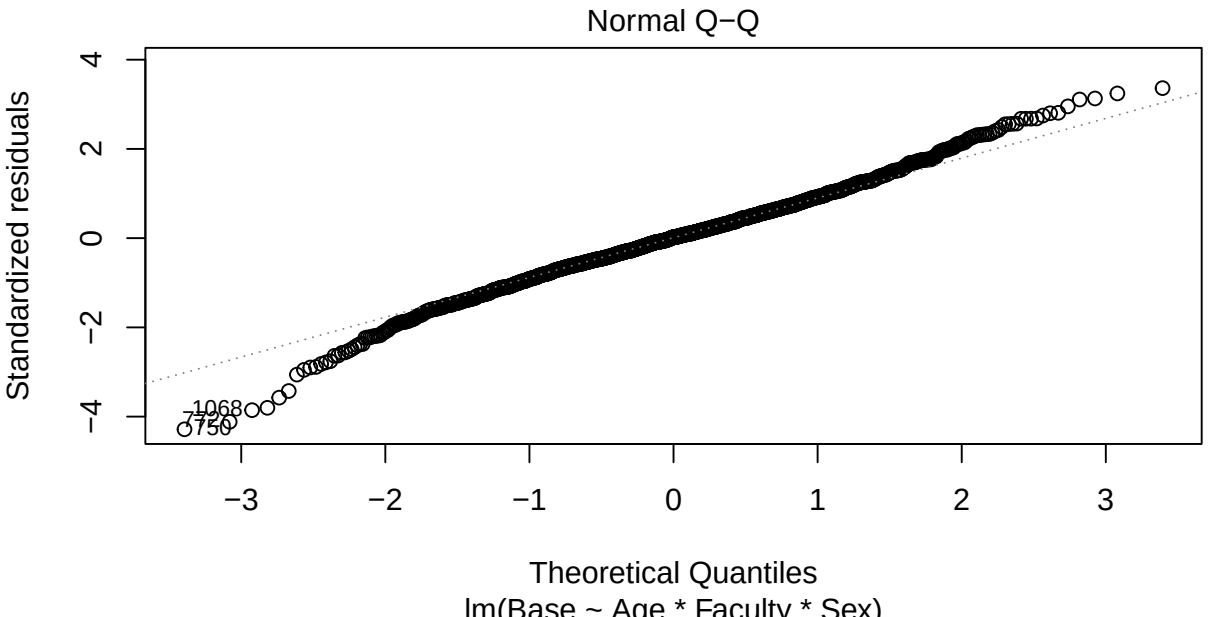

Residuals vs Fitted
Residuals
Fitted values
lm(Base ~ Age * Faculty * Sex)
750
1068
772
Standardized residuals


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
750
1068
772
100
150
200
250
300
350
Fitted values
lm(Base ~ Age * Faculty * Sex)
Standardized residuals
−4
−2
0
2
4
0.000

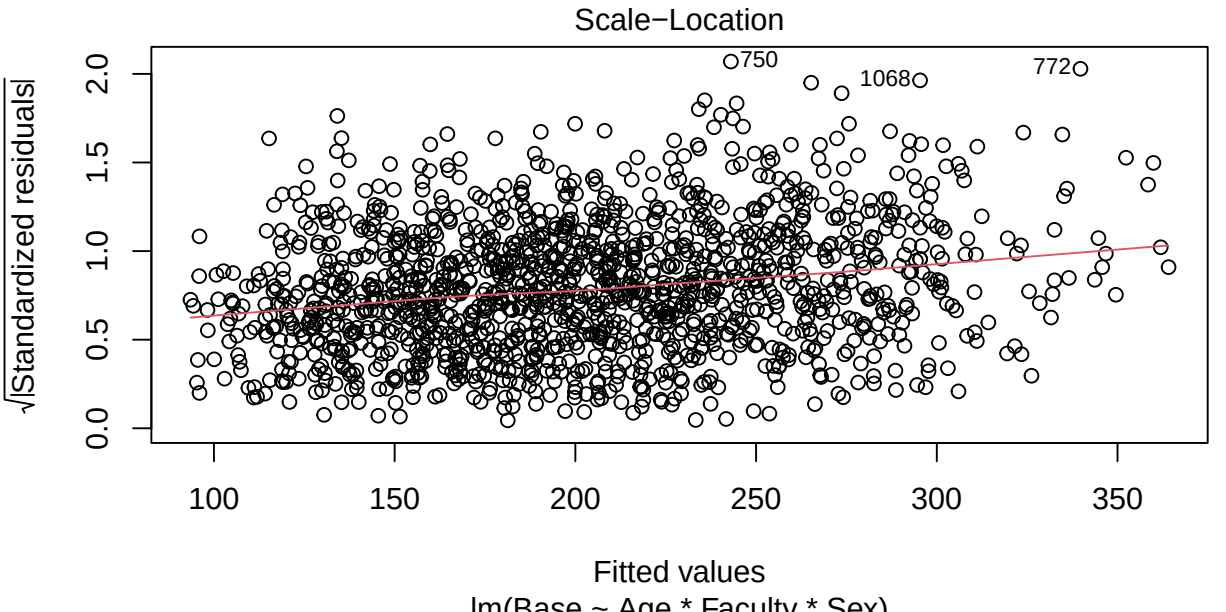
```
plot(fit, 1)
```


Residuals vs Fitted
Residuals
50
0
−50
−100
100
150
200
250
300
350
750
1068
772
Fitted values
lm(Base ~ Age * Faculty * Sex)

