Project Code and Result

Yijie Zhao

2018/12/14

R Markdown

```
library(foreign)
library(readxl)
```

```
## Warning: package 'readxl' was built under R version 3.5.2
```

```
library(dummies)
```

```
## dummies-1.5.6 provided by Decision Patterns
```

```
library(devtools)
```

```
## Warning: package 'devtools' was built under R version 3.5.2
```

```
## Warning: package 'usethis' was built under R version 3.5.2
```

```
library(broom)
```

```
## Warning: package 'broom' was built under R version 3.5.2
```

```
library(readxl)
library(car)
```

```
## Loading required package: carData
```

```
library(het.test)
```

```
## Loading required package: vars
```

```
## Loading required package: MASS
```

```
## Warning: package 'MASS' was built under R version 3.5.2
```

```
## Loading required package: strucchange
```

```
## Loading required package: zoo
```

```
## Warning: package 'zoo' was built under R version 3.5.2
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
## Loading required package: sandwich
```

```
## Warning: package 'sandwich' was built under R version 3.5.2
```

```
## Loading required package: urca
```

```
## Loading required package: lmtest
```

```r
data <- read_excel("data.xlsx")
summary (data$edu)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.000   2.000   1.768   3.000   5.000
```

```r
data$edummy <- as.factor(data$edu)
data$edummy <- relevel(data$edummy, ref = '0')

lm1 <- lm(lgexp~ edummy, data = data)
summary (lm1)
```

```
##
## Call:
## lm(formula = lgexp ~ edummy, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.2990 -0.4586  0.0039  0.4504  3.8641
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  8.64520    0.01478 585.114  < 2e-16 ***
## edummy1      0.12187    0.02324   5.245  1.6e-07 ***
## edummy2      0.17980    0.02175   8.266  < 2e-16 ***
## edummy3      0.26465    0.02215  11.950  < 2e-16 ***
## edummy4      0.44331    0.02808  15.788  < 2e-16 ***
## edummy5      0.93748    0.05540  16.923  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7436 on 9612 degrees of freedom
## Multiple R-squared:  0.05021,    Adjusted R-squared:  0.04971
## F-statistic: 101.6 on 5 and 9612 DF,  p-value: < 2.2e-16
```

```r
lm2 <- lm(lgexp~ edummy + lgincome, data = data)
summary (lm2)
```

```
##
## Call:
## lm(formula = lgexp ~ edummy + lgincome, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.0344 -0.4383 -0.0055  0.4268  3.7207
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 7.318710   0.049781 147.019  < 2e-16 ***
## edummy1     0.094248   0.022379   4.211 2.56e-05 ***
## edummy2     0.114826   0.021057   5.453 5.08e-08 ***
## edummy3     0.150005   0.021702   6.912 5.08e-12 ***
## edummy4     0.253261   0.027865   9.089  < 2e-16 ***
## edummy5     0.616704   0.054531  11.309  < 2e-16 ***
## lgincome    0.159652   0.005742  27.805  < 2e-16 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7154 on 9611 degrees of freedom
## Multiple R-squared:  0.1209, Adjusted R-squared:  0.1204
## F-statistic: 220.3 on 6 and 9611 DF,  p-value: < 2.2e-16
```

```
lm3 <- lm(lgexp~ edummy + lgincome + lgwealth, data = data)
summary (lm3)
```

```
##
## Call:
## lm(formula = lgexp ~ edummy + lgincome + lgwealth, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.1081 -0.4336 -0.0067  0.4228  3.6388
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 6.468764   0.067679  95.580  < 2e-16 ***
