Modelling company's performance based on financial ratios

Slav Emilov Angelov
PhD student
New Bulgarian University

19th European Young Statisticians Meeting
Prague, The Czech Republic

The Goal

1) Predicting wanted firm information based on past observations - financial ratios and macroeconomic variables.
2) A model who can predict current parameter not only for an individual firm but for a group of similar ones.

Why this goal is valuable?

"Wanted firm information" ⇔ :

- ✓ ROA – return on assets
- ✓ ROC – return on capital
- ✓ Leverage – the ratio between the depth and equity
- ✓ Others

Usable by economists, shareholders, managers, employees, banks and even the government.

The Current Project

Gas suppling sector in Bulgaria:

1) 28 licensed private firms;
2) 2 firms are inactive;
3) 13 have poor financial statements(small firms)
4) Others have just started their activity

Result: Only 5 firms are observed for a period of 4 years. These firms are with similar accounting politics and are licensed to supply with gas around 45% of the population in Bulgaria.

What to predict?

- **ROA:**

$$ROA = \frac{net\ income + 90\% * interests}{All\ assets}$$

- **Financial leverage(FL):**

$$FL = \frac{liabilities - cash}{liabilities - cash + equity}$$

Predictors

Sources:

- Yearly financial statements
- Bulgarian national bank
- National statistical institute

Types:

- Financial ratios
- Macroeconomic variables

Primal Difficulties

- Examining the differences between the accounting politics – 2 weeks
- Downloading, reading, converting and extracting information from the annual reports divided into 243 files – 3 weeks
- Based on the extracted information over 90 financial ratios were calculated – over 1 week

Methodology

Multiple linear regression analysis are used for predicting ROA and LF. Each of the models is derived from the current algorithm:

1) Choosing appropriate predictors
2) Checking the Gauss-Markov conditions
3) Fixing multicollinearity problems
4) Examining for normality of the errors

Difficulties

- There are only 21 observations and over 90 variables. Which one to choose?
- The income of one firm depends from many factors so we can not expect to predict ROA with few predictors. In the same time there should be enough degrees of freedom for the regression.
- How to make the regression fit so every condition to be satisfied?

Choosing predictors

Two methodologies were used:

- Stepwise regression (SPSS) – a combination between forward and backward regression based on F-values. This algorithm was modified by manually deleting the first chosen variable and starting it again. Than comparing the results with the older ones.
- All subsets regression (r language) – a search between all possible regression equations out of a given list of variables.

Predicting ROA

Using all subsets regression:

```
Residuals:
       Min         1Q     Median         3Q        Max
-0.0037533 -0.0017833 -0.0006056  0.0011295  0.0061512

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.030226   0.003613   8.365 1.37e-06 ***
RnP         -0.609910   0.041455 -14.713 1.75e-09 ***
IADMA2      -0.437542   0.051759  -8.453 1.22e-06 ***
KSOLP        0.011415   0.001256   9.089 5.39e-07 ***
AL2          0.226243   0.030116   7.512 4.42e-06 ***
BVnKLY       1.226598   0.072713  16.869 3.22e-10 ***
KAA2         0.394957   0.050045   7.892 2.59e-06 ***
InPoP       -0.007515   0.001086  -6.921 1.05e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.00303 on 13 degrees of freedom
Multiple R-squared:  0.9711, Adjusted R-squared:  0.9555
F-statistic: 62.34 on 7 and 13 DF,  p-value: 5.457e-09
```

Predictor variables

Macroeconomic variables:

KSOLP – end value of the basic interest rate given by the Bulgarian national bank

