

Robustness Checks

Colin Linke

2024-06-03

Does including Substance usage as additional Covariates show a similar story? is there a significant difference

There is the option of using standardized transformed and / or logarithmic transformed for the variables. In the following I opted for using the standardized Data, but avoiding an additional logarithmic transformation of the substance use data due to the fact that all models performed almost identically, though with a logarithmic transformation slightly better, but this would lead to a less intuitive interpretation. Furthermore, due to a flat optimization plane and different scales the two substance usage variables lie on, the standadisation of these variables alleviates these problems somewhat.

```
sm.nlme.std.robust2.1 <- lme(std_GPA ~ std_Avg_Drinks * Semester + std_Avg_Drinks * std_Avg_MJ + std_Avg_MJ * Semester, random = ~ 1 | BARCS_ID, data = data.file)

sm.nlme.std.robust2.2 <- lme(std_GPA ~ std_Avg_Drinks * Semester + std_Avg_Drinks * std_Avg_MJ + std_Avg_MJ * Semester, random = ~ 1 | BARCS_ID, data = data.file)

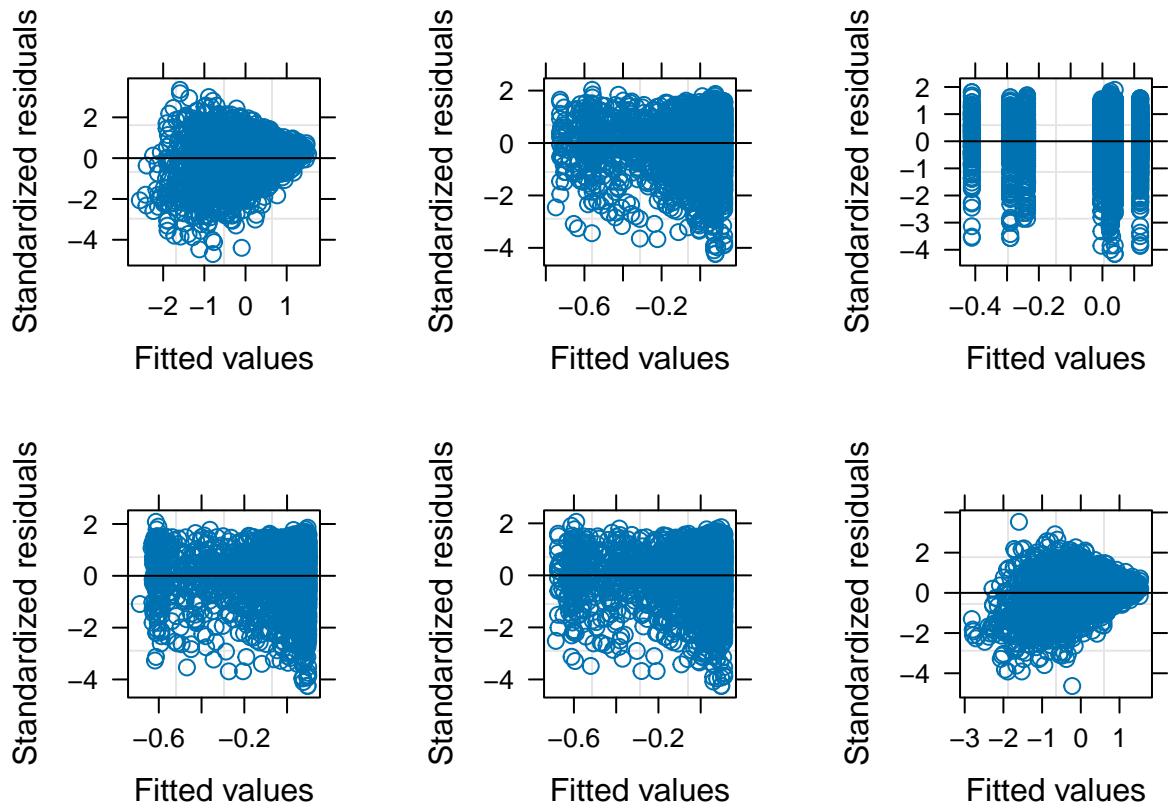
sm.nlme.std.ARMA11 <- lme(std_GPA ~ Cluster_current * Semester, random = ~ 1 | BARCS_ID, data = data.file)

sm.nlme.std.noSemester <- lme(std_GPA ~ std_Avg_Drinks * std_Avg_MJ, random = ~ 1 | BARCS_ID, data = data.file)

sm.nlme.std.noSemester.time <- lme(std_GPA ~ std_Avg_Drinks * std_Avg_MJ + std_Avg_Drinks*Time + std_Avg_MJ * Time, random = ~ 1 | BARCS_ID, data = data.file)

sm.nlme.std.noSemester.timeRE <- lme(std_GPA ~ std_Avg_Drinks * std_Avg_MJ + std_Avg_Drinks*Time + std_Avg_MJ * Time, random = ~ 1 | BARCS_ID, data = data.file)

cowplot::plot_grid(plot(sm.nlme.std.robust2.1), plot(sm.nlme.std.robust2.2), plot(sm.nlme.std.ARMA11),
```



```

## Linear mixed-effects model fit by REML
## Data: data.file.long
##      AIC      BIC    logLik
## 9378.35 9471.948 -4674.175
##
## Random effects:
##   Formula: ~1 | BARCS_ID
##             (Intercept) Residual
## StdDev:  0.7704306 0.630603
##
## Fixed effects: std_GPA ~ std_Avg_Drinks * Semester + std_Avg_Drinks * std_Avg_MJ +
##                 Value Std.Error DF t-value p-value
##   (Intercept) -0.00845156 0.02981378 2650 -0.283478 0.7768
##   std_Avg_Drinks -0.09933159 0.02708363 2650 -3.667588 0.0002
##   Semester2 -0.02057220 0.02679900 2650 -0.767648 0.4428
##   Semester3 -0.00632348 0.02976816 2650 -0.212424 0.8318
##   Semester4 -0.02745734 0.03178949 2650 -0.863724 0.3878
##   std_Avg_MJ -0.15253300 0.02840815 2650 -5.369339 0.0000
##   std_Avg_Drinks:Semester2 0.02190425 0.03145727 2650 0.696317 0.4863
##   std_Avg_Drinks:Semester3 0.04294779 0.03235562 2650 1.327367 0.1845
##   std_Avg_Drinks:Semester4 0.02839279 0.03391358 2650 0.837210 0.4025
##   std_Avg_Drinks:std_Avg_MJ 0.02111682 0.00941777 2650 2.242230 0.0250
##   Semester2:std_Avg_MJ 0.01075971 0.03138780 2650 0.342799 0.7318
##   Semester3:std_Avg_MJ 0.00384125 0.03426492 2650 0.112105 0.9107
##   Semester4:std_Avg_MJ -0.03016358 0.03680673 2650 -0.819513 0.4126
##
## Correlation:
##   (Intr) st_A_D Smstr2 Smstr3 Smstr4 s_A_MJ s_A_D:S2
##   std_Avg_Drinks 0.034

```

```

## Semester2          -0.441  0.005
## Semester3         -0.403 -0.049  0.447
## Semester4         -0.380 -0.054  0.418  0.423
## std_Avg_MJ          0.037 -0.269 -0.005 -0.032 -0.054
## std_Avg_Drinks:Semester2  0.005 -0.524  0.003 -0.008 -0.008  0.232
## std_Avg_Drinks:Semester3  0.010 -0.565 -0.002 -0.087 -0.010  0.262  0.480
## std_Avg_Drinks:Semester4  0.020 -0.581 -0.001 -0.003 -0.109  0.302  0.461
## std_Avg_Drinks:std_Avg_MJ -0.131 -0.260 -0.007  0.042  0.064 -0.233 -0.038
## Semester2:std_Avg_MJ      -0.005  0.229  0.011  0.011  0.012 -0.529 -0.447
## Semester3:std_Avg_MJ      -0.001  0.237  0.007  0.055  0.018 -0.535 -0.200
## Semester4:std_Avg_MJ       0.004  0.260  0.006  0.012  0.063 -0.531 -0.185
##                               s_A_D:S3 s_A_D:S4 s_A_D:_ S2:_A_ S3:_A_
## std_Avg_Drinks
## Semester2
## Semester3
## Semester4
## std_Avg_MJ
## std_Avg_Drinks:Semester2
## std_Avg_Drinks:Semester3
## std_Avg_Drinks:Semester4   0.503
## std_Avg_Drinks:std_Avg_MJ -0.080   -0.148
## Semester2:std_Avg_MJ       -0.206   -0.199  -0.001
## Semester3:std_Avg_MJ       -0.422   -0.205  -0.020   0.461
## Semester4:std_Avg_MJ       -0.198   -0.430  -0.062   0.429   0.460
##
## Standardized Within-Group Residuals:
##           Min        Q1        Med        Q3        Max
## -4.6980545 -0.4151518  0.1077233  0.5367422  3.3468825
##
## Number of Observations: 3802
## Number of Groups: 1140

## Linear mixed-effects model fit by REML
##   Data: data.file.long
##   AIC      BIC      logLik
## 9300.111 9406.189 -4633.056
##
## Random effects:
##   Formula: ~1 | BARCS_ID
##             (Intercept) Residual
##   StdDev: 0.001614964 0.9951262
##
## Correlation Structure: ARMA(1,1)
##   Formula: ~Time | BARCS_ID
##   Parameter estimate(s):
##     Phi1     Theta1
## 0.8584752 -0.3854289
## Fixed effects: std_GPA ~ std_Avg_Drinks * Semester + std_Avg_Drinks * std_Avg_MJ +
##                           Value Std.Error DF t-value p-value
## (Intercept)      -0.00787951 0.02979547 2650 -0.264453 0.7915
## std_Avg_Drinks   -0.09957067 0.02751705 2650 -3.618508 0.0003
## Semester2        -0.02270359 0.02502726 2650 -0.907154 0.3644
## Semester3        -0.00846586 0.03084445 2650 -0.274470 0.7837
## Semester4        -0.02356419 0.03542545 2650 -0.665177 0.5060

```

```

## std_Avg_MJ           -0.15604683 0.02884670 2650 -5.409520  0.0000
## std_Avg_Drinks:Semester2 0.02363643 0.02952478 2650  0.800563  0.4235
## std_Avg_Drinks:Semester3 0.04987121 0.03273176 2650  1.523634  0.1277
## std_Avg_Drinks:Semester4 0.03203528 0.03592943 2650  0.891617  0.3727
## std_Avg_Drinks:std_Avg_MJ 0.02080182 0.00939361 2650  2.214465  0.0269
## Semester2:std_Avg_MJ      0.01277801 0.02939390 2650  0.434716  0.6638
## Semester3:std_Avg_MJ      -0.00017217 0.03475086 2650 -0.004955  0.9960
## Semester4:std_Avg_MJ      -0.03314348 0.03939450 2650 -0.841323  0.4002
## Correlation:
##                               (Intr) st_A_D Smstr2 Smstr3 Smstr4 s_A_MJ s_A_D:S2
## std_Avg_Drinks            0.032
## Semester2                 -0.412  0.006
## Semester3                 -0.427 -0.042  0.508
## Semester4                 -0.439 -0.038  0.428  0.561
## std_Avg_MJ                  0.038 -0.274 -0.005 -0.030 -0.039
## std_Avg_Drinks:Semester2  0.006 -0.522  0.003 -0.016 -0.012  0.239
## std_Avg_Drinks:Semester3  0.011 -0.597 -0.002 -0.081 -0.027  0.276  0.524
## std_Avg_Drinks:Semester4  0.019 -0.595 -0.001 -0.013 -0.107  0.291  0.458
## std_Avg_Drinks:std_Avg_MJ -0.131 -0.248 -0.008  0.042  0.055 -0.243 -0.050
## Semester2:std_Avg_MJ       -0.006  0.230  0.012  0.010  0.006 -0.524 -0.449
## Semester3:std_Avg_MJ       -0.003  0.246  0.008  0.048  0.013 -0.564 -0.215
## Semester4:std_Avg_MJ       0.003  0.250  0.007  0.014  0.053 -0.544 -0.178
##                               s_A_D:S3 s_A_D:S4 s_A_D:_ S2:_A_ S3:_A_
## std_Avg_Drinks
## Semester2
## Semester3
## Semester4
## std_Avg_MJ
## std_Avg_Drinks:Semester2
## std_Avg_Drinks:Semester3
## std_Avg_Drinks:Semester4  0.574
## std_Avg_Drinks:std_Avg_MJ -0.089  -0.142
## Semester2:std_Avg_MJ      -0.223  -0.192   0.007
## Semester3:std_Avg_MJ      -0.413  -0.220  -0.011   0.506
## Semester4:std_Avg_MJ      -0.214  -0.410  -0.050   0.428  0.551
##
## Standardized Within-Group Residuals:
##      Min        Q1        Med        Q3        Max
## -4.2352992 -0.5255693  0.1920014  0.7716773  2.0286823
##
## Number of Observations: 3802
## Number of Groups: 1140
## Linear mixed-effects model fit by REML
## Data: data.file.long
##      AIC      BIC      logLik
## 9307.691 9407.533 -4637.845
##
## Random effects:
## Formula: ~1 | BARCS_ID
##             (Intercept) Residual
## StdDev: 0.001267291 1.004147
##
## Correlation Structure: ARMA(1,1)