## edummy1     0.085183   0.022011   3.870  0.00011 ***
## edummy2     0.093744   0.020738   4.520 6.25e-06 ***
## edummy3     0.110488   0.021449   5.151 2.64e-07 ***
## edummy4     0.194047   0.027592   7.033 2.16e-12 ***
## edummy5     0.481279   0.054134   8.891  < 2e-16 ***
## lgincome    0.127264   0.005920  21.497  < 2e-16 ***
## lgwealth    0.109971   0.006047  18.185  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7035 on 9610 degrees of freedom
## Multiple R-squared:  0.1502, Adjusted R-squared:  0.1495
## F-statistic: 242.6 on 7 and 9610 DF,  p-value: < 2.2e-16
```

```
lm4 <- lm(lgexp~ edummy + lgincome + lgwealth + age, data = data)
summary (lm4)
```

```
##
## Call:
## lm(formula = lgexp ~ edummy + lgincome + lgwealth + age, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.0234 -0.4357 -0.0043  0.4221  3.6367
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  6.707584   0.093219  71.955  < 2e-16 ***
## edummy1      0.077033   0.022105   3.485 0.000494 ***
## edummy2      0.084063   0.020886   4.025 5.75e-05 ***
## edummy3      0.088502   0.022234   3.981 6.93e-05 ***
## edummy4      0.173045   0.028145   6.148 8.14e-10 ***
## edummy5      0.468367   0.054209   8.640  < 2e-16 ***
## lgincome     0.127744   0.005918  21.587  < 2e-16 ***
```

```
## lgwealth     0.107359   0.006084  17.646  < 2e-16 ***
## age         -0.003131   0.000841  -3.723 0.000198 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.703 on 9609 degrees of freedom
## Multiple R-squared:  0.1514, Adjusted R-squared:  0.1507
## F-statistic: 214.3 on 8 and 9609 DF,  p-value: < 2.2e-16
```

```
lm5 <- lm(lgexp~ edummy + lgincome + lgwealth + age + hukou, data = data)
summary (lm5)
```

```
##
## Call:
## lm(formula = lgexp ~ edummy + lgincome + lgwealth + age + hukou,
##     data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.9926 -0.4366 -0.0054  0.4231  3.6646
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  6.9113805  0.0969639  71.278  < 2e-16 ***
## edummy1      0.0704754  0.0220610   3.195 0.001405 **
## edummy2      0.0702553  0.0209120   3.360 0.000784 ***
## edummy3      0.0569221  0.0225800   2.521 0.011721 *
## edummy4      0.1098966  0.0293386   3.746 0.000181 ***
## edummy5      0.3591356  0.0560419   6.408 1.54e-10 ***
## lgincome     0.1170421  0.0060762  19.262  < 2e-16 ***
## lgwealth     0.1032467  0.0060925  16.947  < 2e-16 ***
## age         -0.0043336  0.0008543  -5.072 4.00e-07 ***
## hukou        0.1537082  0.0207973   7.391 1.58e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7011 on 9608 degrees of freedom
## Multiple R-squared:  0.1562, Adjusted R-squared:  0.1554
## F-statistic: 197.6 on 9 and 9608 DF,  p-value: < 2.2e-16
```

```
lm6 <- lm(lgexp~ edummy + lgincome + lgwealth + age + hukou + mar, data = data)
summary (lm6)
```

```
##
## Call:
## lm(formula = lgexp ~ edummy + lgincome + lgwealth + age + hukou +
##     mar, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.0474 -0.4359 -0.0065  0.4209  3.6726
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.0373707  0.1034332  68.038  < 2e-16 ***
```

```
## edummy1      0.0747060  0.0220816   3.383 0.000719 ***
## edummy2      0.0748217  0.0209409   3.573 0.000355 ***
## edummy3      0.0616494  0.0226077   2.727 0.006404 **
## edummy4      0.1151846  0.0293609   3.923 8.80e-05 ***
## edummy5      0.3646451  0.0560317   6.508 8.01e-11 ***
## lgincome     0.1178156  0.0060767  19.388  < 2e-16 ***
## lgwealth     0.1021534  0.0060970  16.755  < 2e-16 ***
## age         -0.0050787  0.0008802  -5.770 8.18e-09 ***
## hukou        0.1529113  0.0207865   7.356 2.05e-13 ***
## mar         -0.0846943  0.0243056  -3.485 0.000495 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7006 on 9607 degrees of freedom
## Multiple R-squared:  0.1572, Adjusted R-squared:  0.1564
## F-statistic: 179.3 on 10 and 9607 DF,  p-value: < 2.2e-16
```

```