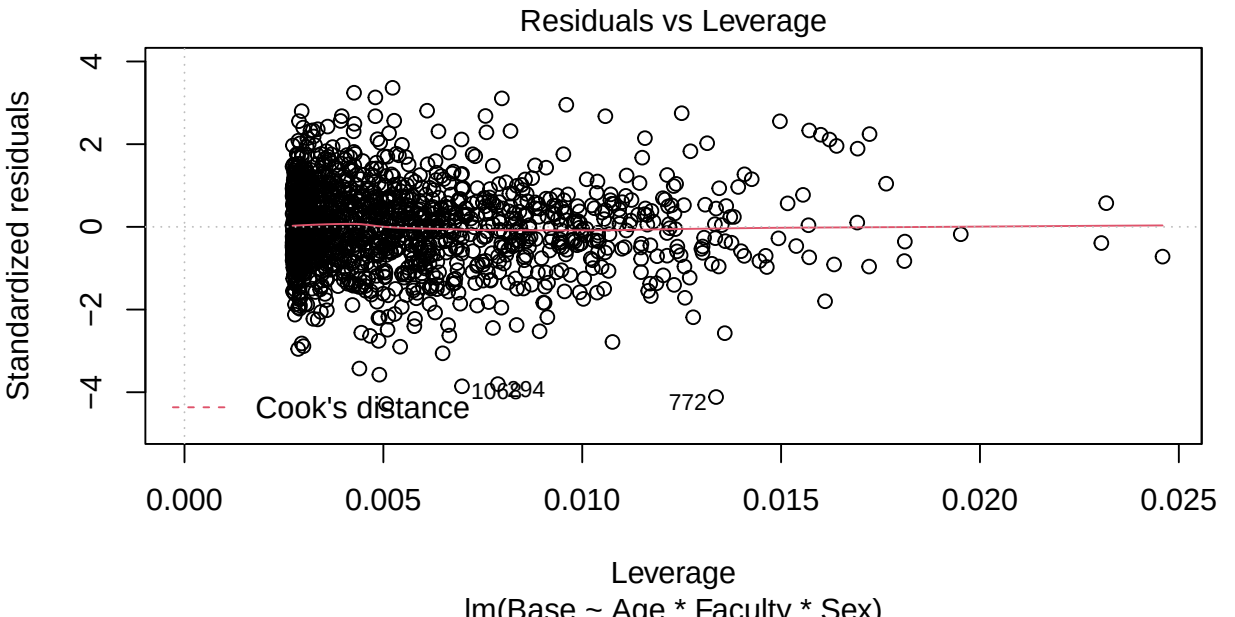
```
plot(fit, 2)
```


Normal Q−Q
Standardized residuals
4
2
0
−2
−4
1068
750
−3
−2
−1
0
1
2
3
Theoretical Quantiles
lm(Base ~ Age * Faculty * Sex)

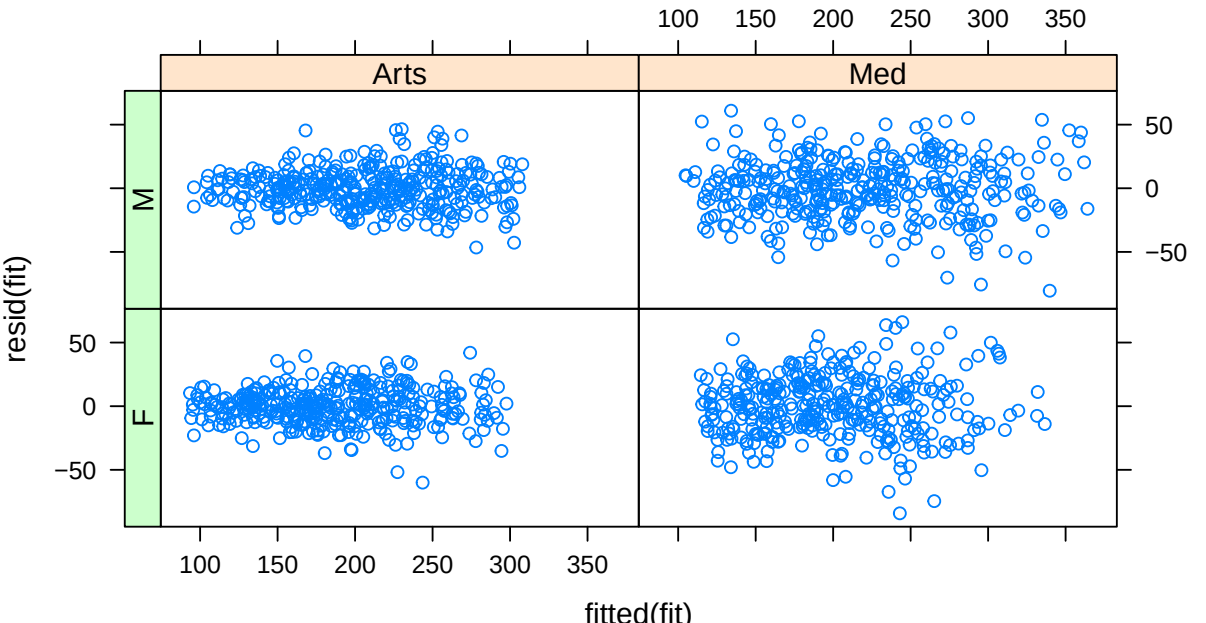
```
plot(fit, 3)
```


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
750
1068
772
100
150
200
250
300
350
Fitted values
lm(Base ~ Age * Faculty * Sex)

```
plot(fit, 5)
```


Residuals vs Leverage
Standardized residuals
4
2
0
−2
−4
Cook's distance
772
0.000
0.005
0.010
0.015
0.020
0.025
Leverage
lm(Base ~ Age * Faculty * Sex)

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


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

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
```

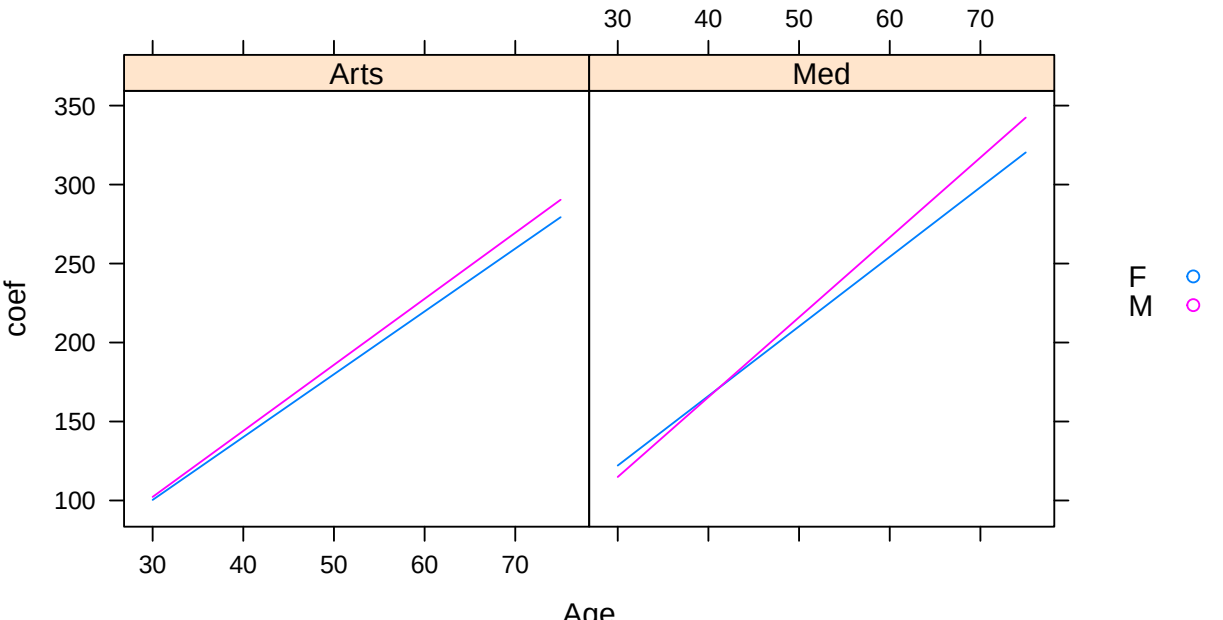
```
pred <- with(dd, pred.grid(Faculty, Sex, Age = seq(30,75,1)))
ww <- as.data.frame(wald(fit, pred = pred))
head(ww)
```