```

```

## Formula: ~Time | BARCS_ID
## Parameter estimate(s):
##      Phi1      Theta1
## 0.8645951 -0.3928098
## Fixed effects: std_GPA ~ Cluster_current * Semester
##                               Value Std.Error DF t-value
## (Intercept)          0.1187900 0.04099713 2651 2.897520
## Cluster_current2nd.cluster -0.1218963 0.05194456 2651 -2.346661
## Cluster_current3rd.cluster -0.4085500 0.07195048 2651 -5.678211
## Semester2           -0.0800065 0.03916589 2651 -2.042760
## Semester3           -0.0930723 0.04600864 2651 -2.022930
## Semester4           -0.0930098 0.05232604 2651 -1.777506
## Cluster_current2nd.cluster:Semester2 0.1063079 0.05887522 2651 1.805648
## Cluster_current3rd.cluster:Semester2 0.1326945 0.07411109 2651 1.790481
## Cluster_current2nd.cluster:Semester3 0.1235550 0.06749057 2651 1.830701
## Cluster_current3rd.cluster:Semester3 0.1347959 0.09076725 2651 1.485072
## Cluster_current2nd.cluster:Semester4 0.1268316 0.07636792 2651 1.660797
## Cluster_current3rd.cluster:Semester4 -0.0279823 0.10318464 2651 -0.271187
##                               p-value
## (Intercept)          0.0038
## Cluster_current2nd.cluster 0.0190
## Cluster_current3rd.cluster 0.0000
## Semester2           0.0412
## Semester3           0.0432
## Semester4            0.0756
## Cluster_current2nd.cluster:Semester2 0.0711
## Cluster_current3rd.cluster:Semester2 0.0735
## Cluster_current2nd.cluster:Semester3 0.0673
## Cluster_current3rd.cluster:Semester3 0.1376
## Cluster_current2nd.cluster:Semester4 0.0969
## Cluster_current3rd.cluster:Semester4 0.7863
## Correlation:
##                               (Intr) Cls_2. Cls_3. Smstr2 Smstr3 Smstr4
## Cluster_current2nd.cluster -0.634
## Cluster_current3rd.cluster -0.498  0.405
## Semester2                 -0.500  0.430  0.299
## Semester3                 -0.511  0.425  0.298  0.524
## Semester4                 -0.503  0.403  0.289  0.448  0.555
## Cluster_current2nd.cluster:Semester2 0.366 -0.591 -0.210 -0.703 -0.360 -0.308
## Cluster_current3rd.cluster:Semester2 0.272 -0.199 -0.592 -0.534 -0.277 -0.237
## Cluster_current2nd.cluster:Semester3 0.368 -0.597 -0.206 -0.371 -0.690 -0.370
## Cluster_current3rd.cluster:Semester3 0.268 -0.205 -0.563 -0.268 -0.508 -0.279
## Cluster_current2nd.cluster:Semester4 0.356 -0.571 -0.209 -0.319 -0.368 -0.687
## Cluster_current3rd.cluster:Semester4 0.259 -0.199 -0.542 -0.231 -0.281 -0.507
##                               C_2.:S2 C_3.:S2 C_2.:S3 C_3.:S3 C_2.:S4
## Cluster_current2nd.cluster
## Cluster_current3rd.cluster
## Semester2
## Semester3
## Semester4
## Cluster_current2nd.cluster:Semester2 0.348
## Cluster_current3rd.cluster:Semester2 0.513  0.184
## Cluster_current2nd.cluster:Semester3 0.175  0.526  0.340

```

```

## Cluster_current2nd.cluster:Semester4  0.438   0.156   0.522   0.183
## Cluster_current3rd.cluster:Semester4  0.153   0.454   0.188   0.541   0.345
##
## Standardized Within-Group Residuals:
##      Min       Q1       Med       Q3       Max
## -4.1574050 -0.5197247  0.1874986  0.7667119  1.8952885
##
## Number of Observations: 3802
## Number of Groups: 1140

## Linear mixed-effects model fit by REML
## Data: data.file.long
##      AIC      BIC    logLik
## 9240.016 9289.954 -4612.008
##
## Random effects:
## Formula: ~1 | BARCS_ID
##             (Intercept) Residual
## StdDev: 0.0004825028 0.9948804
##
## Correlation Structure: ARMA(1,1)
## Formula: ~Time | BARCS_ID
## Parameter estimate(s):
##      Phi1     Theta1
## 0.8595063 -0.3876371
## Fixed effects: std_GPA ~ std_Avg_Drinks * std_Avg_MJ
##                  Value  Std.Error DF t-value p-value
## (Intercept) -0.01903332 0.025383017 2659 -0.749845 0.4534
## std_Avg_Drinks -0.07371449 0.019523511 2659 -3.775678 0.0002
## std_Avg_MJ -0.16071856 0.021333075 2659 -7.533774 0.0000
## std_Avg_Drinks:std_Avg_MJ 0.02081323 0.009197667 2659 2.262882 0.0237
## Correlation:
##              (Intr) st_A_D s_A_MJ
## std_Avg_Drinks 0.017
## std_Avg_MJ 0.049 -0.078
## std_Avg_Drinks:std_Avg_MJ -0.132 -0.465 -0.336
##
## Standardized Within-Group Residuals:
##      Min       Q1       Med       Q3       Max
## -4.2536350 -0.5274084  0.1896226  0.7641827  2.0835088
##
## Number of Observations: 3802
## Number of Groups: 1140

## Linear mixed-effects model fit by REML
## Data: data.file.long
##      AIC      BIC    logLik
## 9265.769 9334.424 -4621.884
##
## Random effects:
## Formula: ~1 | BARCS_ID
##             (Intercept) Residual
## StdDev: 0.0007436189 0.9953149
##
## Correlation Structure: ARMA(1,1)

```

```

## Formula: ~Time | BARCS_ID
## Parameter estimate(s):
##      Phi1      Theta1
## 0.8590500 -0.3862338
## Fixed effects: std_GPA ~ std_Avg_Drinks * std_Avg_MJ + std_Avg_Drinks * Time + std_Avg_MJ * Time
##                                     Value Std.Error DF t-value p-value
## (Intercept)           -0.00403700 0.03596490 2656 -0.112248 0.9106
## std_Avg_Drinks        -0.10292445 0.03462650 2656 -2.972418 0.0030
## std_Avg_MJ             -0.13724141 0.03658170 2656 -3.751642 0.0002
## Time                  -0.00697455 0.01143913 2656 -0.609709 0.5421
## std_Avg_Drinks:std_Avg_MJ 0.01949314 0.00936830 2656 2.080756 0.0376
## std_Avg_Drinks:Time    0.01240169 0.01162291 2656 1.067004 0.2861
## std_Avg_MJ:Time        -0.00921548 0.01270849 2656 -0.725143 0.4684
## Correlation:
##                               (Intr) st_A_D st_A_MJ Time   s_A_D:_ s_A_D:T
## std_Avg_Drinks          0.013
## std_Avg_MJ               0.044 -0.345
## Time                   -0.708  0.003 -0.053
## std_Avg_Drinks:std_Avg_MJ -0.134 -0.144 -0.166  0.059
## std_Avg_Drinks:Time     0.055 -0.822  0.379 -0.092 -0.144
## std_Avg_MJ:Time         -0.017  0.342 -0.810  0.048 -0.046 -0.402
##
## Standardized Within-Group Residuals:
##      Min       Q1       Med       Q3       Max
## -4.2531896 -0.5245321  0.1886440  0.7684498  2.0578214
##
## Number of Observations: 3802
## Number of Groups: 1140
## Linear mixed-effects model fit by REML
## Data: data.file.long
##      AIC      BIC      logLik
## 9265.392 9334.047 -4621.696
##
## Random effects:
## Formula: ~1 + Time | BARCS_ID
## Structure: General positive-definite, Log-Cholesky parametrization
##             StdDev   Corr
## (Intercept) 0.8856061 (Intr)
## Time        0.2186274 -0.471
## Residual    0.5699322
##
## Fixed effects: std_GPA ~ std_Avg_Drinks * std_Avg_MJ + std_Avg_Drinks * Time + std_Avg_MJ * Time
##                                     Value Std.Error DF t-value p-value
## (Intercept)           -0.00058688 0.03496095 2656 -0.016787 0.9866
## std_Avg_Drinks        -0.10364462 0.03415540 2656 -3.034501 0.0024
## std_Avg_MJ             -0.14009924 0.03605918 2656 -3.885258 0.0001
## Time                  -0.00967801 0.01163204 2656 -0.832013 0.4055
## std_Avg_Drinks:std_Avg_MJ 0.02020329 0.00940928 2656 2.147166 0.0319
## std_Avg_Drinks:Time    0.01232718 0.01168925 2656 1.054574 0.2917
## std_Avg_MJ:Time        -0.00705841 0.01284187 2656 -0.549640 0.5826
## Correlation:
##                               (Intr) st_A_D st_A_MJ Time   s_A_D:_ s_A_D:T
## std_Avg_Drinks          0.012

```

```

## std_Avg_MJ          0.047 -0.350
## Time                 -0.685  0.004 -0.054
## std_Avg_Drinks:std_Avg_MJ -0.140 -0.145 -0.166   0.059
## std_Avg_Drinks:Time      0.062 -0.817  0.382 -0.095 -0.141
## std_Avg_MJ:Time        -0.018  0.342 -0.804   0.047 -0.049 -0.400
##
## Standardized Within-Group Residuals:
##           Min         Q1         Med         Q3         Max
## -4.63535095 -0.40385478  0.09032549  0.50664820  3.53275927
##
## Number of Observations: 3802
## Number of Groups: 1140
AIC_values.sm.std.r2 <- AIC(sm.nlme.std.robust2.1, sm.nlme.std.robust2.2, sm.nlme.std.ARMA11, sm.nlme.s

## Warning in AIC.default(sm.nlme.std.robust2.1, sm.nlme.std.robust2.2,
## sm.nlme.std.ARMA11, : models are not all fitted to the same number of
## observations

BIC_values.sm.std.r2 <- BIC(sm.nlme.std.robust2.1, sm.nlme.std.robust2.2, sm.nlme.std.ARMA11, sm.nlme.s

## Warning in BIC.default(sm.nlme.std.robust2.1, sm.nlme.std.robust2.2,
## sm.nlme.std.ARMA11, : models are not all fitted to the same number of
## observations

df_scores_sm.std.r2 <- data.frame(AIC_values.sm.std.r2, BIC_values.sm.std.r2)
df_scores_sm.std.r2

##                               df      AIC df.1      BIC
## sm.nlme.std.robust2.1    15 9378.350  15 9471.948
## sm.nlme.std.robust2.2    17 9300.111  17 9406.189
## sm.nlme.std.ARMA11       16 9307.691  16 9407.533
## sm.nlme.std.noSemester   8 9240.016   8 9289.954
## sm.nlme.std.noSemester.time 11 9265.769  11 9334.424
## sm.nlme.std.noSemester.timeRE 11 9265.392  11 9334.047

no time / Semester > Time > Semester. Semester Pretty bad. The same holds for the full model.

sm.outlier.barebones <- lme(std_GPA ~ std_Avg_Drinks + std_Avg_MJ, random = ~ 1 | BARCS_ID, data = outlier)

sm.outlier <- lme(std_GPA ~ std_Avg_Drinks * std_Avg_MJ, random = ~ 1 | BARCS_ID, data = outlier.removed)

sm.outlier.semester <- lme(std_GPA ~ std_Avg_Drinks * Semester + std_Avg_Drinks * std_Avg_MJ + std_Avg_MJ * Semester, random = ~ 1 | BARCS_ID, data = outlier.semester)

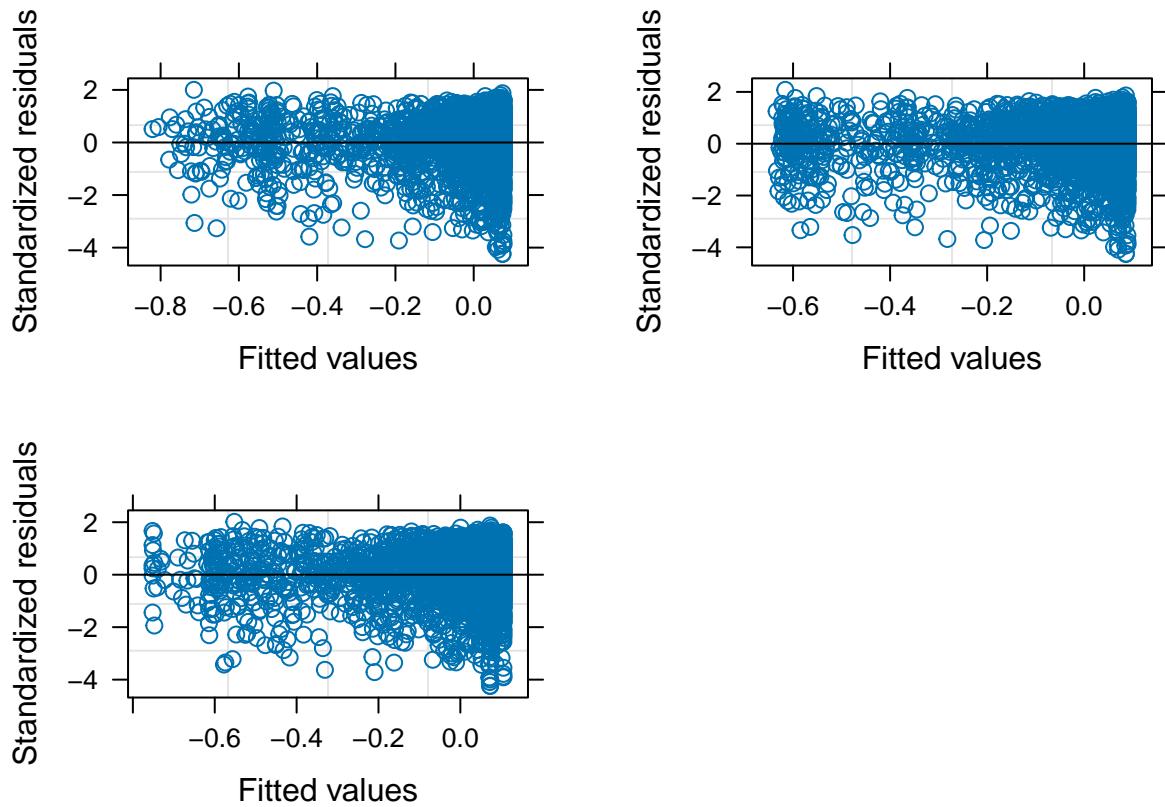
fm.outlier.barebones <- lme(std_GPA ~ std_Avg_Drinks + std_Avg_MJ + Sex + Age1stround + SATMath + SATVerbal + SATWriting, random = ~ 1 | BARCS_ID, data = outlier)

fm.outlier <- lme(std_GPA ~ std_Avg_Drinks * std_Avg_MJ + Sex + Age1stround + SATMath + SATVerbal + SATWriting, random = ~ 1 | BARCS_ID, data = outlier)

fm.outlier.semester <- lme(std_GPA ~ std_Avg_Drinks * Semester + std_Avg_Drinks * std_Avg_MJ + std_Avg_MJ * Semester, random = ~ 1 | BARCS_ID, data = outlier.semester)

cowplot::plot_grid(plot(sm.outlier.barebones), plot(sm.outlier), plot(sm.outlier.semester))