lm7 <- lm(lgexp~ edummy + lgincome + lgwealth + age + hukou + mar + com, data = data)
summary (lm7)
```

```
##
## Call:
## lm(formula = lgexp ~ edummy + lgincome + lgwealth + age + hukou +
##     mar + com, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.0308 -0.4344 -0.0070  0.4235  3.6801
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.0839511  0.1039873  68.123  < 2e-16 ***
## edummy1      0.0706412  0.0220867   3.198  0.00139 **
## edummy2      0.0689052  0.0209750   3.285  0.00102 **
## edummy3      0.0514113  0.0227303   2.262  0.02373 *
## edummy4      0.0950251  0.0297560   3.193  0.00141 **
## edummy5      0.3193814  0.0570898   5.594 2.28e-08 ***
## lgincome     0.1171832  0.0060739  19.293  < 2e-16 ***
## lgwealth     0.1016803  0.0060933  16.687  < 2e-16 ***
## age         -0.0056143  0.0008894  -6.313 2.87e-10 ***
## hukou        0.1478211  0.0208078   7.104 1.30e-12 ***
## mar         -0.0885701  0.0243049  -3.644  0.00027 ***
## com          0.0984588  0.0242931   4.053 5.10e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7001 on 9606 degrees of freedom
## Multiple R-squared:  0.1587, Adjusted R-squared:  0.1577
## F-statistic: 164.7 on 11 and 9606 DF,  p-value: < 2.2e-16
```

```
lm8 <- lm(lgexp~ edummy + lgincome + lgwealth + age + hukou + mar + com + gender, data = data)
summary (lm8)
```

```
##
## Call:
```

```
## lm(formula = lgexp ~ edummy + lgincome + lgwealth + age + hukou +
##     mar + com + gender, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.0304 -0.4345 -0.0072  0.4234  3.6797
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.0841119  0.1040715  68.070  < 2e-16 ***
## edummy1      0.0704623  0.0225426   3.126 0.001779 **
## edummy2      0.0686534  0.0219129   3.133 0.001735 **
## edummy3      0.0511168  0.0239103   2.138 0.032554 *
## edummy4      0.0947036  0.0308390   3.071 0.002140 **
## edummy5      0.3190358  0.0577522   5.524 3.40e-08 ***
## lgincome     0.1171975  0.0060848  19.261  < 2e-16 ***
## lgwealth     0.1016872  0.0060961  16.681  < 2e-16 ***
## age         -0.0056205  0.0009029  -6.225 5.02e-10 ***
## hukou        0.1478873  0.0208756   7.084 1.50e-12 ***
## mar         -0.0886467  0.0243824  -3.636 0.000279 ***
## com          0.0982951  0.0246418   3.989 6.69e-05 ***
## gender       0.0006262  0.0157674   0.040 0.968323
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7001 on 9605 degrees of freedom
## Multiple R-squared:  0.1587, Adjusted R-squared:  0.1576
## F-statistic:   151 on 12 and 9605 DF,  p-value: < 2.2e-16
```

```
lm01 <- lm(engel~ edummy, data = data)
summary (lm01)
```

```
##
## Call:
## lm(formula = engel ~ edummy, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -59.508 -16.104   2.546  17.959  45.095
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  60.2508     0.4459 135.128  < 2e-16 ***
## edummy1      -0.4562     0.7013  -0.651  0.51534
## edummy2      -1.8570     0.6564  -2.829  0.00468 **
## edummy3      -4.1309     0.6683  -6.181 6.62e-10 ***
## edummy4      -6.7842     0.8473  -8.007 1.32e-15 ***
## edummy5     -11.3273     1.6717  -6.776 1.31e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 22.44 on 9612 degrees of freedom
## Multiple R-squared:  0.01261,    Adjusted R-squared:  0.01209
## F-statistic: 24.54 on 5 and 9612 DF,  p-value: < 2.2e-16
```

```r
lm02 <- lm(engel~ edummy + lgincome, data = data)
summary (lm02)
```

```
##
## Call:
## lm(formula = engel ~ edummy + lgincome, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -61.500 -15.939   2.552  17.797  45.142
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  72.7117     1.5559  46.734  < 2e-16 ***