```
            coef       se       U2        L2 Faculty Sex Age L.(Intercept) L.Age
```

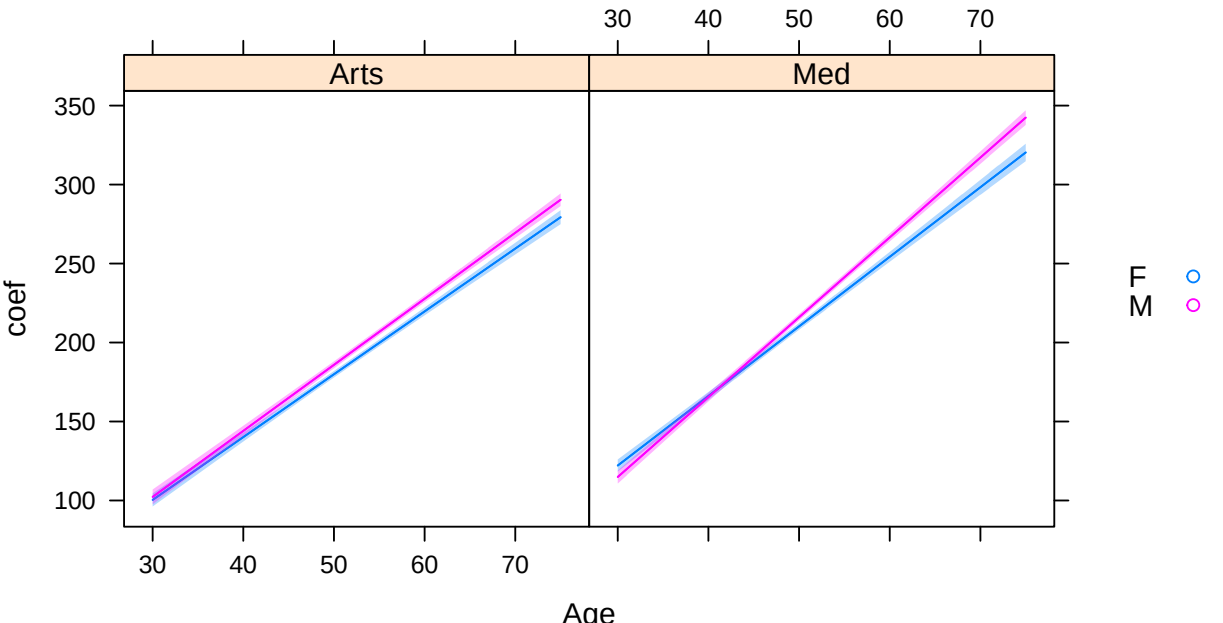
```
1 100.3628 2.080263 104.5233  96.20224    Arts   F  30             1    30
2 122.1229 1.940117 126.0031 118.24268     Med   F  30             1    30
3 102.2655 2.350772 106.9671  97.56400    Arts   M  30             1    30
4 114.8848 2.030886 118.9465 110.82299     Med   M  30             1    30
5 104.3403 2.008011 108.3563 100.32424    Arts   F  31             1    31
6 126.5274 1.862586 130.2526 122.80225     Med   F  31             1    31
  L.FacultyMed L.SexM L.Age:FacultyMed L.Age:SexM L.FacultyMed:SexM
1            0      0                0          0                  0
2            1      0               30          0                  0
3            0      1                0         30                  0
4            1      1               30         30                  1
5            0      0                0          0                  0
6            1      0               31          0                  0
  L.Age:FacultyMed:SexM
1                     0
2                     0
3                     0
4                    30
5                     0
```

6 0

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


Arts
Med
coef
Age
F
M
100
150
200
250
300
350
30
40
50
60
70

```
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
F
M
100
150
200
250
300
350
30
40
50
60
70

Analyzing the Gap

```
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  1442 15.39867 <.00001
   Estimate  Std.Error DF   t-value   p-value Lower 0.95 Upper 0.95
3   1.902778 3.139048  1442  0.606164 0.54450  -4.254810  8.060367
4  -7.238153 2.808656  1442 -2.577087 0.01006 -12.747642 -1.728663
7   2.105720 3.033885  1442  0.694067 0.48775  -3.845580  8.057020
8  -6.586926 2.702361  1442 -2.437471 0.01491 -11.887906 -1.285946
11  2.308662 2.929824  1442  0.787986 0.43083  -3.438512  8.055835
12 -5.935700 2.597841  1442 -2.284858 0.02247 -11.031653 -0.839747
15  2.511603 2.826987  1442  0.888438 0.37445  -3.033845  8.057051
```

```
16  -5.284473 2.495320  1442 -2.117754 0.03437 -10.179319 -0.389628
19   2.714545 2.725513  1442  0.995976 0.31943  -2.631851  8.060940
20  -4.633247 2.395053  1442 -1.934507 0.05325  -9.331408  0.064914
23   2.917486 2.625560  1442  1.111186 0.26667  -2.232839  8.067811
24  -3.982021 2.297337  1442 -1.733321 0.08325  -8.488500  0.524459
27   3.120428 2.527307  1442  1.234685 0.21715  -1.837164  8.078019
28  -3.330794 2.202509  1442 -1.512272 0.13068  -7.651260  0.989671
31   3.323369 2.430961  1442  1.367101 0.17181  -1.445230  8.091968
32  -2.679568 2.110961  1442 -1.269359 0.20452  -6.820452  1.461316
35   3.526311 2.336759  1442  1.509061 0.13150  -1.057499  8.110121
36  -2.028341 2.023137  1442 -1.002572 0.31624  -5.996949  1.940266
39   3.729253 2.244969  1442  1.661160 0.09690  -0.674502  8.133007
40  -1.377115 1.939544  1442 -0.710020 0.47781  -5.181744  2.427514
43   3.932194 2.155900  1442  1.823922 0.06837  -0.296843  8.161231
44  -0.725889 1.860750  1442 -0.390105 0.69652  -4.375956  2.924178
47   4.135136 2.069904  1442  1.997742 0.04593   0.074790  8.195481
48  -0.074662 1.787392  1442 -0.041772 0.96669  -3.580829  3.431505
51   4.338077 1.987379  1442  2.182813 0.02921   0.439613  8.236541
52   0.576564 1.720165  1442  0.335180 0.73754  -2.797729  3.950857
```