```



```
summary(sm.outlier.barebones)
```

```
## Linear mixed-effects model fit by maximum likelihood
##   Data: outlier.removed
##       AIC     BIC   logLik
##   9102.64 9146.258 -4544.32
##
## Random effects:
##   Formula: ~1 | BARCS_ID
##             (Intercept) Residual
## StdDev: 0.0007762924 0.9977145
##
## Correlation Structure: ARMA(1,1)
##   Formula: ~Time | BARCS_ID
## Parameter estimate(s):
##   Phi1    Theta1
## 0.8579115 -0.3780934
## Fixed effects: std_GPA ~ std_Avg_Drinks + std_Avg_MJ
##                  Value Std.Error DF t-value p-value
## (Intercept) -0.01264347 0.02538871 2616 -0.497995 0.6185
## std_Avg_Drinks -0.04735407 0.01693829 2616 -2.795681 0.0052
## std_Avg_MJ -0.13872714 0.01994380 2616 -6.955904 0.0000
## Correlation:
##   (Intr) st_A_D
## std_Avg_Drinks -0.066
## std_Avg_MJ      0.001 -0.270
##
## Standardized Within-Group Residuals:
```

```

##      Min       Q1       Med       Q3       Max
## -4.2443825 -0.5179469  0.1871437  0.7696898  2.0026611
##
## Number of Observations: 3756
## Number of Groups: 1138
summary(sm.outlier)

## Linear mixed-effects model fit by maximum likelihood
## Data: outlier.removed
##      AIC      BIC   logLik
## 9101.108 9150.957 -4542.554
##
## Random effects:
## Formula: ~1 | BARCS_ID
##             (Intercept) Residual
## StdDev: 0.0007558861 0.9975136
##
## Correlation Structure: ARMA(1,1)
## Formula: ~Time | BARCS_ID
## Parameter estimate(s):
##     Phi1    Theta1
## 0.8582221 -0.3784458
## Fixed effects: std_GPA ~ std_Avg_Drinks * std_Avg_MJ
##                  Value Std.Error DF t-value p-value
## (Intercept)      -0.01974456 0.02566993 2615 -0.769171 0.4419
## std_Avg_Drinks  -0.05932821 0.01809179 2615 -3.279289 0.0011
## std_Avg_MJ       -0.15514073 0.02177872 2615 -7.123500 0.0000
## std_Avg_Drinks:std_Avg_MJ 0.01962392 0.01044398 2615  1.878970 0.0604
## Correlation:
##              (Intr) st_A_D s_A_MJ
## std_Avg_Drinks -0.010
## std_Avg_MJ      0.060 -0.089
## std_Avg_Drinks:std_Avg_MJ -0.147 -0.352 -0.402
##
## Standardized Within-Group Residuals:
##      Min       Q1       Med       Q3       Max
## -4.2577059 -0.5217214  0.1901809  0.7682227  2.0789886
##
## Number of Observations: 3756
## Number of Groups: 1138
summary(sm.outlier.semester)

## Linear mixed-effects model fit by maximum likelihood
## Data: outlier.removed
##      AIC      BIC   logLik
## 9115.009 9220.938 -4540.505
##
## Random effects:
## Formula: ~1 | BARCS_ID
##             (Intercept) Residual
## StdDev: 0.001521384 0.9976363
##
## Correlation Structure: ARMA(1,1)

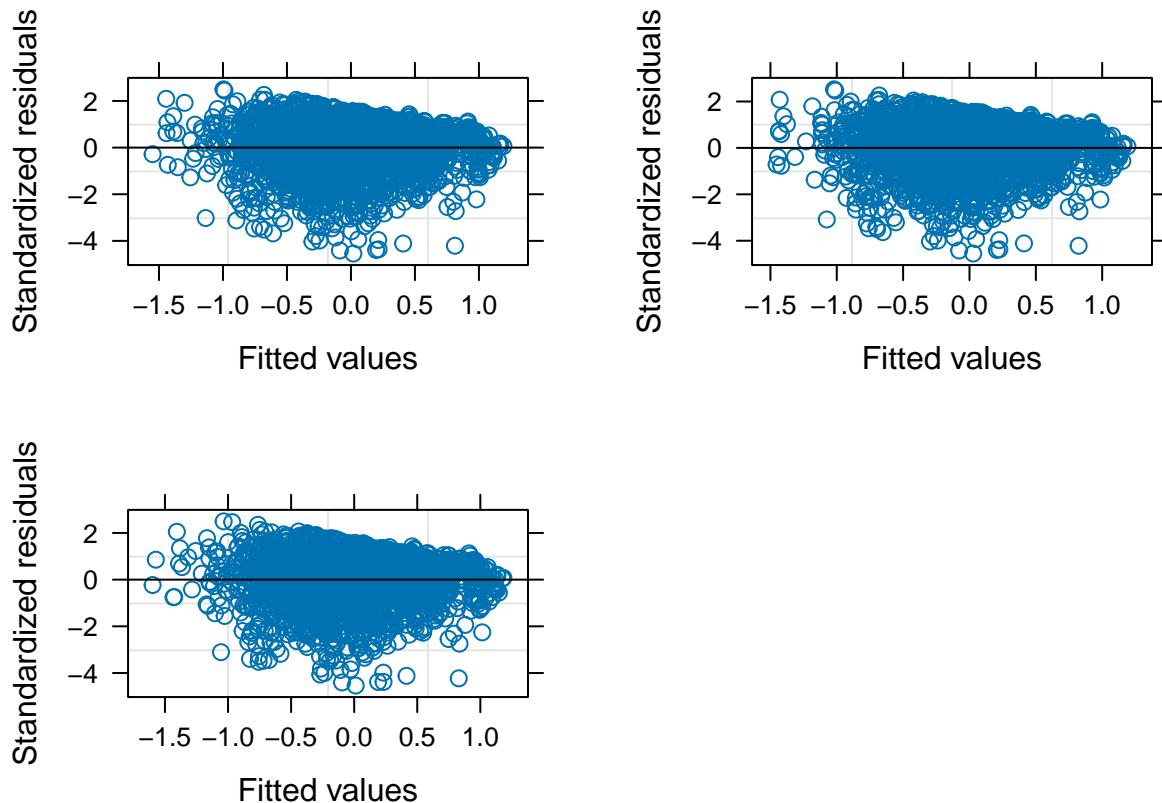
```

```

## Formula: ~Time | BARCS_ID
## Parameter estimate(s):
##      Phi1      Theta1
## 0.8582922 -0.3776696
## Fixed effects: std_GPA ~ std_Avg_Drinks * Semester + std_Avg_Drinks * std_Avg_MJ +
##                                         Value Std.Error DF t-value p-value
## (Intercept)           -0.01092126 0.03011496 2606 -0.362652 0.7169
## std_Avg_Drinks        -0.06989461 0.02718331 2606 -2.571232 0.0102
## Semester2             -0.02087550 0.02506876 2606 -0.832730 0.4051
## Semester3             -0.00464759 0.03107674 2606 -0.149552 0.8811
## Semester4             -0.01969896 0.03586045 2606 -0.549323 0.5828
## std_Avg_MJ            -0.15360563 0.02955179 2606 -5.197846 0.0000
## std_Avg_Drinks:Semester2 0.00136962 0.02984079 2606 0.045898 0.9634
## std_Avg_Drinks:Semester3 0.02559479 0.03276464 2606 0.781171 0.4348
## std_Avg_Drinks:Semester4 0.00847614 0.03649883 2606 0.232230 0.8164
## std_Avg_Drinks:std_Avg_MJ 0.02184203 0.01069853 2606 2.041592 0.0413
## Semester2:std_Avg_MJ    0.01837749 0.02955321 2606 0.621844 0.5341
## Semester3:std_Avg_MJ    -0.00199076 0.03471046 2606 -0.057353 0.9543
## Semester4:std_Avg_MJ    -0.03171829 0.03959044 2606 -0.801160 0.4231
## Correlation:
##                               (Intr) st_A_D Smstr2 Smstr3 Smstr4 s_A_MJ s_A_D:S2
## std_Avg_Drinks           0.021
## Semester2                -0.411  0.001
## Semester3                -0.425 -0.045  0.510
## Semester4                -0.435 -0.040  0.427  0.563
## std_Avg_MJ                0.047 -0.287 -0.008 -0.037 -0.041
## std_Avg_Drinks:Semester2 0.004 -0.568  0.009 -0.013 -0.008  0.247
## std_Avg_Drinks:Semester3 0.007 -0.648  0.004 -0.078 -0.022  0.269  0.547
## std_Avg_Drinks:Semester4 0.017 -0.619  0.003 -0.011 -0.120  0.272  0.474
## std_Avg_Drinks:std_Avg_MJ -0.145 -0.163 -0.003  0.047  0.050 -0.289 -0.066
## Semester2:std_Avg_MJ     -0.009  0.259  0.014  0.009  0.005 -0.536 -0.442
## Semester3:std_Avg_MJ     -0.004  0.271  0.009  0.045  0.013 -0.581 -0.224
## Semester4:std_Avg_MJ     0.008  0.262  0.007  0.012  0.057 -0.536 -0.182
##                               s_A_D:S3 s_A_D:S4 s_A_D:_ S2:_A_ S3:_A_
## std_Avg_Drinks
## Semester2
## Semester3
## Semester4
## std_Avg_MJ
## std_Avg_Drinks:Semester2
## std_Avg_Drinks:Semester3
## std_Avg_Drinks:Semester4  0.596
## std_Avg_Drinks:std_Avg_MJ -0.072  -0.123
## Semester2:std_Avg_MJ     -0.239  -0.210   0.033
## Semester3:std_Avg_MJ     -0.383  -0.236   0.005   0.521
## Semester4:std_Avg_MJ     -0.220  -0.377  -0.085   0.435  0.567
## Standardized Within-Group Residuals:
##      Min       Q1       Med       Q3       Max
## -4.2438666 -0.5226909  0.1922433  0.7702130  2.0147415
## 
## Number of Observations: 3756
## Number of Groups: 1138