## edummy1      -0.1967     0.6994  -0.281   0.7785
## edummy2      -1.2466     0.6581  -1.894   0.0582 .
## edummy3      -3.0540     0.6783  -4.503 6.79e-06 ***
## edummy4      -4.9989     0.8709  -5.740 9.76e-09 ***
## edummy5      -8.3140     1.7043  -4.878 1.09e-06 ***
## lgincome     -1.4997     0.1795  -8.357  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 22.36 on 9611 degrees of freedom
## Multiple R-squared:  0.01973,    Adjusted R-squared:  0.01912
## F-statistic: 32.24 on 6 and 9611 DF,  p-value: < 2.2e-16
```

```r
lm03 <- lm(engel~ edummy + lgincome + lgwealth, data = data)
summary (lm03)
```

```
##
## Call:
## lm(formula = engel ~ edummy + lgincome + lgwealth, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -62.167 -15.950   2.503  17.609  46.509
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  82.2662     2.1467  38.322  < 2e-16 ***
## edummy1      -0.0948     0.6982  -0.136 0.891993
## edummy2      -1.0096     0.6578  -1.535 0.124836
## edummy3      -2.6098     0.6804  -3.836 0.000126 ***
## edummy4      -4.3333     0.8752  -4.951 7.50e-07 ***
## edummy5      -6.7917     1.7171  -3.955 7.70e-05 ***
## lgincome     -1.1357     0.1878  -6.048 1.52e-09 ***
## lgwealth     -1.2362     0.1918  -6.445 1.21e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 22.31 on 9610 degrees of freedom
## Multiple R-squared:  0.02395,    Adjusted R-squared:  0.02324
## F-statistic: 33.68 on 7 and 9610 DF,  p-value: < 2.2e-16
```

```r
lm04 <- lm(engel~ edummy + lgincome + lgwealth + age, data = data)
summary (lm04)
```

```
##
## Call:
## lm(formula = engel ~ edummy + lgincome + lgwealth + age, data = data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -60.99 -15.99   2.54  17.62  48.03
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 92.20260    2.95530  31.199  < 2e-16 ***
## edummy1     -0.43392    0.70077  -0.619   0.5358
## edummy2     -1.41243    0.66215  -2.133   0.0329 *
## edummy3     -3.52450    0.70487  -5.000 5.83e-07 ***
## edummy4     -5.20710    0.89227  -5.836 5.53e-09 ***
## edummy5     -7.32889    1.71855  -4.265 2.02e-05 ***
## lgincome    -1.11573    0.18760  -5.947 2.82e-09 ***
## lgwealth    -1.34490    0.19288  -6.973 3.31e-12 ***
## age         -0.13027    0.02666  -4.886 1.05e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 22.29 on 9609 degrees of freedom
## Multiple R-squared:  0.02637,    Adjusted R-squared:  0.02556
## F-statistic: 32.53 on 8 and 9609 DF,  p-value: < 2.2e-16
```

```r