55	4.541019	1.908776	1442	2.379021	0.01749	0.796744	8.285293
56	1.227791	1.659813	1442	0.739716	0.45959	-2.028116	4.483698
59	4.743960	1.834598	1442	2.585831	0.00981	1.145194	8.342727
60	1.879017	1.607112	1442	1.169188	0.24252	-1.273512	5.031546
63	4.946902	1.765404	1442	2.802136	0.00514	1.483867	8.409937
64	2.530244	1.562837	1442	1.619007	0.10566	-0.535433	5.595920
67	5.149843	1.701801	1442	3.026113	0.00252	1.811572	8.488115
68	3.181470	1.527718	1442	2.082498	0.03747	0.184681	6.178258
71	5.352785	1.644439	1442	3.255082	0.00116	2.127036	8.578535
72	3.832696	1.502400	1442	2.551049	0.01084	0.885573	6.779820
75	5.555727	1.593992	1442	3.485417	0.00051	2.428936	8.682518
76	4.483923	1.487382	1442	3.014641	0.00262	1.566258	7.401587
79	5.758668	1.551133	1442	3.712555	0.00021	2.715949	8.801388
80	5.135149	1.482978	1442	3.462729	0.00055	2.226125	8.044174
83	5.961610	1.516508	1442	3.931144	0.00009	2.986812	8.936407
84	5.786376	1.489281	1442	3.885350	0.00011	2.864987	8.707764
87	6.164551	1.490689	1442	4.135371	0.00004	3.240401	9.088702
88	6.437602	1.506157	1442	4.274191	0.00002	3.483109	9.392095
91	6.367493	1.474139	1442	4.319466	0.00002	3.475807	9.259179

```
92   7.088828 1.533257  1442  4.623380 <.00001   4.081176 10.096481
95   6.570434 1.467172  1442  4.478298 0.00001   3.692414  9.448455
96   7.740055 1.570052  1442  4.929809 <.00001   4.660225 10.819885
99   6.773376 1.469925  1442  4.607974 <.00001   3.889956  9.656796
100  8.391281 1.615879  1442  5.193013 <.00001   5.221556 11.561007
103  6.976318 1.482343  1442  4.706278 <.00001   4.068538  9.884097
104  9.042508 1.669996  1442  5.414689 <.00001   5.766627 12.318389
107  7.179259 1.504187  1442  4.772851 <.00001   4.228631 10.129888
108  9.693734 1.731624  1442  5.598058 <.00001   6.296961 13.090507
111  7.382201 1.535054  1442  4.809082 <.00001   4.371023 10.393379
112 10.344961 1.799994  1442  5.747220 <.00001   6.814074 13.875847
115  7.585142 1.574414  1442  4.817755 <.00001   4.496755 10.673530
116 10.996187 1.874367  1442  5.866615 <.00001   7.319410 14.672964
119  7.788084 1.621649  1442  4.802571 <.00001   4.607040 10.969127
120 11.647413 1.954057  1442  5.960631 <.00001   7.814314 15.480512
123  7.991025 1.676093  1442  4.767652 <.00001   4.703185 11.278866
124 12.298640 2.038442  1442  6.033353 <.00001   8.300011 16.297269
127  8.193967 1.737067  1442  4.717127 <.00001   4.786517 11.601417
128 12.949866 2.126963  1442  6.088432 <.00001   8.777594 17.122138
```

```
131  8.396909 1.803911  1442  4.654834 <.00001   4.858337 11.935480
132 13.601093 2.219124  1442  6.129037 <.00001   9.248036 17.954149
135  8.599850 1.875997  1442  4.584149 <.00001   4.919874 12.279826
136 14.252319 2.314491  1442  6.157863 <.00001   9.712189 18.792449
139  8.802792 1.952744  1442  4.507908 0.00001   4.972268 12.633316
140 14.903545 2.412684  1442  6.177164 <.00001  10.170799 19.636292
143  9.005733 2.033625  1442  4.428413 0.00001   5.016553 12.994914
144 15.554772 2.513372  1442  6.188807 <.00001  10.624516 20.485028
147  9.208675 2.118166  1442  4.347475 0.00001   5.053657 13.363692
148 16.205998 2.616266  1442  6.194324 <.00001  11.073904 21.338093
151  9.411616 2.205947  1442  4.266474 0.00002   5.084408 13.738825
152 16.857225 2.721116  1442  6.194967 <.00001  11.519455 22.194995
155  9.614558 2.296596  1442  4.186439 0.00003   5.109532 14.119584
156 17.508451 2.827705  1442  6.191752 <.00001  11.961594 23.055308
159  9.817500 2.389786  1442  4.108108 0.00004   5.129670 14.505329
160 18.159678 2.935844  1442  6.185505 <.00001  12.400695 23.918660
163 10.020441 2.485232  1442  4.031994 0.00006   5.145384 14.895498
164 18.810904 3.045367  1442  6.176893 <.00001  12.837081 24.784727
167 10.223383 2.582684  1442  3.958434 0.00008   5.157163 15.289602
```

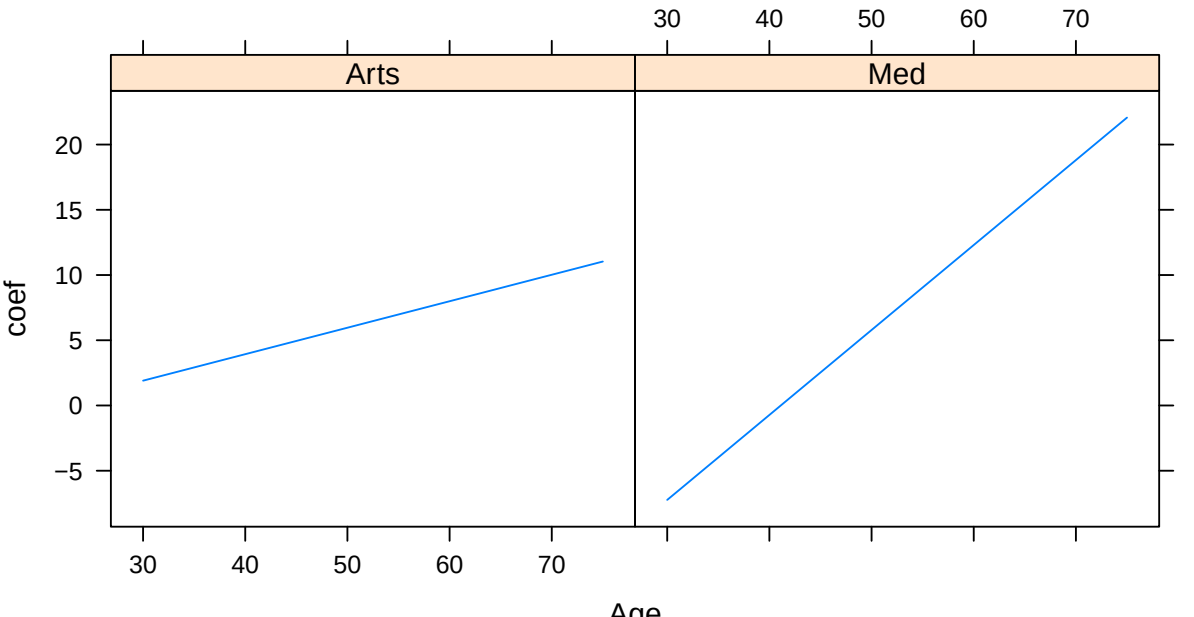
```
168 19.462130 3.156130  1442  6.166454 <.00001  13.271033 25.653227
171 10.426324 2.681923  1442  3.887631 0.00011   5.165437 15.687212
172 20.113357 3.268007  1442  6.154625 <.00001  13.702801 26.523913
175 10.629266 2.782757  1442  3.819688 0.00014   5.170580 16.087952
176 20.764583 3.380887  1442  6.141756 <.00001  14.132599 27.396567
179 10.832207 2.885021  1442  3.754638 0.00018   5.172921 16.491494
180 21.415810 3.494674  1442  6.128128 <.00001  14.560620 28.270999
183 11.035149 2.988566  1442  3.692457 0.00023   5.172747 16.897551
184 22.067036 3.609281  1442  6.113970 <.00001  14.987032 29.147040
```