```

```
cowplot::plot_grid(plot(fm.outlier.barebones), plot(fm.outlier), plot(fm.outlier.semester))
```



```
summary(sm.outlier.barebones)
```

```
## Linear mixed-effects model fit by maximum likelihood
##   Data: outlier.removed
##         AIC     BIC   logLik
##   9102.64 9146.258 -4544.32
##
## Random effects:
##   Formula: ~1 | BARCS_ID
##             (Intercept) Residual
## StdDev: 0.0007762924 0.9977145
##
## Correlation Structure: ARMA(1,1)
##   Formula: ~Time | BARCS_ID
##   Parameter estimate(s):
##     Phi1     Theta1
## 0.8579115 -0.3780934
## Fixed effects: std_GPA ~ std_Avg_Drinks + std_Avg_MJ
##                  Value Std.Error DF t-value p-value
## (Intercept) -0.01264347 0.02538871 2616 -0.497995 0.6185
## std_Avg_Drinks -0.04735407 0.01693829 2616 -2.795681 0.0052
## std_Avg_MJ    -0.13872714 0.01994380 2616 -6.955904 0.0000
## Correlation:
##   (Intr) st_A_D
## std_Avg_Drinks -0.066
## std_Avg_MJ      0.001 -0.270
```

```

## 
## Standardized Within-Group Residuals:
##      Min       Q1       Med       Q3       Max
## -4.2443825 -0.5179469  0.1871437  0.7696898  2.0026611
## 
## Number of Observations: 3756
## Number of Groups: 1138
summary(sm.outlier)

## Linear mixed-effects model fit by maximum likelihood
## Data: outlier.removed
##      AIC      BIC    logLik
## 9101.108 9150.957 -4542.554
##
## Random effects:
## Formula: ~1 | BARCS_ID
##             (Intercept) Residual
## StdDev: 0.0007558861 0.9975136
##
## Correlation Structure: ARMA(1,1)
## Formula: ~Time | BARCS_ID
## Parameter estimate(s):
##      Phi1     Theta1
## 0.8582221 -0.3784458
## Fixed effects: std_GPA ~ std_Avg_Drinks * std_Avg_MJ
##                  Value Std.Error DF t-value p-value
## (Intercept)      -0.01974456 0.02566993 2615 -0.769171 0.4419
## std_Avg_Drinks   -0.05932821 0.01809179 2615 -3.279289 0.0011
## std_Avg_MJ        -0.15514073 0.02177872 2615 -7.123500 0.0000
## std_Avg_Drinks:std_Avg_MJ 0.01962392 0.01044398 2615  1.878970 0.0604
## Correlation:
##              (Intr) st_A_D s_A_MJ
## std_Avg_Drinks -0.010
## std_Avg_MJ      0.060 -0.089
## std_Avg_Drinks:std_Avg_MJ -0.147 -0.352 -0.402
##
## Standardized Within-Group Residuals:
##      Min       Q1       Med       Q3       Max
## -4.2577059 -0.5217214  0.1901809  0.7682227  2.0789886
## 
## Number of Observations: 3756
## Number of Groups: 1138
summary(sm.outlier.semester)

## Linear mixed-effects model fit by maximum likelihood
## Data: outlier.removed
##      AIC      BIC    logLik
## 9115.009 9220.938 -4540.505
##
## Random effects:
## Formula: ~1 | BARCS_ID
##             (Intercept) Residual
## StdDev: 0.001521384 0.9976363

```

```

##
## Correlation Structure: ARMA(1,1)
## Formula: ~Time | BARCS_ID
## Parameter estimate(s):
##      Phi1      Theta1
## 0.8582922 -0.3776696
## Fixed effects: std_GPA ~ std_Avg_Drinks * Semester + std_Avg_Drinks * std_Avg_MJ +
##                                     Value Std.Error DF t-value p-value
## (Intercept) -0.01092126 0.03011496 2606 -0.362652 0.7169
## std_Avg_Drinks -0.06989461 0.02718331 2606 -2.571232 0.0102
## Semester2 -0.02087550 0.02506876 2606 -0.832730 0.4051
## Semester3 -0.00464759 0.03107674 2606 -0.149552 0.8811
## Semester4 -0.01969896 0.03586045 2606 -0.549323 0.5828
## std_Avg_MJ -0.15360563 0.02955179 2606 -5.197846 0.0000
## std_Avg_Drinks:Semester2 0.00136962 0.02984079 2606 0.045898 0.9634
## std_Avg_Drinks:Semester3 0.02559479 0.03276464 2606 0.781171 0.4348
## std_Avg_Drinks:Semester4 0.00847614 0.03649883 2606 0.232230 0.8164
## std_Avg_Drinks:std_Avg_MJ 0.02184203 0.01069853 2606 2.041592 0.0413
## Semester2:std_Avg_MJ 0.01837749 0.02955321 2606 0.621844 0.5341
## Semester3:std_Avg_MJ -0.00199076 0.03471046 2606 -0.057353 0.9543
## Semester4:std_Avg_MJ -0.03171829 0.03959044 2606 -0.801160 0.4231
## Correlation:
##                               (Intr) st_A_D Smstr2 Smstr3 Smstr4 s_A_MJ s_A_D:S2
## std_Avg_Drinks           0.021
## Semester2                 -0.411  0.001
## Semester3                 -0.425 -0.045  0.510
## Semester4                 -0.435 -0.040  0.427  0.563
## std_Avg_MJ                  0.047 -0.287 -0.008 -0.037 -0.041
## std_Avg_Drinks:Semester2  0.004 -0.568  0.009 -0.013 -0.008  0.247
## std_Avg_Drinks:Semester3  0.007 -0.648  0.004 -0.078 -0.022  0.269  0.547
## std_Avg_Drinks:Semester4  0.017 -0.619  0.003 -0.011 -0.120  0.272  0.474
## std_Avg_Drinks:std_Avg_MJ -0.145 -0.163 -0.003  0.047  0.050 -0.289 -0.066
## Semester2:std_Avg_MJ      -0.009  0.259  0.014  0.009  0.005 -0.536 -0.442
## Semester3:std_Avg_MJ      -0.004  0.271  0.009  0.045  0.013 -0.581 -0.224
## Semester4:std_Avg_MJ      0.008  0.262  0.007  0.012  0.057 -0.536 -0.182
##                               s_A_D:S3 s_A_D:S4 s_A_D:_ S2:_A_ S3:_A_
## std_Avg_Drinks
## Semester2
## Semester3
## Semester4
## std_Avg_MJ
## std_Avg_Drinks:Semester2
## std_Avg_Drinks:Semester3
## std_Avg_Drinks:Semester4  0.596
## std_Avg_Drinks:std_Avg_MJ -0.072   -0.123
## Semester2:std_Avg_MJ     -0.239   -0.210   0.033
## Semester3:std_Avg_MJ     -0.383   -0.236   0.005   0.521
## Semester4:std_Avg_MJ     -0.220   -0.377   -0.085  0.435   0.567
##
## Standardized Within-Group Residuals:
##      Min        Q1        Med        Q3        Max
## -4.2438666 -0.5226909  0.1922433  0.7702130  2.0147415
##
## Number of Observations: 3756

```

```

## Number of Groups: 1138
AIC_sm.outlier <- AIC(sm.outlier.barebones, sm.outlier, sm.outlier.semester, fm.outlier.barebones, fm.o

## Warning in AIC.default(sm.outlier.barebones, sm.outlier, sm.outlier.semester, :
## models are not all fitted to the same number of observations
BIC_sm.outlier <- BIC(sm.outlier.barebones, sm.outlier, sm.outlier.semester, fm.outlier.barebones, fm.o

## Warning in BIC.default(sm.outlier.barebones, sm.outlier, sm.outlier.semester, :
## models are not all fitted to the same number of observations
df.sm.out <- data.frame(AIC_sm.outlier, BIC_sm.outlier)
df.sm.out

##          df      AIC df.1      BIC
## sm.outlier.barebones 7 9102.640    7 9146.258
## sm.outlier           8 9101.108    8 9150.957
## sm.outlier.semester 17 9115.009   17 9220.938
## fm.outlier.barebones 17 7832.435   17 7936.271
## fm.outlier           18 7832.599   18 7942.543
## fm.outlier.semester 27 7845.024   27 8009.941

## probably useful to just make a table with the relevant coefficient values of the models
Anova(sm.outlier, white.adjust = "hc3", correction = "Sidak")

## Analysis of Deviance Table (Type II tests)
##
## Response: std_GPA
##                  Chisq Df Pr(>Chisq)
## std_Avg_Drinks      7.8301  1  0.005138 **
## std_Avg_MJ          48.4266  1  3.429e-12 ***
## std_Avg_Drinks:std_Avg_MJ 3.5343  1  0.060112 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
anova(sm.outlier, sm.outlier.barebones)

##          Model df      AIC      BIC logLik  Test L.Ratio
## sm.outlier        1 8 9101.108 9150.957 -4542.554
## sm.outlier.barebones 2 7 9102.640 9146.258 -4544.320 1 vs 2 3.531973
## p-value
## sm.outlier
## sm.outlier.barebones 0.0602
Anova(sm.outlier.semester, white.adjust = "hc3", correction = "Sidak")

## Analysis of Deviance Table (Type II tests)
##
## Response: std_GPA
##                  Chisq Df Pr(>Chisq)
## std_Avg_Drinks      7.4732  1  0.006262 **
## Semester            1.1061  3  0.775610
## std_Avg_MJ          47.9706  1  4.327e-12 ***
## std_Avg_Drinks:Semester 0.8591  3  0.835296
## std_Avg_Drinks:std_Avg_MJ 4.1826  1  0.040842 *
## Semester:std_Avg_MJ    1.8171  3  0.611224
## ---

```

```

## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
anova(sm.outlier.semester, sm.outlier.barebones)

##                               Model df      AIC      BIC    logLik   Test L.Ratio
## sm.outlier.semester      1 17 9115.009 9220.938 -4540.505
## sm.outlier.barebones     2  7 9102.640 9146.258 -4544.320 1 vs 2 7.631014
##                               p-value
## sm.outlier.semester
## sm.outlier.barebones  0.6648
anova(sm.outlier, sm.outlier.semester)

##                               Model df      AIC      BIC    logLik   Test L.Ratio
## sm.outlier           1  8 9101.108 9150.957 -4542.554
## sm.outlier.semester  2 17 9115.009 9220.938 -4540.505 1 vs 2 4.099041
##                               p-value
## sm.outlier
## sm.outlier.semester  0.9048
Anova(fm.outlier, white.adjust = "hc3", correction = "Sidak")

## Analysis of Deviance Table (Type II tests)
##
## Response: std_GPA
##                               Chisq Df Pr(>Chisq)
## std_Avg_Drinks          11.1678 1  0.0008323 ***
## std_Avg_MJ                43.1579 1  5.050e-11 ***
## Sex                      15.0154 1  0.0001066 ***
## Age1stround              3.7010 1  0.0543808 .
## SATMath                  27.6641 1  1.443e-07 ***
## SATVerbal                 1.5898 1  0.2073574
## SATWriting                19.4864 1  1.013e-05 ***
## Fager4_binary             5.0731 1  0.0242997 *
## FH_binary                 2.2826 1  0.1308286
## STAI_SELF_Total           0.7438 1  0.3884593
## BDI_SELF_Total            4.9218 1  0.0265194 *
## Parental_SES              4.8808 1  0.0271577 *
## std_Avg_Drinks:std_Avg_MJ 1.8381 1  0.1751755
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
anova(fm.outlier, fm.outlier.barebones)

##                               Model df      AIC      BIC    logLik   Test L.Ratio
## fm.outlier           1 18 7832.599 7942.543 -3898.299
## fm.outlier.barebones  2 17 7832.435 7936.271 -3899.217 1 vs 2 1.835931
##                               p-value
## fm.outlier
## fm.outlier.barebones  0.1754
Anova(fm.outlier.semester, white.adjust = "hc3", correction = "Sidak")

## Analysis of Deviance Table (Type II tests)
##
## Response: std_GPA
##                               Chisq Df Pr(>Chisq)
## std_Avg_Drinks          11.3512 1  0.000754 ***