lm05 <- lm(engel~ edummy + lgincome + lgwealth + age + hukou, data = data)
summary (lm05)
```

```
##
## Call:
## lm(formula = engel ~ edummy + lgincome + lgwealth + age + hukou,
##     data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -60.826 -15.990   2.512  17.622  47.704
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 86.67160    3.07634  28.174  < 2e-16 ***
## edummy1     -0.25595    0.69992  -0.366 0.714604
## edummy2     -1.03769    0.66347  -1.564 0.117840
## edummy3     -2.66742    0.71639  -3.723 0.000198 ***
## edummy4     -3.49326    0.93081  -3.753 0.000176 ***
## edummy5     -4.36436    1.77802  -2.455 0.014121 *
## lgincome    -0.82530    0.19278  -4.281 1.88e-05 ***
## lgwealth    -1.23330    0.19329  -6.380 1.85e-10 ***
## age         -0.09763    0.02711  -3.602 0.000318 ***
## hukou       -4.17161    0.65983  -6.322 2.69e-10 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 22.24 on 9608 degrees of freedom
## Multiple R-squared:  0.0304, Adjusted R-squared:  0.02949
## F-statistic: 33.47 on 9 and 9608 DF,  p-value: < 2.2e-16
```

```
lm06 <- lm(engel~ edummy + lgincome + lgwealth + age + hukou + mar, data = data)
summary (lm06)
```

```
##
## Call:
## lm(formula = engel ~ edummy + lgincome + lgwealth + age + hukou +
##     mar, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -60.619 -15.959   2.531  17.599  47.840
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 87.75394    3.28351  26.726  < 2e-16 ***
## edummy1     -0.21961    0.70099  -0.313 0.754069
## edummy2     -0.99846    0.66477  -1.502 0.133140
## edummy3     -2.62681    0.71768  -3.660 0.000253 ***
## edummy4     -3.44783    0.93207  -3.699 0.000218 ***
## edummy5     -4.31703    1.77874  -2.427 0.015242 *
## lgincome    -0.81865    0.19291  -4.244 2.22e-05 ***
## lgwealth    -1.24269    0.19355  -6.420 1.42e-10 ***
## age         -0.10403    0.02794  -3.723 0.000198 ***
## hukou       -4.17845    0.65987  -6.332 2.52e-10 ***
## mar         -0.72758    0.77159  -0.943 0.345721
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 22.24 on 9607 degrees of freedom
## Multiple R-squared:  0.03049,    Adjusted R-squared:  0.02948
## F-statistic: 30.21 on 10 and 9607 DF,  p-value: < 2.2e-16
```

```
lm07 <- lm(engel~ edummy + lgincome + lgwealth + age + hukou + mar + com, data = data)
summary (lm07)
```

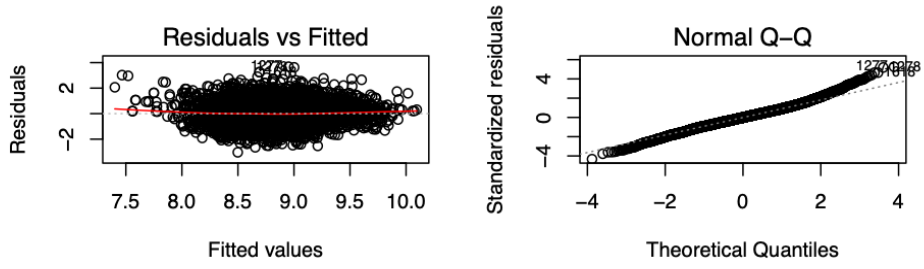
```
##
## Call:
## lm(formula = engel ~ edummy + lgincome + lgwealth + age + hukou +
##     mar + com, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -60.608 -15.968   2.532  17.605  47.902
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 87.80144    3.30392  26.575  < 2e-16 ***