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

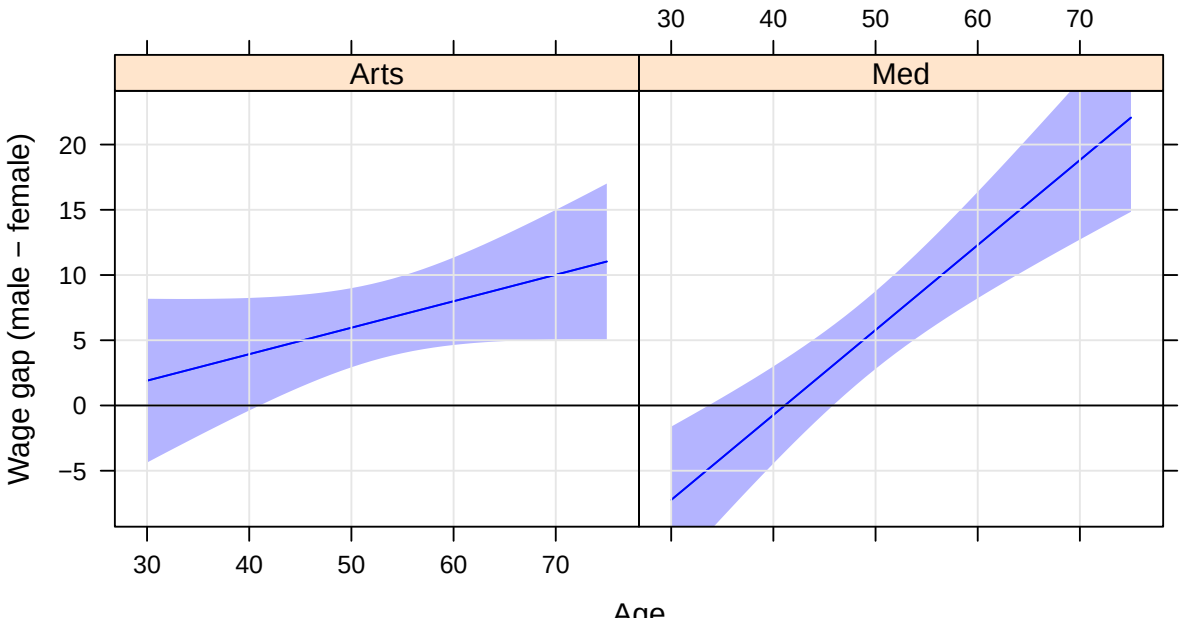
```
        coef       se        U2         L2 Faculty Age L.(Intercept) L.Age
3   1.902778 3.139048  8.180873  -4.375317    Arts  30             0     0
4  -7.238153 2.808656 -1.620840 -12.855465     Med  30             0     0
7   2.105720 3.033885  8.173489  -3.962049    Arts  31             0     0
8  -6.586926 2.702361 -1.182204 -11.991649     Med  31             0     0
11  2.308662 2.929824  8.168310  -3.550986    Arts  32             0     0
12 -5.935700 2.597841 -0.740017 -11.131383     Med  32             0     0
```

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

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


Arts
Med
30
40
50
60
70
20
15
10
5
0
−5
coef
Age

```
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
20
15
10
5
0
−5
30
40
50
60
70

3 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
  12552.91 12621.48 -6263.454

Variance function:
 Structure: Constant plus power of variance covariate, different strata
 Formula: ~fitted(.) | Faculty
 Parameter estimates:
               Arts        Med
const 9.749647e-06 29.7003504
power 7.307494e-01  0.7216105
```

```
Coefficients:
                        Value Std.Error  t-value p-value
(Intercept)         -19.762576  2.784153 -7.09824  0.0000
Age                   3.993222  0.057807 69.07861  0.0000
FacultyMed            8.149613  5.857826  1.39123  0.1644
SexM                 -4.803421  4.190896 -1.14616  0.2519
Age:FacultyMed        0.445143  0.125251  3.55401  0.0004
Age:SexM              0.213444  0.083970  2.54190  0.0111
FacultyMed:SexM     -20.688326  8.384991 -2.46730  0.0137
Age:FacultyMed:SexM   0.410238  0.174064  2.35682  0.0186

 Correlation:
                    (Intr) Age    FcltyM SexM   Ag:FcM Ag:SxM FcM:SM
Age                 -0.969
FacultyMed          -0.475  0.461
SexM                -0.664  0.644  0.316
Age:FacultyMed       0.447 -0.462 -0.971 -0.297
Age:SexM             0.667 -0.688 -0.317 -0.970  0.318
```

```
FacultyMed:SexM      0.332 -0.322 -0.699 -0.500  0.679  0.485
Age:FacultyMed:SexM -0.322  0.332  0.699  0.468 -0.720 -0.482 -0.971

Standardized residuals:
        Min          Q1         Med          Q3         Max
-3.53559716 -0.67370477  0.02089916  0.67207429  3.51634551

Residual standard error: 0.3070914
Degrees of freedom: 1451 total; 1443 residual
```

```
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 12780.75 12828.22 -6381.373
fitgroups           2 10 12621.74 12674.48 -6300.868 1 vs 2 161.00929  <.0001
fitpower            3 11 12557.45 12615.47 -6267.727 2 vs 3  66.28162  <.0001
fitconpower         4 13 12552.91 12621.48 -6263.454 3 vs 4   8.54737  0.0139
```

```
library(car)
```

```
Loading required package: carData
```

```
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)          -18.96  -18.96  -19.57  -19.76
SE                     4.44    3.31    2.92    2.78

Age                  3.9775  3.9775  3.9893  3.9932
```