```

```

## Semester          0.9823  3   0.805524
## std_Avg_MJ      42.7801  1   6.125e-11 ***
## Sex              14.8239  1   0.000118 ***
## Age1stround     3.6197  1   0.057100 .
## SATMath          27.6019  1   1.490e-07 ***
## SATVerbal        1.4960  1   0.221288
## SATWriting       19.6764  1   9.172e-06 ***
## Fager4_binary    5.1421  1   0.023353 *
## FH_binary        2.2182  1   0.136395
## STAI_SELF_Total  0.7471  1   0.387389
## BDI_SELF_Total   4.9600  1   0.025940 *
## Parental_SES     4.6396  1   0.031243 *
## std_Avg_Drinks:Semester 1.9065  3   0.592036
## std_Avg_Drinks:std_Avg_MJ 2.5800  1   0.108219
## Semester:std_Avg_MJ    1.7775  3   0.619845
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
anova(fm.outlier.semester, fm.outlier.barebones)

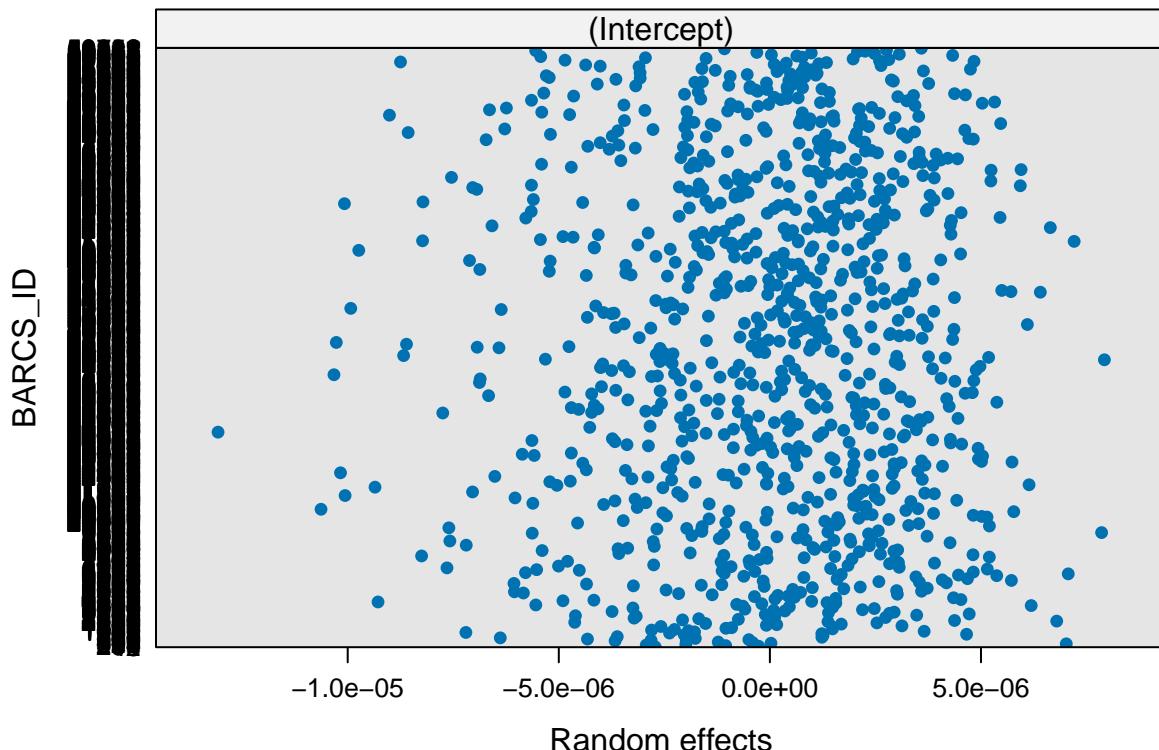
##                               Model df      AIC      BIC    logLik   Test L.Ratio
## fm.outlier.semester      1 27 7845.024 8009.941 -3895.512
## fm.outlier.barebones     2 17 7832.435 7936.271 -3899.217 1 vs 2 7.410643
##                               p-value
## fm.outlier.semester
## fm.outlier.barebones  0.6862
anova(fm.outlier, fm.outlier.semester)

##                               Model df      AIC      BIC    logLik   Test L.Ratio
## fm.outlier           1 18 7832.599 7942.543 -3898.299
## fm.outlier.semester  2 27 7845.024 8009.941 -3895.512 1 vs 2 5.574712
##                               p-value
## fm.outlier
## fm.outlier.semester  0.7816

inclusion of the Interaction Effect leads to an stronger negative coefficients of the substance use parameters, but the interaction is always positive and not statistically significant. Models perform slightly better with the interaction effect, but the LRT test still rejects the inclusion for a pvalue =.05
plot(ranef(fm.outlier),
      main = "Random Intercept")

```

Random Intercept



```
dim(ranef(fm.outlier))

## [1] 1002      1
print(range(ranef(fm.outlier)))

## [1] -1.307062e-05  7.922518e-06
print(r.squaredGLMM(fm.outlier))

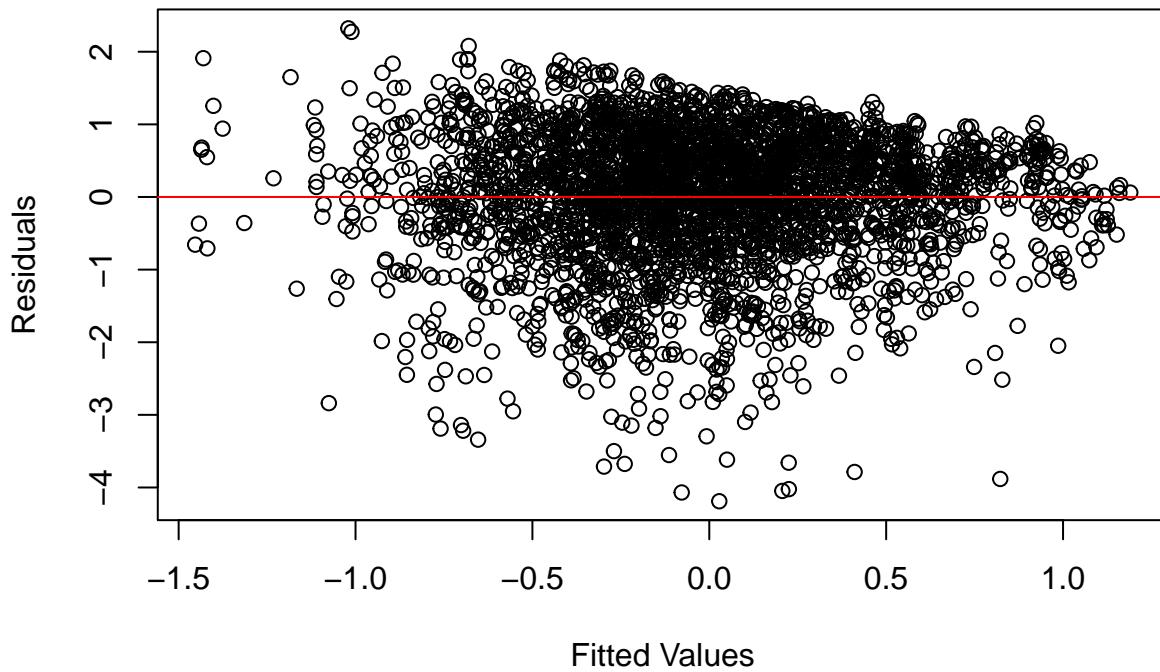
## Warning: 'r.squaredGLMM' now calculates a revised statistic. See the help page.
##          R2m          R2c
## [1,] 0.181672 0.1816743
```

of course not.

Residuals:

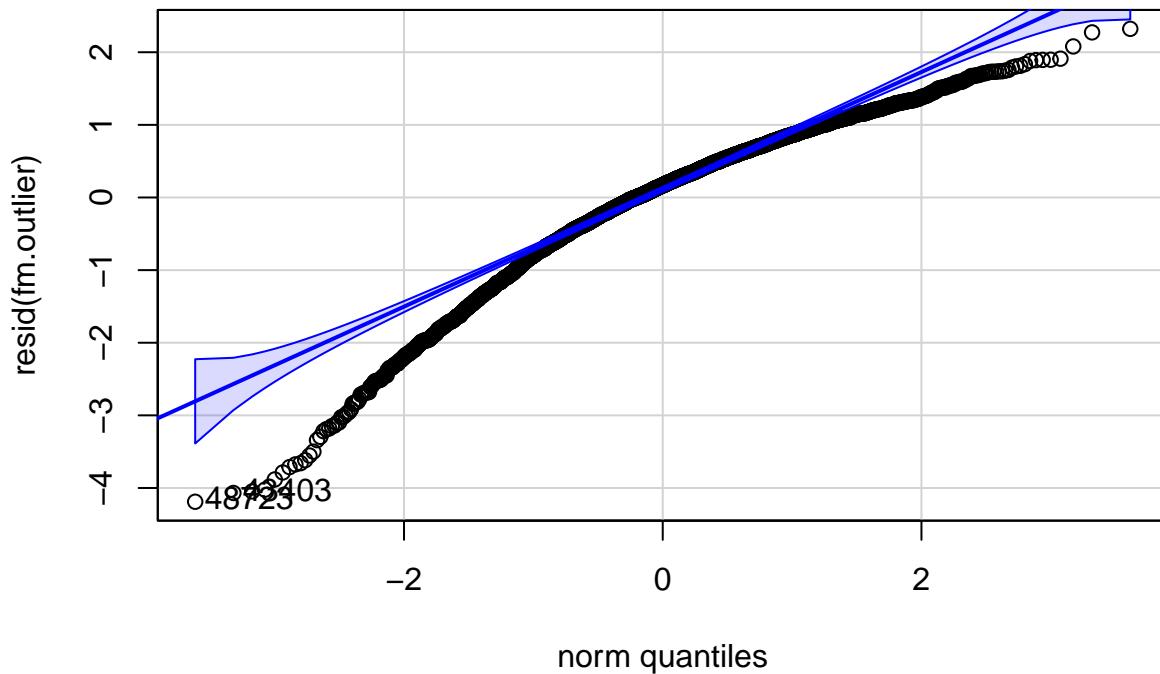
```
plot(fitted(fm.outlier), resid(fm.outlier),
     xlab = "Fitted Values",
     ylab = "Residuals",
     main = "Residuals vs Fitted Plot")
abline(h = 0, col = "red")
```

Residuals vs Fitted Plot



```
qqPlot(resid(fm.outlier),  
       main = "Normal Q-Q Plot of Residuals")
```

Normal Q-Q Plot of Residuals



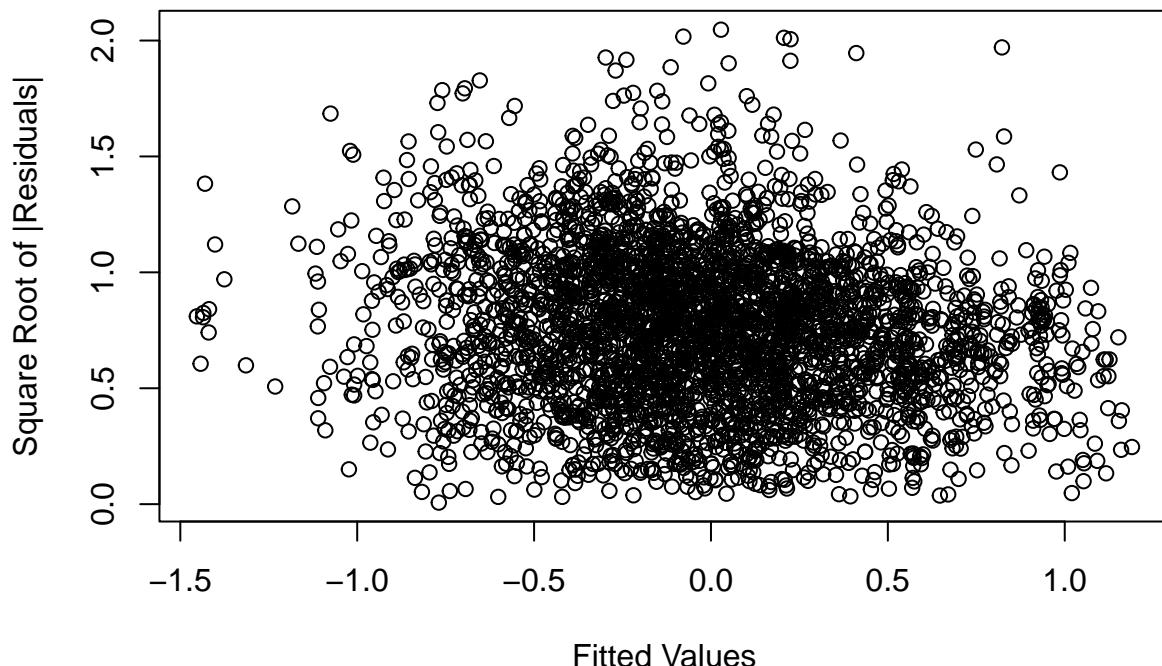
```
## 48723 43403  
## 1400 1394
```

```

sqrt_abs_resid <- sqrt(abs(resid(fm.outlier)))
plot(fitted(fm.outlier), sqrt_abs_resid,
      xlab = "Fitted Values",
      ylab = "Square Root of |Residuals|",
      main = "Scale-Location Plot")

```

Scale-Location Plot

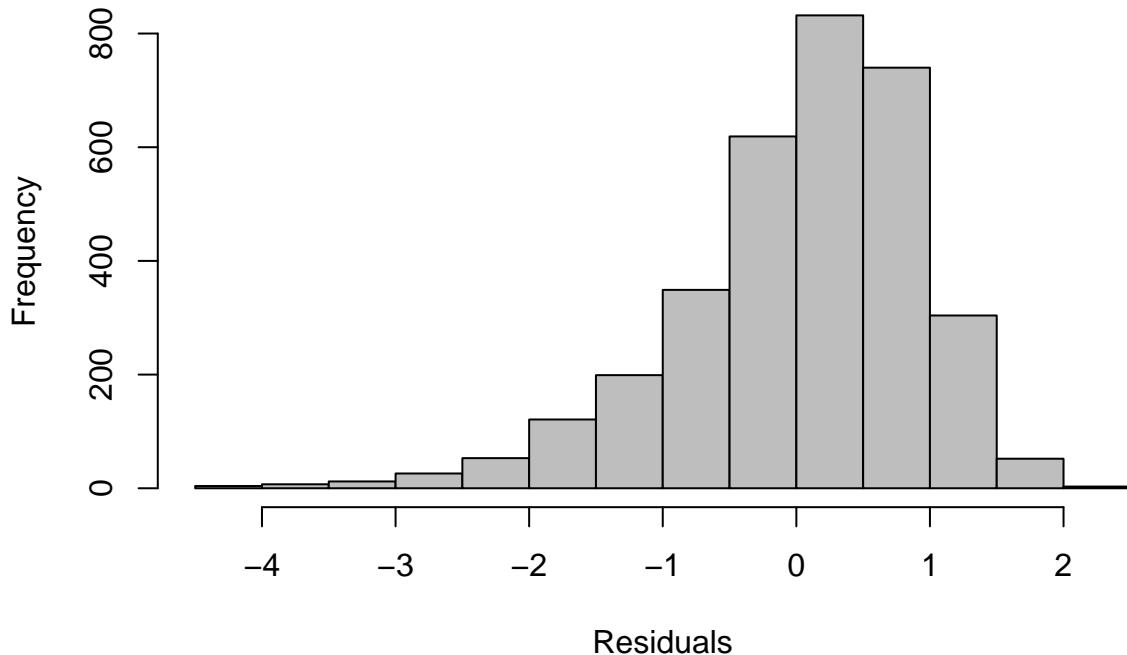


```

hist(resid(fm.outlier),
     main = "Histogram of Residuals",
     xlab = "Residuals",
     breaks = 10,
     col = "gray")

```

Histogram of Residuals



The introduction of the covariates might help with the Random Effects.