Financial ratios:

$$\boldsymbol{RnP} = \frac{net\ income}{liabilities},$$

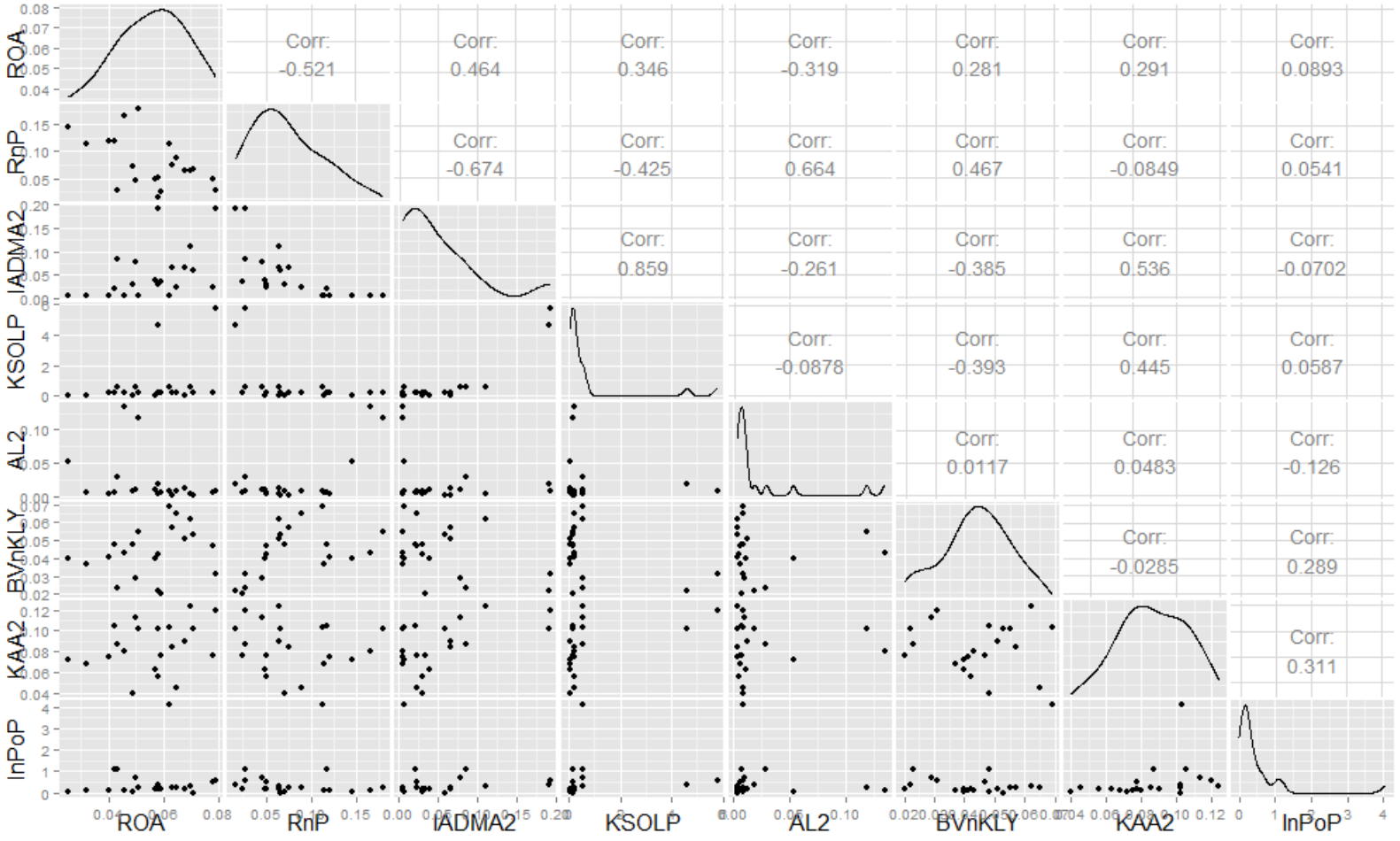
$$\boldsymbol{IADMA2} = \frac{Average\ plants\ and\ equipment\ in\ construction}{Average\ fixed\ assets},$$

$$\boldsymbol{Al2} = \frac{Average\ cash\ and\ marketable\ securities}{Average\ current\ liabilities},$$

$$\textbf{BVnKLY} = \frac{net\ income}{average\ net\ working\ capital + average\ fixed\ assets},$$

$$\boldsymbol{KAA2} = \frac{average\ current\ assets}{average\ assets}, \qquad \boldsymbol{InPoP} = \frac{sales_i - sales_{i-1}}{sales_{i-1}}$$

Multicollinearity Problem



Detecting Multicollinearity

$\boldsymbol{R_j^2}$ ⇔ the value of R^2 between x_j and all other predictors from the model.