SE	0.0837	0.0624	0.0589	0.0578
FacultyMed	8.95	8.95	7.19	8.15
SE	6.35	6.40	5.81	5.86
SexM	-4.19	-4.19	-4.63	-4.80
SE	6.53	4.87	4.37	4.19
Age:FacultyMed	0.427	0.427	0.465	0.445
SE	0.125	0.128	0.125	0.125
Age:SexM	0.2029	0.2029	0.2104	0.2134
SE	0.1195	0.0891	0.0852	0.0840
FacultyMed:SexM	-22.59	-22.59	-20.21	-20.69
SE	9.10	9.07	8.33	8.38
Age:FacultyMed:SexM	0.448	0.448	0.399	0.410
SE	0.174	0.176	0.175	0.174

4 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  1442 14.1574 <.00001
    Estimate  Std.Error DF   t-value   p-value Lower 0.95 Upper 0.95
3   1.599889 1.850517  1442  0.864563 0.38742  -2.030104  5.229882
4  -6.781289 3.023151  1442 -2.243120 0.02504 -12.711533 -0.851044
7   1.813333 1.780923  1442  1.018198 0.30875  -1.680144  5.306809
8  -6.157607 2.899415  1442 -2.123741 0.03386 -11.845130 -0.470083
11  2.026776 1.712618  1442  1.183437 0.23683  -1.332713  5.386266
12 -5.533925 2.778538  1442 -1.991668 0.04660 -10.984334 -0.083516
```

```
15   2.240220 1.645764  1442  1.361204 0.17366  -0.988128  5.468568
16  -4.910243 2.660908  1442 -1.845326 0.06520 -10.129908  0.309422
19   2.453664 1.580544  1442  1.552418 0.12078  -0.646747  5.554075
20  -4.286561 2.546976  1442 -1.683000 0.09259  -9.282735  0.709613
23   2.667107 1.517168  1442  1.757951 0.07897  -0.308986  5.643201
24  -3.662879 2.437259  1442 -1.502868 0.13309  -8.443832  1.118074
27   2.880551 1.455879  1442  1.978565 0.04806   0.024684  5.736418
28  -3.039197 2.332354  1442 -1.303060 0.19276  -7.614367  1.535973
31   3.093995 1.396950  1442  2.214822 0.02693   0.353723  5.834266
32  -2.415515 2.232938  1442 -1.081765 0.27954  -6.795670  1.964640
35   3.307438 1.340692  1442  2.466963 0.01374   0.677522  5.937355
36  -1.791833 2.139777  1442 -0.837392 0.40251  -5.989242  2.405576
39   3.520882 1.287457  1442  2.734757 0.00632   0.995393  6.046371
40  -1.168151 2.053722  1442 -0.568797 0.56958  -5.196754  2.860451
43   3.734326 1.237633  1442  3.017312 0.00259   1.306571  6.162080
44  -0.544469 1.975701  1442 -0.275583 0.78291  -4.420026  3.331087
47   3.947769 1.191650  1442  3.312860 0.00095   1.610216  6.285323
48   0.079213 1.906702  1442  0.041544 0.96687  -3.660994  3.819419
51   4.161213 1.149967  1442  3.618549 0.00031   1.905425  6.417001
```

```
52   0.702895 1.847735  1442  0.380409 0.70370  -2.921642  4.327431
55   4.374657 1.113069  1442  3.930266 0.00009   2.191249  6.558064
56   1.326577 1.799786  1442  0.737074 0.46120  -2.203903  4.857057
59   4.588100 1.081444  1442  4.242569 0.00002   2.466728  6.709472
60   1.950259 1.763755  1442  1.105742 0.26902  -1.509542  5.410059
63   4.801544 1.055567  1442  4.548781 0.00001   2.730932  6.872156
64   2.573940 1.740382  1442  1.478952 0.13937  -0.840011  5.987892
67   5.014988 1.035869  1442  4.841333 <.00001   2.983016  7.046960
68   3.197622 1.730179  1442  1.848145 0.06479  -0.196315  6.591560
71   5.228431 1.022707  1442  5.112344 <.00001   3.222278  7.234585
72   3.821304 1.733380  1442  2.204539 0.02764   0.421088  7.221521
75   5.441875 1.016335  1442  5.354410 <.00001   3.448221  7.435529
76   4.444986 1.749911  1442  2.540122 0.01119   1.012343  7.877629
79   5.655319 1.016880  1442  5.561439 <.00001   3.660595  7.650042
80   5.068668 1.779399  1442  2.848527 0.00445   1.578180  8.559157
83   5.868762 1.024332  1442  5.729354 <.00001   3.859421  7.878103
84   5.692350 1.821217  1442  3.125574 0.00181   2.119831  9.264869
87   6.082206 1.038542  1442  5.856485 <.00001   4.044991  8.119421
88   6.316032 1.874539  1442  3.369378 0.00077   2.638916  9.993148
```

```
91   6.295650 1.059238  1442  5.943567 <.00001   4.217838  8.373461
92   6.939714 1.938416  1442  3.580095 0.00035   3.137296 10.742132
95   6.509093 1.086048  1442  5.993374 <.00001   4.378690  8.639497
96   7.563396 2.011843  1442  3.759436 0.00018   3.616943 11.509849
99   6.722537 1.118535  1442  6.010129 <.00001   4.528408  8.916666
100  8.187078 2.093815  1442  3.910124 0.00010   4.079828 12.294328
103  6.935981 1.156218  1442  5.998852 <.00001   4.667931  9.204030
104  8.810760 2.183371  1442  4.035394 0.00006   4.527837 13.093683
107  7.149424 1.198609  1442  5.964769 <.00001   4.798221  9.500628
108  9.434442 2.279615  1442  4.138611 0.00004   4.962724 13.906159
111  7.362868 1.245226  1442  5.912878 <.00001   4.920220  9.805516
112 10.058124 2.381739  1442  4.223017 0.00003   5.386080 14.730168
115  7.576312 1.295613  1442  5.847664 <.00001   5.034823 10.117800
116 10.681806 2.489018  1442  4.291575 0.00002   5.799322 15.564289
119  7.789755 1.349349  1442  5.772974 <.00001   5.142858 10.436652
120 11.305488 2.600814  1442  4.346904 0.00001   6.203704 16.407272
123  8.003199 1.406049  1442  5.691979 <.00001   5.245079 10.761319
124 11.929170 2.716570  1442  4.391262 0.00001   6.600318 17.258022
127  8.216643 1.465369  1442  5.607219 <.00001   5.342160 11.091125
```