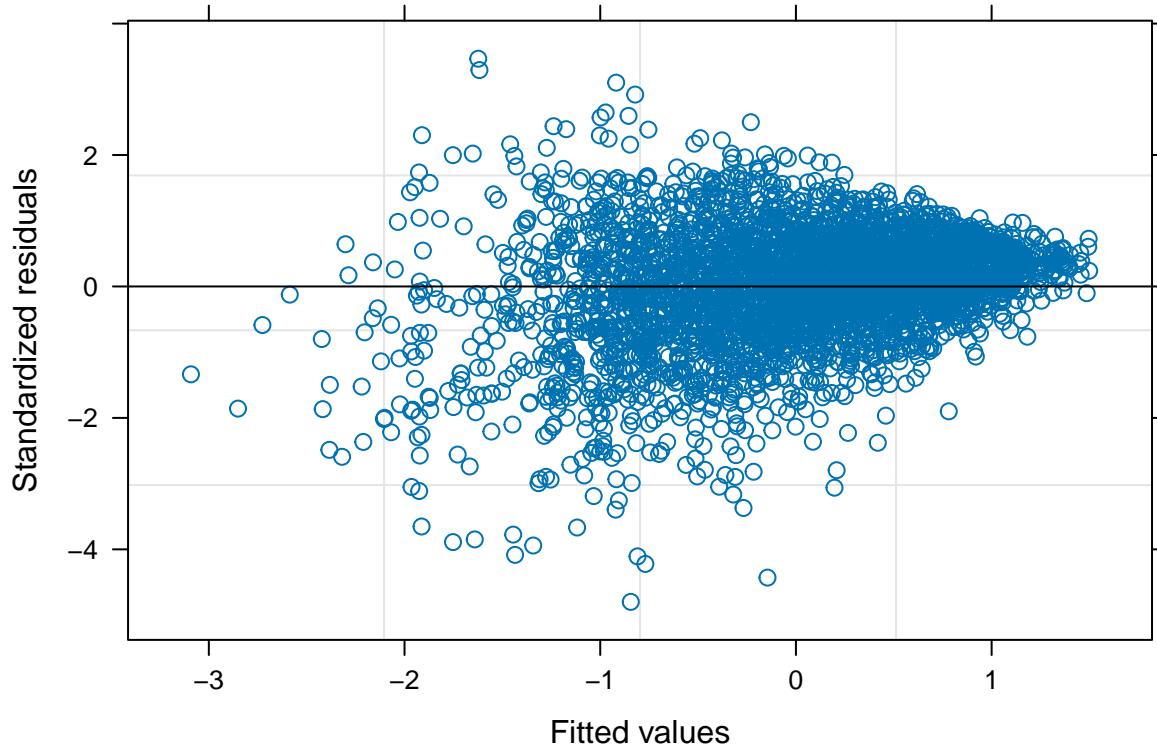
```
sm.outlier.barebones <- lme(std_GPA ~ std_Avg_Drinks + std_Avg_MJ, random = ~ 1 | BARCS_ID, data = outliers)

sm.outlier.barebones.2 <- lme(std_GPA ~ std_Avg_Drinks + std_Avg_MJ, random = ~ 1 + std_Avg_Drinks + std_Avg_MJ)

#sm.outlier.barebones.3 <- lme(std_GPA ~ std_Avg_Drinks + std_Avg_MJ, random = ~ 1 + std_Avg_Drinks + std_Avg_MJ)

anova(sm.outlier.barebones, sm.outlier.barebones.2)

##                                     Model df      AIC      BIC    logLik   Test  L.Ratio
## sm.outlier.barebones       1  7 9120.540 9164.152 -4553.270
## sm.outlier.barebones.2     2 10 9179.471 9241.774 -4579.736 1 vs 2 52.93075
##                                     p-value
## sm.outlier.barebones
## sm.outlier.barebones.2 <.0001
plot(sm.outlier.barebones.2)
```



```

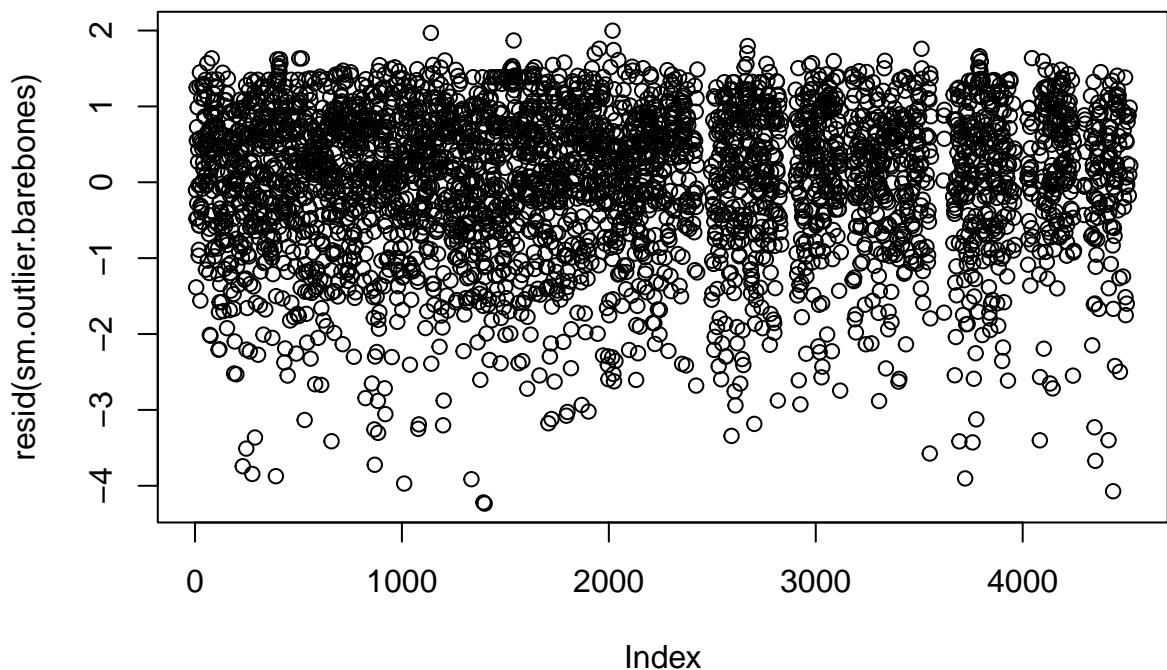
sm.outlier.barebones <- lme(std_GPA ~ std_Avg_Drinks + std_Avg_MJ, random = ~ 1 | BARCS_ID, data = outlier)

sm.outlier.barebones.2 <- lme(std_GPA ~ std_Avg_Drinks + std_Avg_MJ, random = ~ 1 + std_Avg_Drinks + std_Avg_MJ, data = outlier)

AIC(sm.outlier.barebones, sm.outlier.barebones.2)

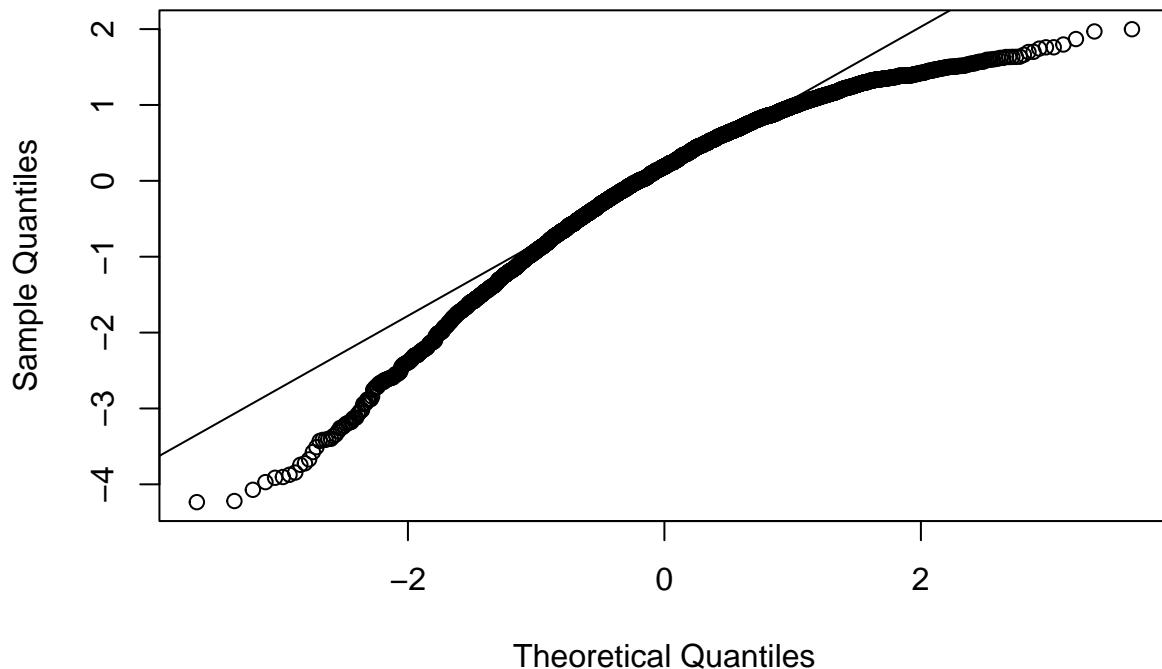
##                      df      AIC
## sm.outlier.barebones    7 9102.640
## sm.outlier.barebones.2 10 9162.055
plot(resid(sm.outlier.barebones))

```

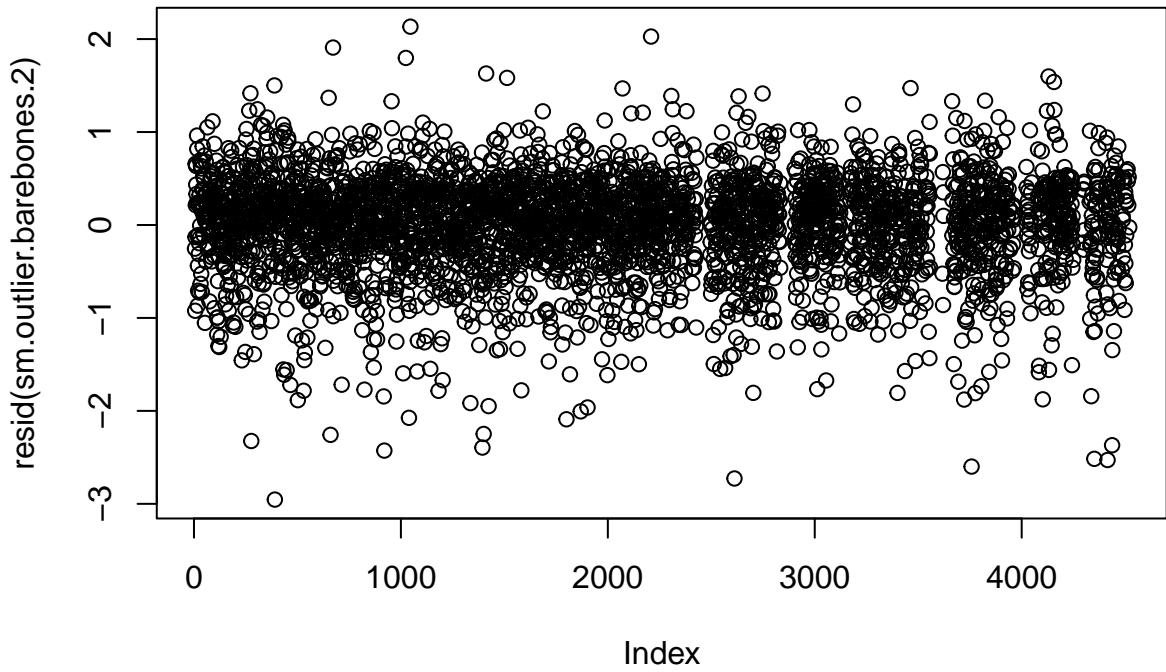


```
qqnorm(resid(sm.outlier.barebones))  
qqline(resid(sm.outlier.barebones))
```

Normal Q-Q Plot

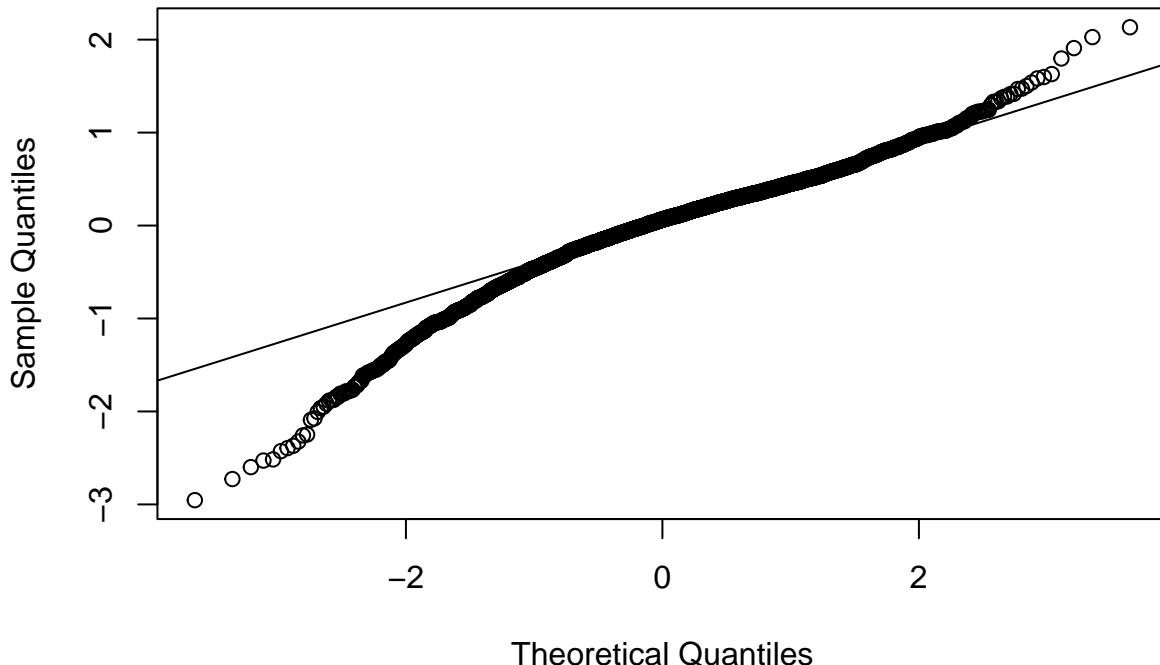


```
plot(resid(sm.outlier.barebones).2)
```



```
qqnorm(resid(sm.outlier.barebones.2))
qqline(resid(sm.outlier.barebones.2))
```

Normal Q-Q Plot



is problematic, they aren't really nested.