The tolerance (TOL): $TOL_j = 1 - R_j^2$

Variance inflation factor (VIF): $VIF_j = TOL_j^{-1}$

In the ROA model VIFS are:

RnP	IADMA2	KSOLP	AL2	BVnKLY	KAA2	InPoP
8.309777	17.833190	7.835377	2.615069	2.212346	2.949692	1.981728

Multicollinearity correction algorithm

1) The first predictor(RnP) is taken as a dependent variable and a linear regression is run with the all others 6 predictors.
2) RnP is replaced by the residuals from that regression
3) The second predictor(IADMA2) is taken and again a regression is run with the 5 left predictors
4) ….

Note: The only changeable variable during the regressions is the intercept.

Multicollinearity correction

Replacements:

- $\boldsymbol{RnP} \Leftrightarrow \boldsymbol{RnP} - (0.975086 * \boldsymbol{IADMA2} + 0.018285 * \boldsymbol{KSOLP} + 0.481844 * \boldsymbol{AL2} + 1.181413 * \boldsymbol{BVnKLY} + 0.714054 * \boldsymbol{KAA2} - 0.011794 * \boldsymbol{InPoP})$

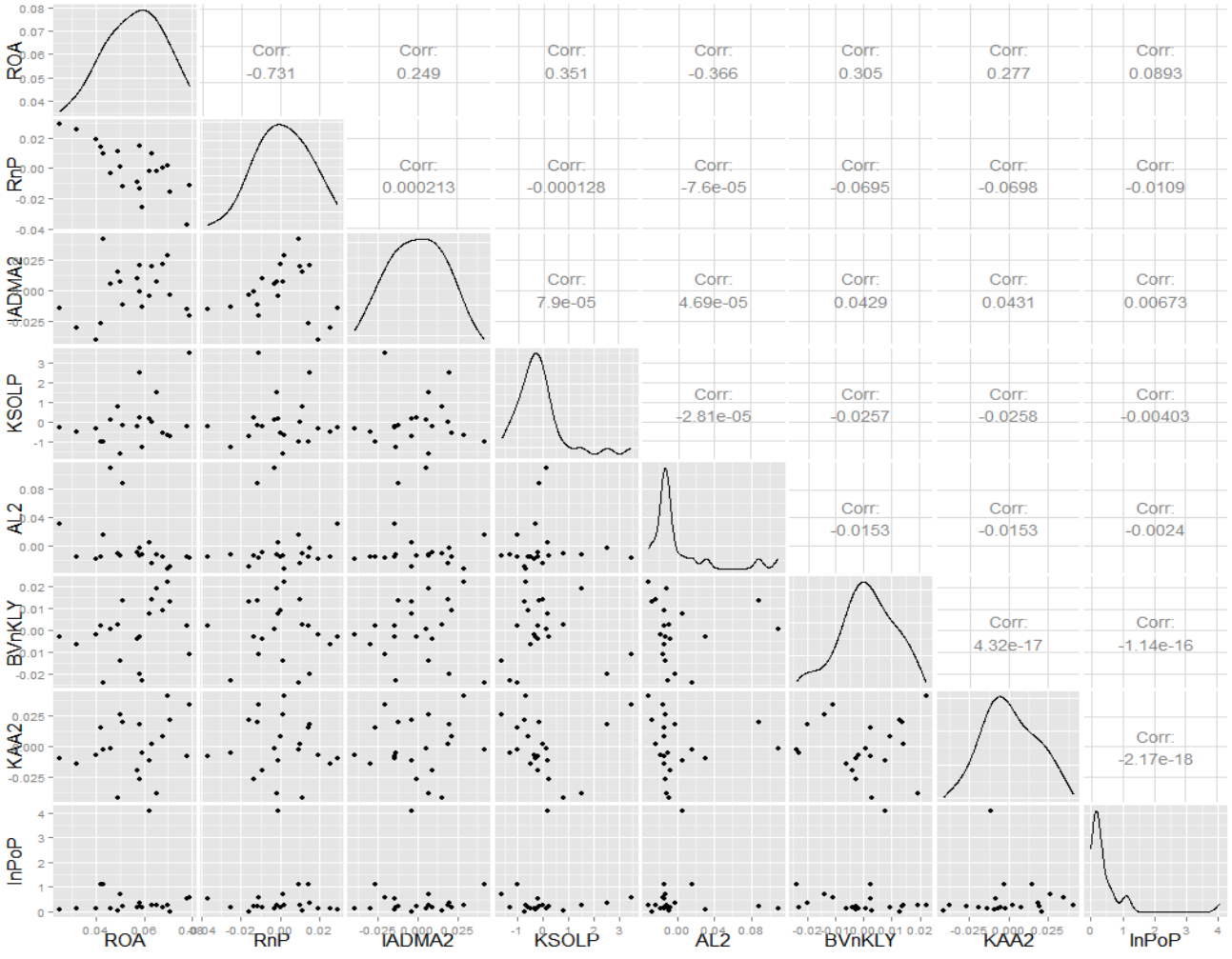
- $\boldsymbol{IADMA2} \Leftrightarrow \boldsymbol{IADMA2} - (0.025675 * \boldsymbol{KSOLP} - 0.369912 * \boldsymbol{AL2} - 0.194863 * \boldsymbol{BVnKLY} + 0.681236 * \boldsymbol{KAA2} - 0.013827 * \boldsymbol{InPoP})$