```
128 12.552852 2.835801  1442  4.426563 0.00001   6.990115 18.115588
131  8.430086 1.527003  1442  5.520673 <.00001   5.434700 11.425472
132 13.176534 2.958086  1442  4.454412 0.00001   7.373921 18.979146
135  8.643530 1.590684  1442  5.433845 <.00001   5.523228 11.763832
136 13.800215 3.083063  1442  4.476138 0.00001   7.752447 19.847984
139  8.856974 1.656174  1442  5.347851 <.00001   5.608204 12.105743
140 14.423897 3.210417  1442  4.492843 0.00001   8.126311 20.721484
143  9.070417 1.723268  1442  5.263497 <.00001   5.690036 12.450799
144 15.047579 3.339875  1442  4.505431 0.00001   8.496045 21.599114
147  9.283861 1.791786  1442  5.181345 <.00001   5.769075 12.798647
148 15.671261 3.471204  1442  4.514648 0.00001   8.862112 22.480411
151  9.497305 1.861569  1442  5.101773 <.00001   5.845631 13.148978
152 16.294943 3.604197  1442  4.521102 0.00001   9.224912 23.364974
155  9.710748 1.932482  1442  5.025014 <.00001   5.919971 13.501525
156 16.918625 3.738678  1442  4.525296 0.00001   9.584795 24.252456
159  9.924192 2.004404  1442  4.951194 <.00001   5.992333 13.856051
160 17.542307 3.874492  1442  4.527641 0.00001   9.942063 25.142551
163 10.137636 2.077230  1442  4.880363 <.00001   6.062920 14.212352
164 18.165989 4.011503  1442  4.528475 0.00001  10.296983 26.034995
```

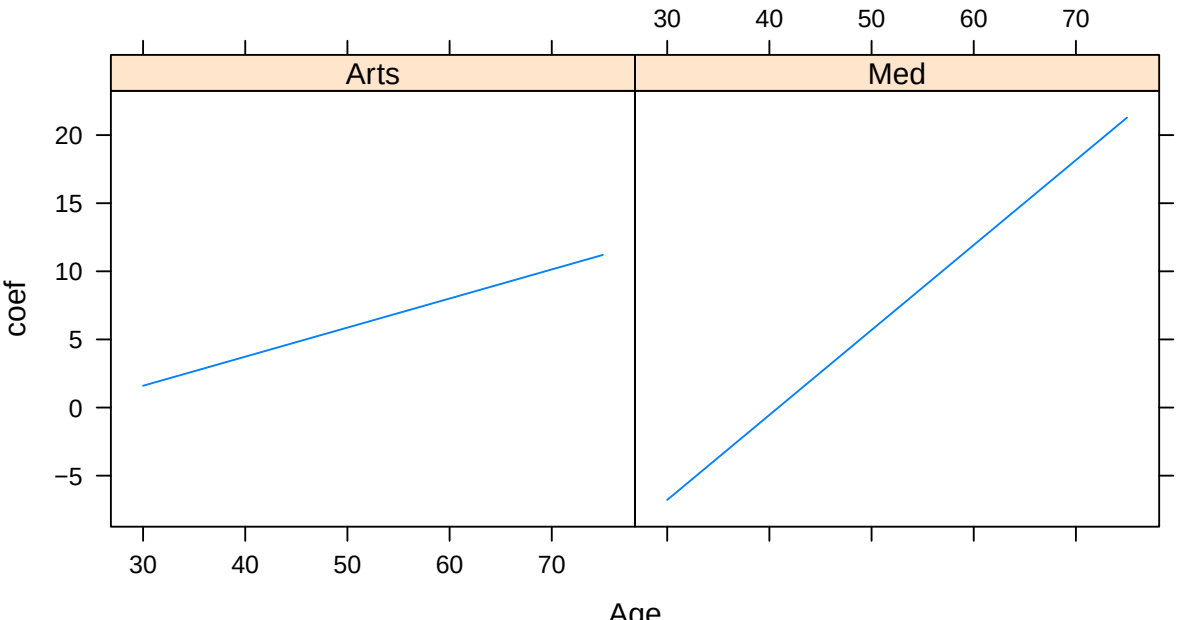
```
167 10.351079 2.150869  1442  4.812511 <.00001   6.131913 14.570246
168 18.789671 4.149592  1442  4.528077 0.00001  10.649788 26.929555
171 10.564523 2.225239  1442  4.747589 <.00001   6.199470 14.929576
172 19.413353 4.288656  1442  4.526675 0.00001  11.000680 27.826026
175 10.777967 2.300271  1442  4.685521 <.00001   6.265731 15.290202
176 20.037035 4.428603  1442  4.524460 0.00001  11.349841 28.724229
179 10.991410 2.375901  1442  4.626208 <.00001   6.330819 15.652002
180 20.660717 4.569351  1442  4.521587 0.00001  11.697430 29.624004
183 11.204854 2.452073  1442  4.569543 0.00001   6.394841 16.014866
184 21.284399 4.710829  1442  4.518185 0.00001  12.043587 30.525210
```

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

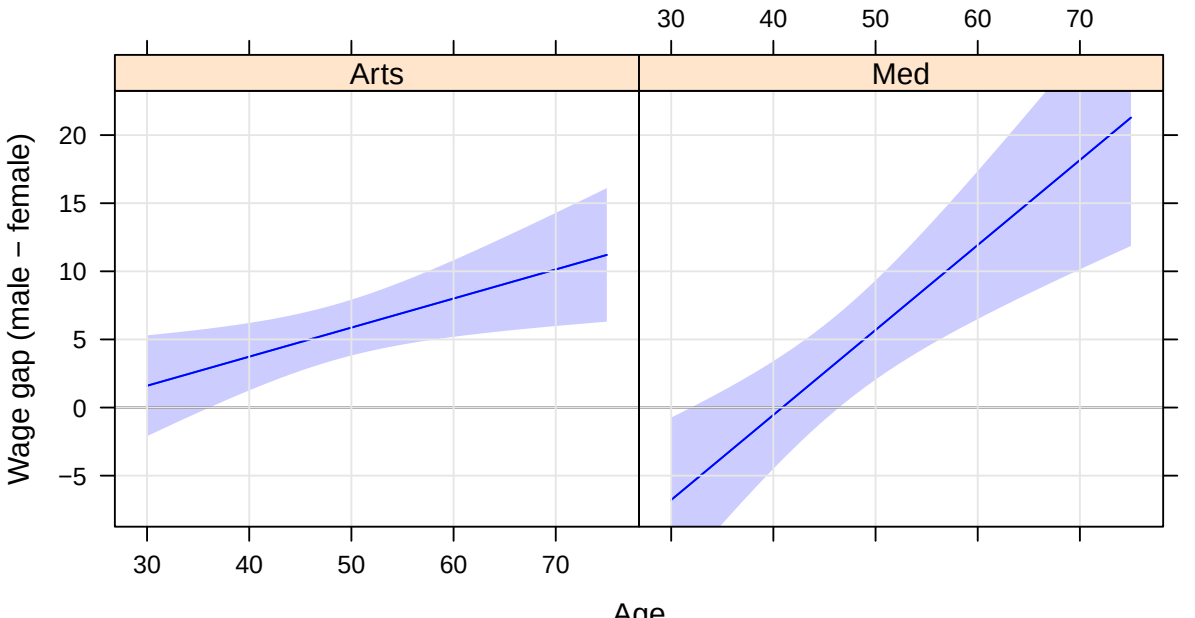
```
        coef       se          U2         L2 Faculty Age L.(Intercept) L.Age
3   1.599889 1.850517  5.30092255  -2.101144    Arts  30             0     0
4  -6.781289 3.023151 -0.73498710 -12.827590     Med  30             0     0
7   1.813333 1.780923  5.37517823  -1.748513    Arts  31             0     0
8  -6.157607 2.899415 -0.35877599 -11.956438     Med  31             0     0
11  2.026776 1.712618  5.45201298  -1.398460    Arts  32             0     0
```