This

```
#diff.robust <- lme(diff_GPA ~ diff_Avg_Drinks_current * diff_Avg_MJ_current * I(Time - 1) + factor(BARCS)

diff.robust.2<- lme(diff_GPA ~ diff_Avg_Drinks_current * diff_Avg_MJ_current * I(Time - 1), random = ~ 1)
```

```

diff.robust.3 <- lm(diff_GPA ~ diff_Avg_Drinks_current * diff_Avg_MJ_current * I(Time -1) +factor(BARCS_ID))

diff.robust.4<- lme(diff_GPA ~ diff_Avg_Drinks_current * diff_Avg_MJ_current, random = ~ 0 + diff_Avg_Drinks_current | BARCS_ID)

diff.robust.5<- lme(diff_GPA ~ diff_Avg_Drinks_current + diff_Avg_MJ_current, random = ~ 0 + diff_Avg_Drinks_current | BARCS_ID)

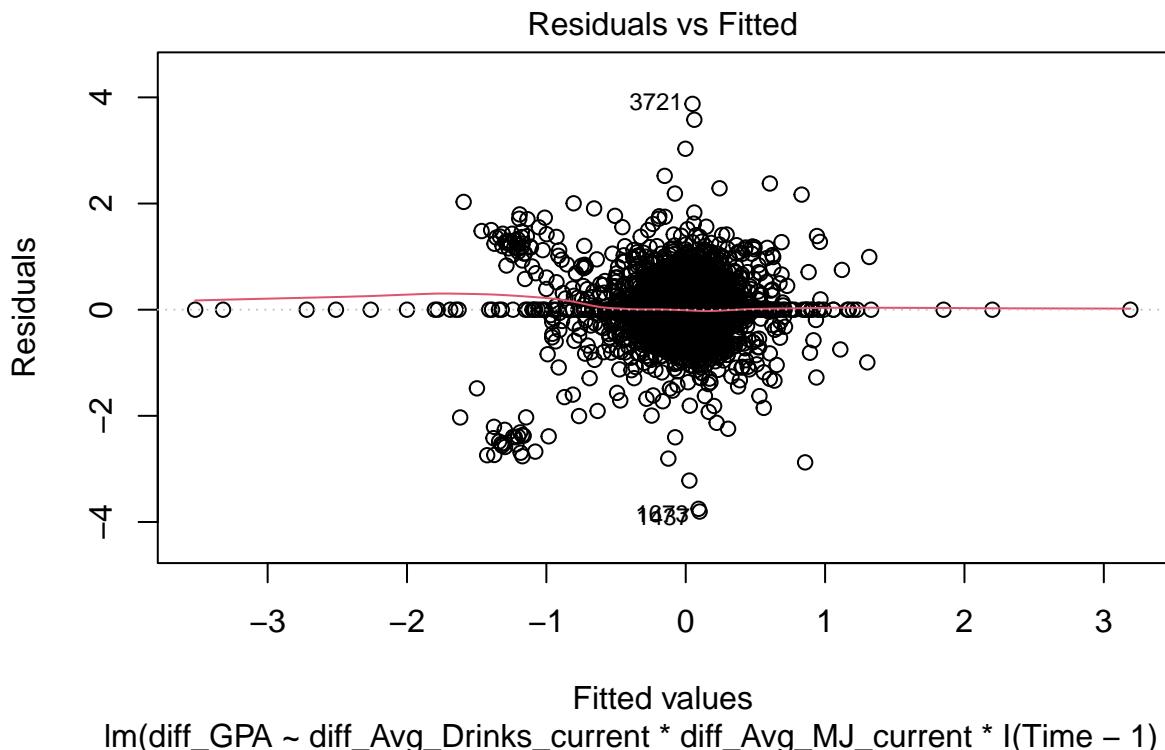
diff.robust.6<- lme(diff_GPA ~ diff_Avg_MJ_current, random = ~ 1 + diff_Avg_MJ_current | BARCS_ID, data = diff)

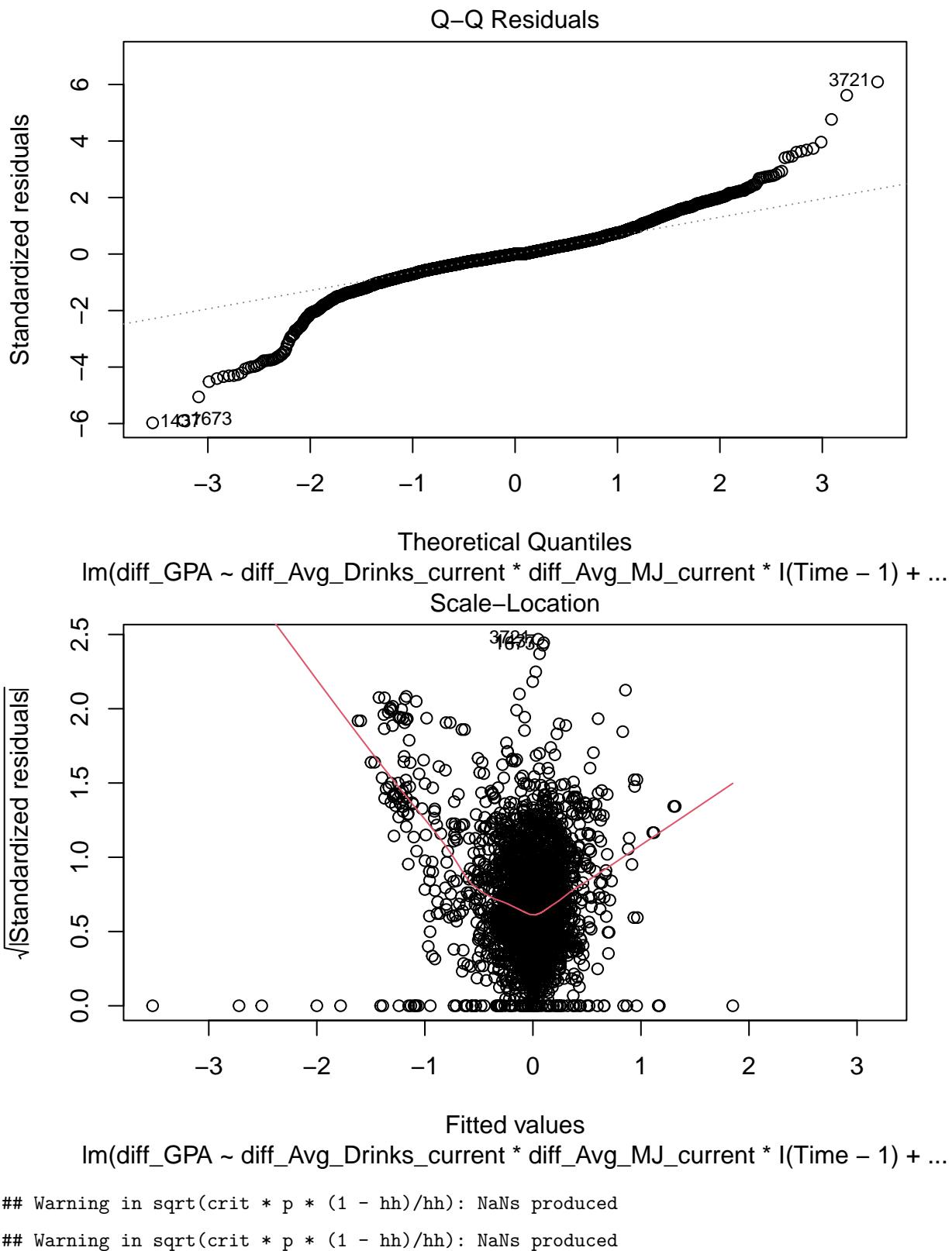
diff.robust.7 <- lme(diff_GPA ~ diff_Avg_MJ_current*diff_Avg_Drinks_current - diff_Avg_Drinks_current, random = ~ 0 + diff_Avg_MJ_current | BARCS_ID)

cowplot::plot_grid(plot(diff.robust.2), plot(diff.robust.3), plot(diff.robust.4), plot(diff.robust.5), plot(diff.robust.6), plot(diff.robust.7))

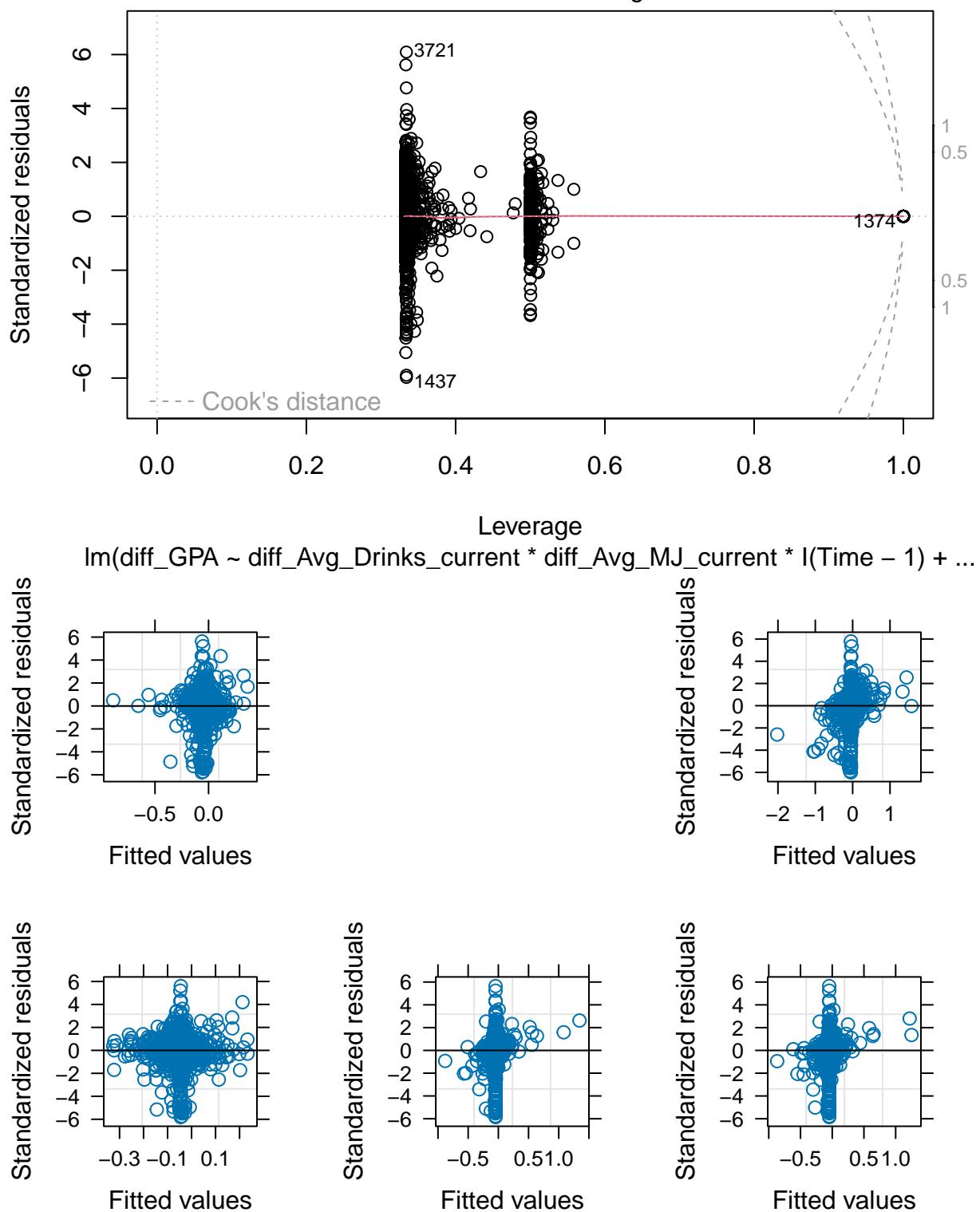
```