## edummy1     -0.22376    0.70175  -0.319 0.749843
## edummy2     -1.00450    0.66642  -1.507 0.131768
```

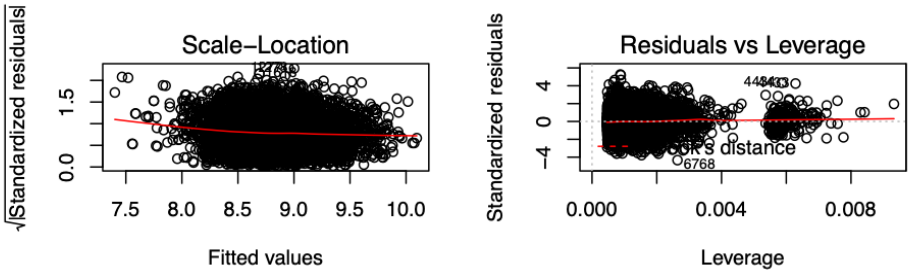
```
## edummy3     -2.63725    0.72219  -3.652 0.000262 ***
## edummy4     -3.46839    0.94542  -3.669 0.000245 ***
## edummy5     -4.36319    1.81388  -2.405 0.016171 *
## lgincome    -0.81930    0.19298  -4.246 2.20e-05 ***
## lgwealth    -1.24317    0.19360  -6.421 1.41e-10 ***
## age         -0.10458    0.02826  -3.701 0.000216 ***
## hukou       -4.18364    0.66111  -6.328 2.59e-10 ***
## mar         -0.73153    0.77222  -0.947 0.343506
## com          0.10040    0.77185   0.130 0.896507
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 22.24 on 9606 degrees of freedom
## Multiple R-squared:  0.03049,    Adjusted R-squared:  0.02938
## F-statistic: 27.46 on 11 and 9606 DF,  p-value: < 2.2e-16
```

```
lm08 <- lm(engel~ edummy + lgincome + lgwealth + age + hukou + mar + com + gender, data = data)
summary (lm08)
```

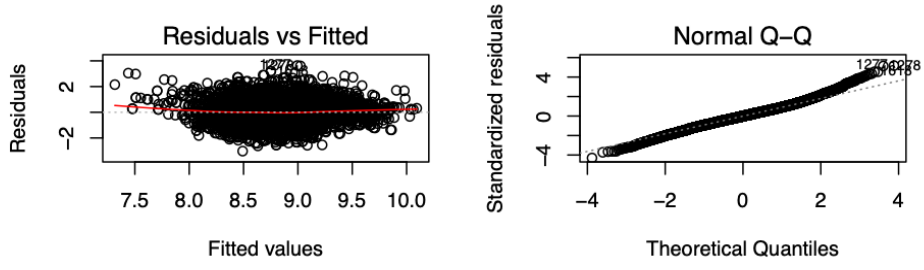
```
##
## Call:
## lm(formula = engel ~ edummy + lgincome + lgwealth + age + hukou +
##     mar + com + gender, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -61.704 -15.940   2.454  17.655  47.221
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 88.29138    3.30409  26.722  < 2e-16 ***
## edummy1     -0.76901    0.71569  -1.075 0.282623
## edummy2     -1.77168    0.69570  -2.547 0.010892 *
## edummy3     -3.53467    0.75911  -4.656 3.26e-06 ***
## edummy4     -4.44815    0.97909  -4.543 5.61e-06 ***
## edummy5     -5.41643    1.83354  -2.954 0.003144 **
## lgincome    -0.77588    0.19318  -4.016 5.96e-05 ***
## lgwealth    -1.22189    0.19354  -6.313 2.85e-10 ***
## age         -0.12338    0.02867  -4.304 1.69e-05 ***
## hukou       -3.98175    0.66277  -6.008 1.95e-09 ***
## mar         -0.96472    0.77410  -1.246 0.212702
## com         -0.39867    0.78234  -0.510 0.610350
## gender       1.90829    0.50059   3.812 0.000139 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 22.23 on 9605 degrees of freedom
## Multiple R-squared:  0.03196,    Adjusted R-squared:  0.03075
## F-statistic: 26.42 on 12 and 9605 DF,  p-value: < 2.2e-16
```

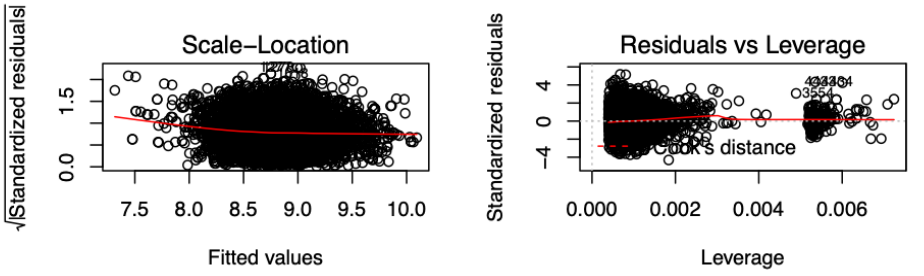
```
par(mfrow=c(2,2))
plot(lm8)
```






```
par(mfrow=c(2,2))
plot(lm4)
```





```r
vif(lm8)
```

```