```
12 -5.533925 2.778538  0.02315111 -11.091001     Med  32             0     0
   L.FacultyMed L.SexM L.Age:FacultyMed L.Age:SexM L.FacultyMed:SexM
3             0      1                 0         30                 0
4             0      1                 0         30                 1
7             0      1                 0         31                 0
8             0      1                 0         31                 1
11            0      1                 0         32                 0
12            0      1                 0         32                 1
   L.Age:FacultyMed:SexM
3                      0
4                     30
7                      0
8                     31
11                     0
12                    32
```

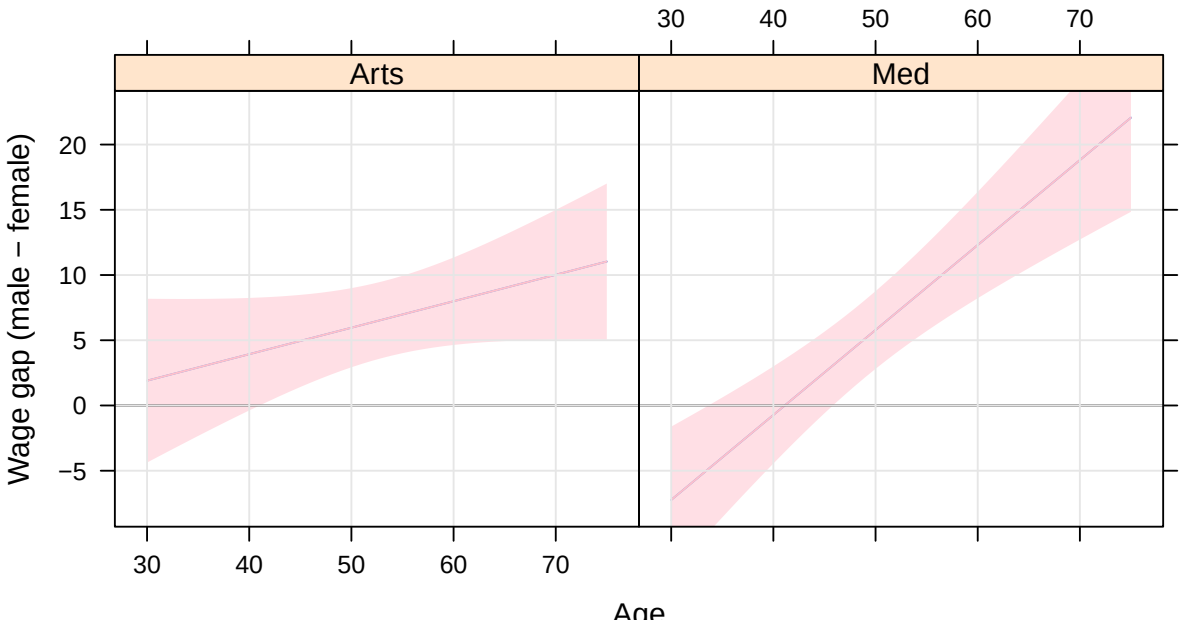
```
xyplot(coef ~ Age | Faculty, wgap2,
       type = 'l', auto.key = list(space='right'))
```


Arts
Med
coef
Age
20
15
10
5
0
−5
30
40
50
60
70

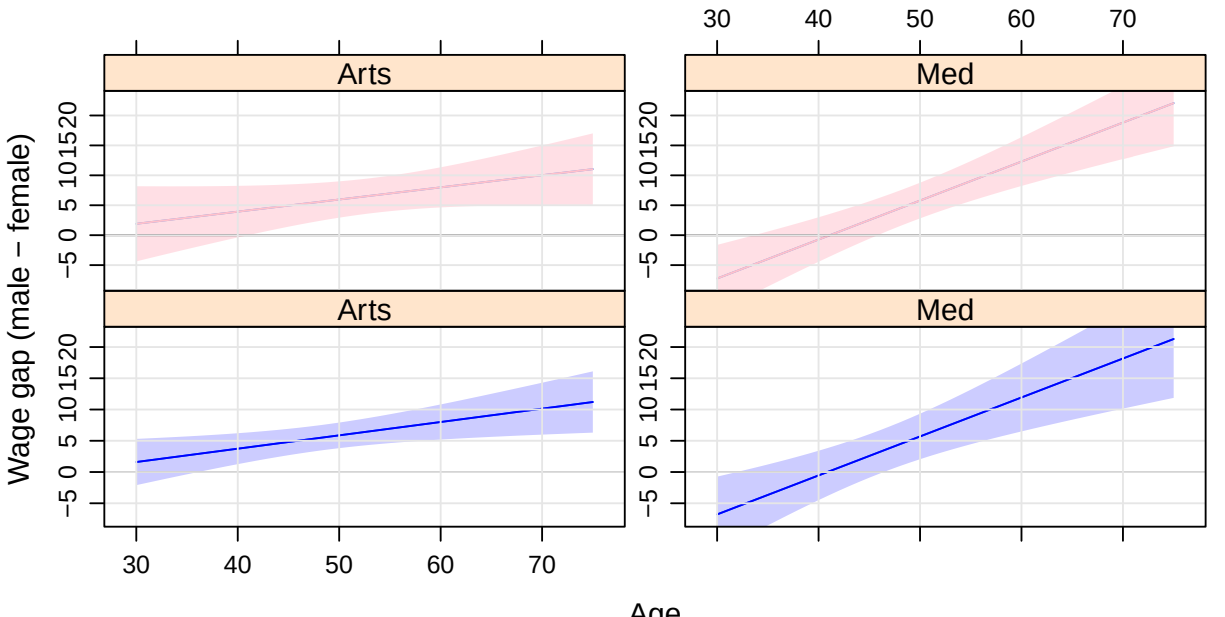
```
xyplot(coef ~ Age | Faculty, wgap2,
       type = 'l', auto.key = list(space='right'),
       ylab = 'Wage gap (male - female)',
       fit = wgap2$coef,
       lower = wgap2$L2,
       upper = wgap2$U2,
       subscripts = TRUE) +
  layer(panel.fit(..., alpha = .2)) +
  layer(panel.abline(h=0))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


Arts
Med
Wage gap (male – female)
Age
30
40
50
60
70
20
15
10
5
0
−5

```
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
20
15
10
5
0
−5
30
40
50
60
70

```
c(plhet, plnohet)
```


Arts
Med
Arts
Med
Wage gap (male – female)
−5 0 5 10 15 20
30 40 50 60 70
Age

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.