Warning: not plotting observations with leverage one:
1155, 1205, 1207, 1208, 1212, 1213, 1214, 1272, 1275, 1276, 1277, 1279, 1280, 1281, 1282, 1288, 1300





Residuals vs Leverage



```
AIC(diff.robust.2, diff.robust.3, diff.robust.4, diff.robust.5, diff.robust.6, diff.robust.7)
```

```
##          df      AIC
## diff.robust.2 10 5761.954
## diff.robust.3 1125 7046.543
```

```

## diff.robust.4 11 5742.030
## diff.robust.5 10 5767.947
## diff.robust.6 6 5754.585
## diff.robust.7 7 5754.721
c(range(ranef(diff.robust.2)), range(ranef(diff.robust.4)), range(ranef(diff.robust.5)), range(ranef(diff.robust.6)), range(ranef(diff.robust.7)))
## [1] -6.924241e-09 5.263689e-09 -4.287642e-02 1.964699e-02 -2.318021e-40
## [6] 3.992313e-40 -5.649102e-02 2.656564e-02 -5.779194e-02 2.547510e-02
c(r.squaredGLMM(diff.robust.2), r.squaredGLMM(diff.robust.4), r.squaredGLMM(diff.robust.5), r.squaredGLMM(diff.robust.6), r.squaredGLMM(diff.robust.7))
## [1] 0.005218532 0.005218534 0.003111525 0.077428616 0.002983748 0.002983748
## [7] 0.002171006 0.018764465 0.003064903 0.018968644

this is a dead end.

Perhaps the students vary too much among themselves too much. Taking a look when the student data is averaged over all 4 Semester

sm.avgd <- lme(GPA ~ Avg_Drinks_current*Avg_MJ_current, data = average.data, random = ~1 | BARCS_ID)
r.squaredGLMM(sm.avgd)

##           R2m      R2c
## [1,] 0.06604796 0.8848429

sm.avgd.2 <- lme(GPA ~ Avg_Drinks_current*Avg_MJ_current, data = (average.data %>% filter(Avg_Drinks_current == 2)), random = ~1 | BARCS_ID)
r.squaredGLMM(sm.avgd.2)

##           R2m      R2c
## [1,] 0.0606483 0.8841957

sm.avgd.3 <- lme(GPA ~ 0 + Avg_Drinks_current*Avg_MJ_current, data = average.data, random = ~1 | BARCS_ID)
r.squaredGLMM(sm.avgd.3)

##           R2m      R2c
## [1,] 0.2444616 0.9068415

sm.avgd.4 <- lme(GPA ~ 0 + Avg_Drinks_current*Avg_MJ_current, data = (average.data %>% filter(Avg_Drinks_current == 4)), random = ~1 | BARCS_ID)
r.squaredGLMM(sm.avgd.4)

##           R2m      R2c
## [1,] 0.2541252 0.9080428

sm.avgd.5 <- lme(GPA ~ I(Avg_Drinks_current^2) + I(Avg_MJ_current^2), data = average.data, random = ~1 | BARCS_ID)
r.squaredGLMM(sm.avgd.5)

##           R2m      R2c
## [1,] 0.059579 0.8840577

fm.avgd <- lme(GPA ~ Avg_Drinks_current*Avg_MJ_current + Sex + Age1stround + SATMath + SATVerbal + SATWriting, data = average.data, random = ~1 | BARCS_ID)
r.squaredGLMM(fm.avgd)

##           R2m      R2c
## [1,] 0.303826 0.9141703

fm.avgd2 <- lme(GPA ~ Avg_Drinks_current*Avg_MJ_current + Sex + Age1stround + SATMath + SATVerbal + SATWriting, data = (average.data %>% filter(Avg_Drinks_current == 2)), random = ~1 | BARCS_ID)
r.squaredGLMM(fm.avgd2)

##           R2m      R2c
## [1,] 0.3003502 0.9137418

```

```

fm.avgd3 <- lme(GPA ~ O + Avg_Drinks_current*Avg_MJ_current + Sex + Age1stround + SATMath + SATVerbal +
r.squaredGLMM(fm.avgd3)

##          R2m      R2c
## [1,] 0.303826 0.9141703

fm.avgd4 <- lme(GPA ~ O + Avg_Drinks_current*Avg_MJ_current + Sex + Age1stround + SATMath + SATVerbal +
r.squaredGLMM(fm.avgd4)

##          R2m      R2c
## [1,] 0.303826 0.9141703

fm.avgd.Cluster <- lme(GPA ~ Avg_Drinks_current*Avg_MJ_current + Sex + Age1stround + SATMath + SATVerbal +
r.squaredGLMM(fm.avgd.Cluster)

##          R2m      R2c
## [1,] 0.3043111 0.9142393

Cluster.avgd <- lme(GPA ~ Sex + Age1stround + SATMath + SATVerbal + SATWriting + Fager4_binary + FH_binary +
r.squaredGLMM(Cluster.avgd)

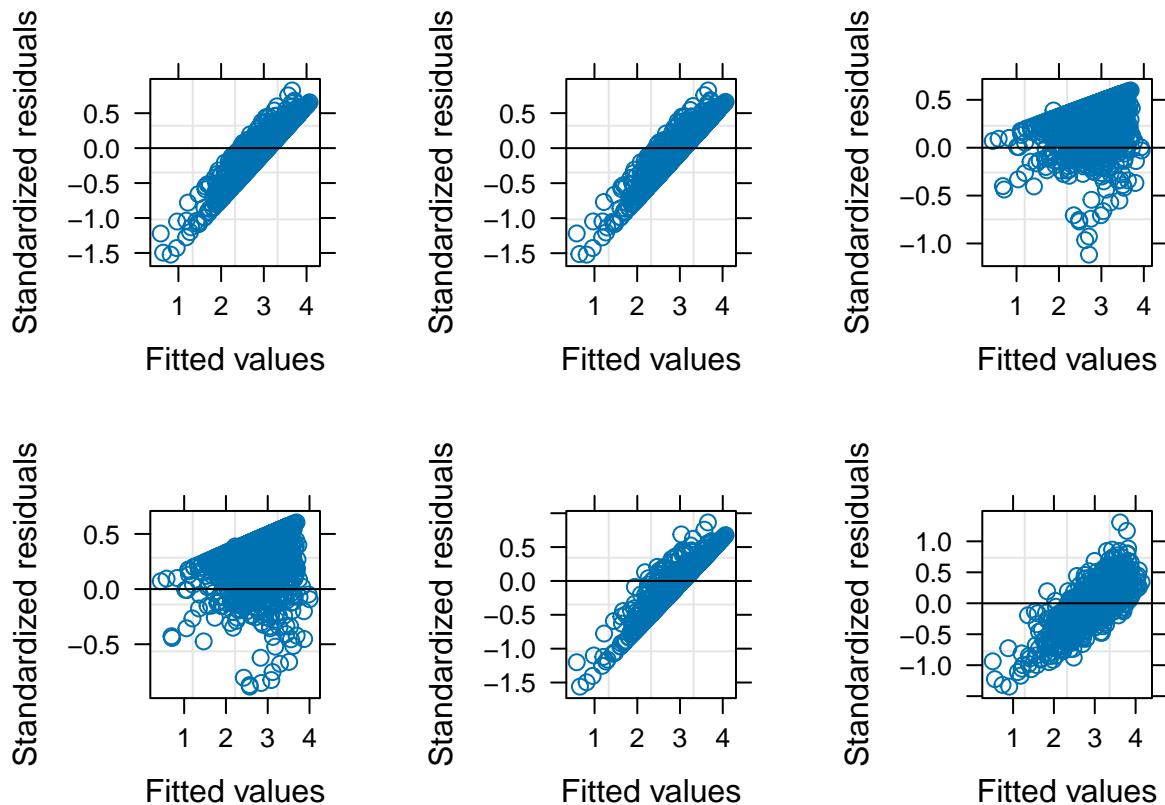
##          R2m      R2c
## [1,] 0.2842217 0.9117486

noSub.avgd <- lme(GPA ~ Sex + Age1stround + SATMath + SATVerbal + SATWriting + Fager4_binary + FH_binary +
r.squaredGLMM(noSub.avgd)

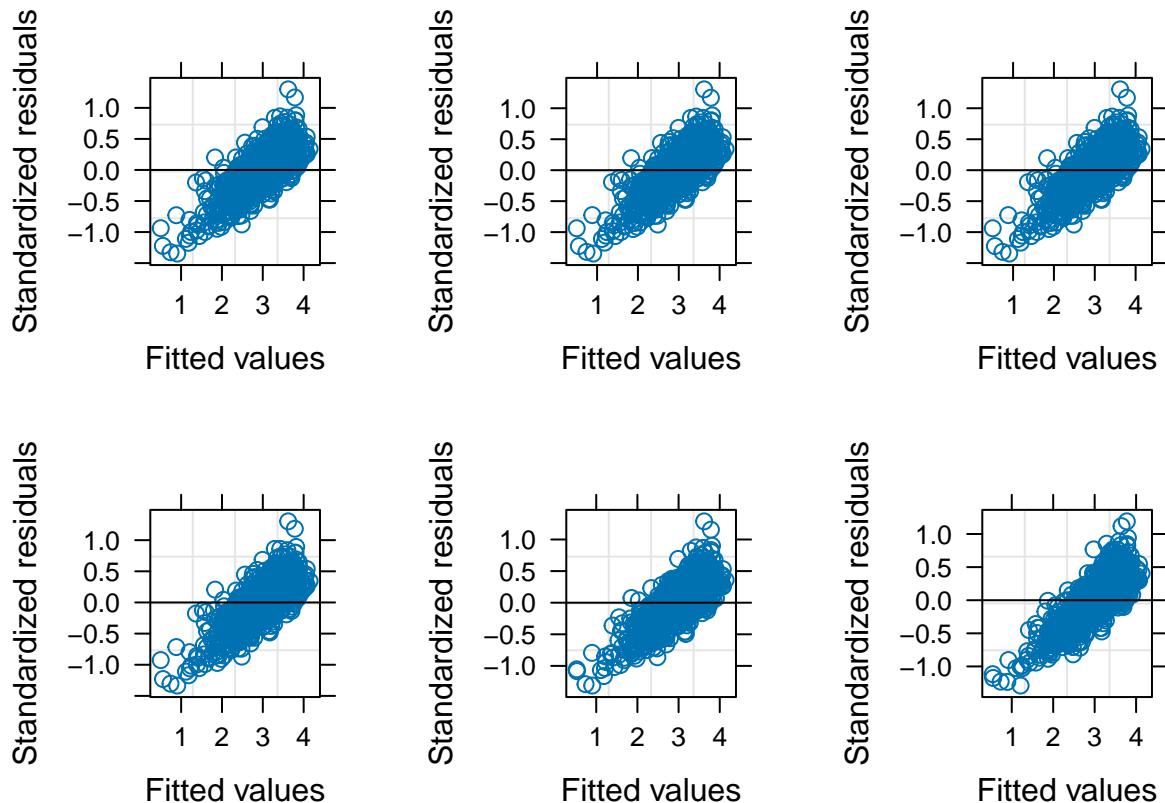
##          R2m      R2c
## [1,] 0.2575405 0.9084639

cowplot:::plot_grid(plot(sm.avgd), plot(sm.avgd.2), plot(sm.avgd.3), plot(sm.avgd.4), plot(sm.avgd.5), p

```



```
cowplot::plot_grid(plot(fm.avgd2), plot(fm.avgd3), plot(fm.avgd4), plot(fm.avgd.Cluster), plot(Cluster.avgd))
```



```
anova(noSub.avgd, fm.avgd)
```

```
## Warning in anova.lme(noSub.avgd, fm.avgd): fitted objects with different fixed
## effects. REML comparisons are not meaningful.
```

```
##          Model df      AIC      BIC logLik  Test L.Ratio p-value
## noSub.avgd     1 14 1723.074 1791.670 -847.5368
## fm.avgd       2 17 1703.489 1786.732 -834.7443 1 vs 2  25.585 <.0001
```

```
noSub.avgd <- lme(GPA ~ Sex + Age1strand + SATMath + SATVerbal + SATWriting + Fager4_binary + FH_binary)
```

```
anova(noSub.avgd, Cluster.avgd)
```

```
## Warning in anova.lme(noSub.avgd, Cluster.avgd): fitted objects with different
## fixed effects. REML comparisons are not meaningful.
```

```
##          Model df      AIC      BIC logLik  Test L.Ratio p-value
## noSub.avgd     1 14 1717.533 1786.072 -844.7665
## Cluster.avgd   2 16 1695.745 1774.043 -831.8723 1 vs 2  25.78826 <.0001
```

There is clearly a confounding effect at play here from variable(s) that were not observed in the study. The inference from the models therefor has to be taken with scepticism as the parameters may be biased and it is unclear to what extend and in which direction they were biased.

The LRT at least leads to the inclusion of Substance usage data, but again, the model assumptions are violated and confounded.