##              GVIF Df GVIF^(1/(2*Df))
## edummy   1.764293  5        1.058418
## lgincome 1.308325  1        1.143820
## lgwealth 1.218664  1        1.103931
## age      1.306428  1        1.142991
## hukou    1.384815  1        1.176782
## mar      1.100167  1        1.048889
## com      1.149032  1        1.071929
## gender   1.219239  1        1.104191
```

```r
bptest(lm8)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  lm8
## BP = 80.281, df = 12, p-value = 3.648e-12
```

```r
covmtx <- hccm(lm8)
coeftest(lm8, vcov = covmtx)
```

```
##
## t test of coefficients:
##
##                Estimate  Std. Error t value  Pr(>|t|)
## (Intercept)  7.08411189  0.10890236 65.0501 < 2.2e-16 ***
## edummy1      0.07046232  0.02305434  3.0564 0.0022466 **
## edummy2      0.06865343  0.02216397  3.0975 0.0019571 **
## edummy3      0.05111683  0.02379249  2.1484 0.0317034 *
## edummy4      0.09470356  0.02998201  3.1587 0.0015898 **
## edummy5      0.31903575  0.05057384  6.3083 2.945e-10 ***
## lgincome     0.11719748  0.00750862 15.6084 < 2.2e-16 ***
## lgwealth     0.10168724  0.00646965 15.7176 < 2.2e-16 ***
## age         -0.00562051  0.00091283 -6.1572 7.699e-10 ***
## hukou        0.14788731  0.02016436  7.3341 2.413e-13 ***
## mar         -0.08864667  0.02666407 -3.3246 0.0008889 ***
## com          0.09829507  0.02495465  3.9389 8.242e-05 ***
## gender       0.00062617  0.01579529  0.0396 0.9683785
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
vif(lm4)
```

```
##              GVIF Df GVIF^(1/(2*Df))
## edummy   1.255089  5        1.022981
## lgincome 1.227291  1        1.107832
## lgwealth 1.203856  1        1.097204
## age      1.124246  1        1.060305
```

```r
bptest(lm4)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  lm4
```

```
## BP = 60.521, df = 8, p-value = 3.684e-10
```

```
covmtx <- hccm(lm4)
coeftest(lm4, vcov = covmtx)
```

```
##
## t test of coefficients:
##
##                Estimate  Std. Error t value  Pr(>|t|)
## (Intercept)  6.70758370  0.09744425 68.8351 < 2.2e-16 ***
## edummy1      0.07703280  0.02265726  3.3999 0.0006768 ***
## edummy2      0.08406278  0.02123177  3.9593 7.571e-05 ***
## edummy3      0.08850234  0.02205930  4.0120 6.066e-05 ***
## edummy4      0.17304535  0.02748446  6.2961 3.185e-10 ***
## edummy5      0.46836752  0.04646722 10.0795 < 2.2e-16 ***
## lgincome     0.12774345  0.00735989 17.3567 < 2.2e-16 ***
## lgwealth     0.10735864  0.00649069 16.5404 < 2.2e-16 ***
## age         -0.00313094  0.00084902 -3.6877 0.0002275 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```