- $\boldsymbol{KSOLP} \Leftrightarrow \boldsymbol{KSOLP} - (-4.1725 * \boldsymbol{AL2} - 39.7558 * \boldsymbol{BVnKLY} + 29.6410 * \boldsymbol{KAA2} + 0.0217 * \boldsymbol{InPoP})$

- $\boldsymbol{Al2} \Leftrightarrow \boldsymbol{Al2} - (0.215438 * \boldsymbol{BVnKLY} + 0.193447 * \boldsymbol{KAA2} - 0.007706 * \boldsymbol{InPoP})$ □

- $\boldsymbol{BVnKLY}$ $\boldsymbol{BVnKLY} - (-0.078142 * \boldsymbol{KAA2} + 0.005200 * \boldsymbol{InPoP} + 0.047984)$
- $\boldsymbol{KAA2}$ $\boldsymbol{KAA2} - (0.008238 * \boldsymbol{InPoP} + 0.080906)$

Results from the correction algorithm



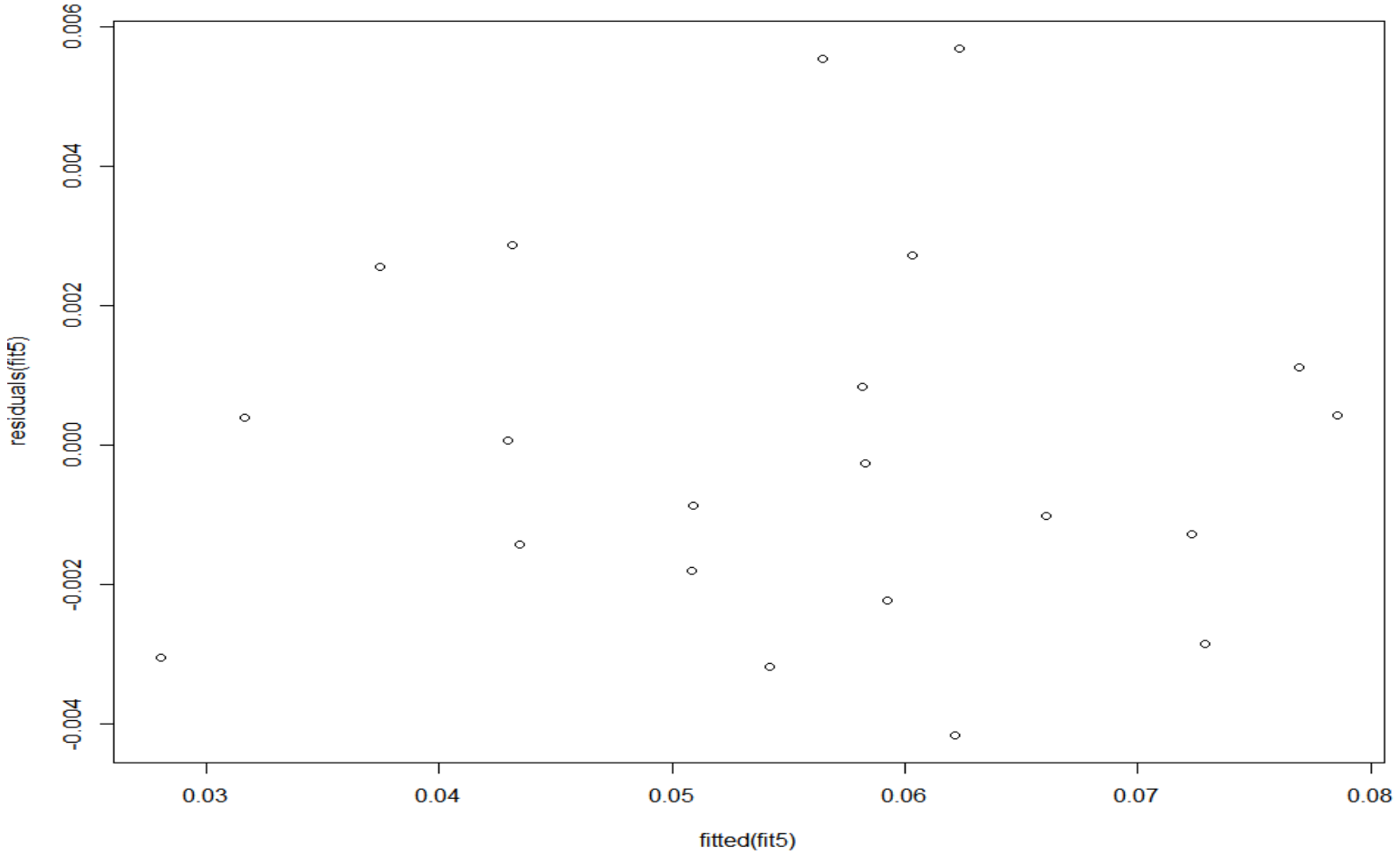
The model without multicollinearity

```
Residuals:
       Min         1Q     Median         3Q        Max
-0.0041629 -0.0018074 -0.0002712  0.0011129  0.0056858

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.0557282  0.0007055  78.996  < 2e-16 ***
RnP         -0.6106919  0.0442184 -13.811 1.51e-09 ***
IADMA2       0.1575532  0.0347585   4.533 0.000469 ***
KSOLP        0.0042944  0.0005958   7.208 4.51e-06 ***
AL2         -0.1437924  0.0201539  -7.135 5.05e-06 ***
BVnKLY       0.2734652  0.0551463   4.959 0.000210 ***
KAA2         0.1446839  0.0328339   4.407 0.000597 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.003232 on 14 degrees of freedom
Multiple R-squared:  0.9646, Adjusted R-squared:  0.9494
F-statistic: 63.49 on 6 and 14 DF,  p-value: 2.378e-09
```

VIFs:

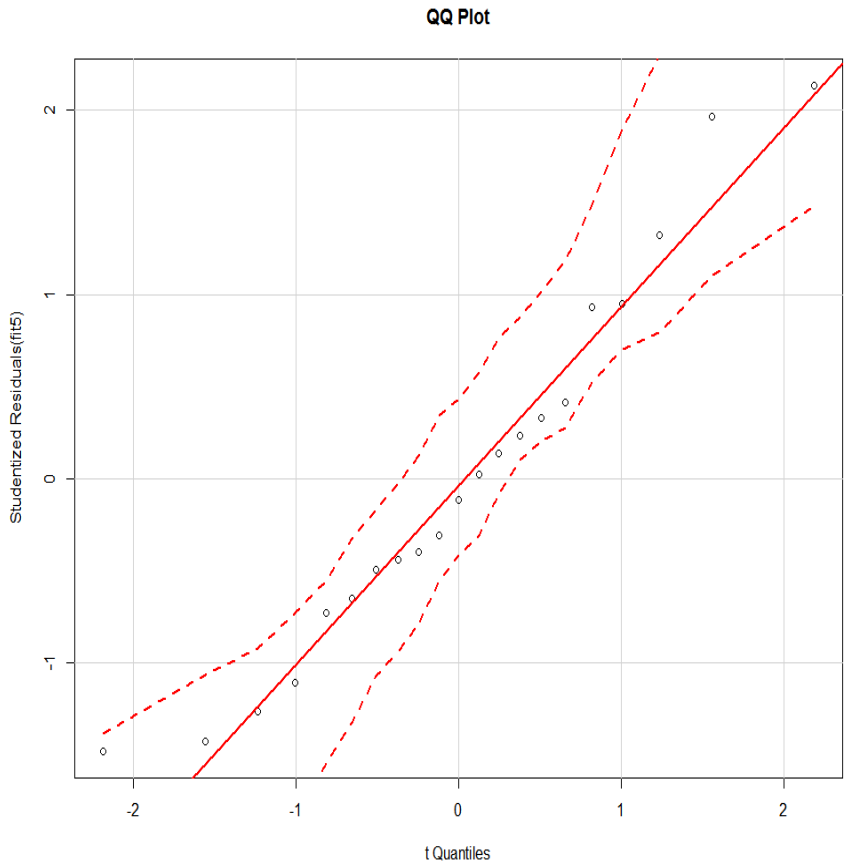
```
     RnP    IADMA2     KSOLP       AL2    BVnKLY      KAA2
1.009849 1.003758 1.001348 1.000476 1.007684 1.007744
```

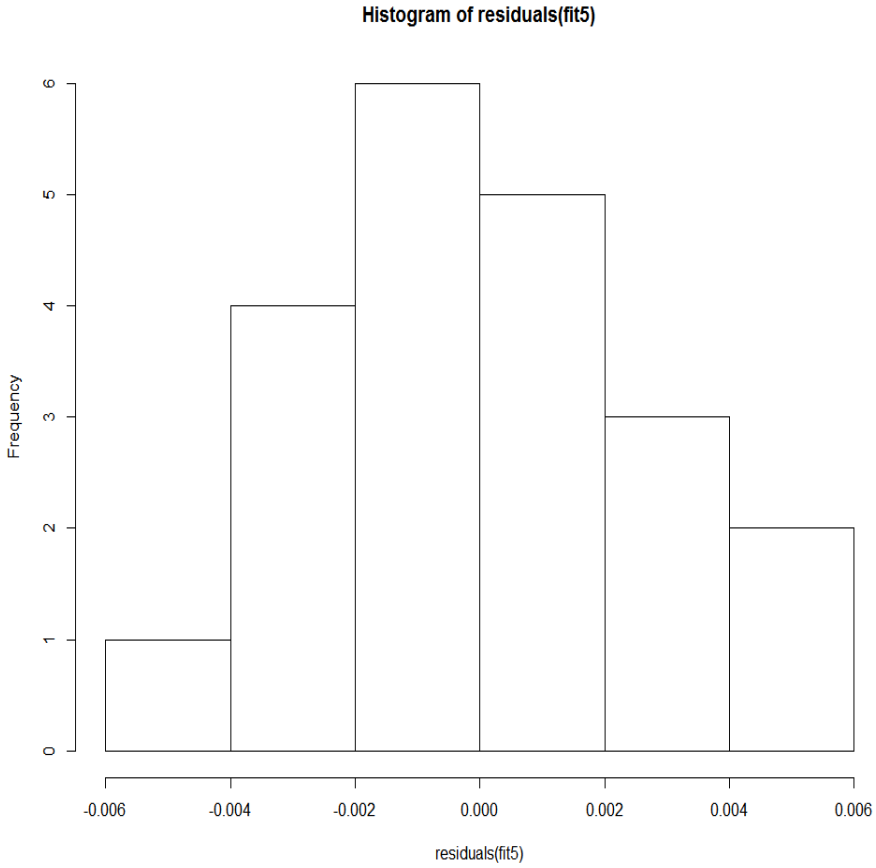
Heteroscedasticity



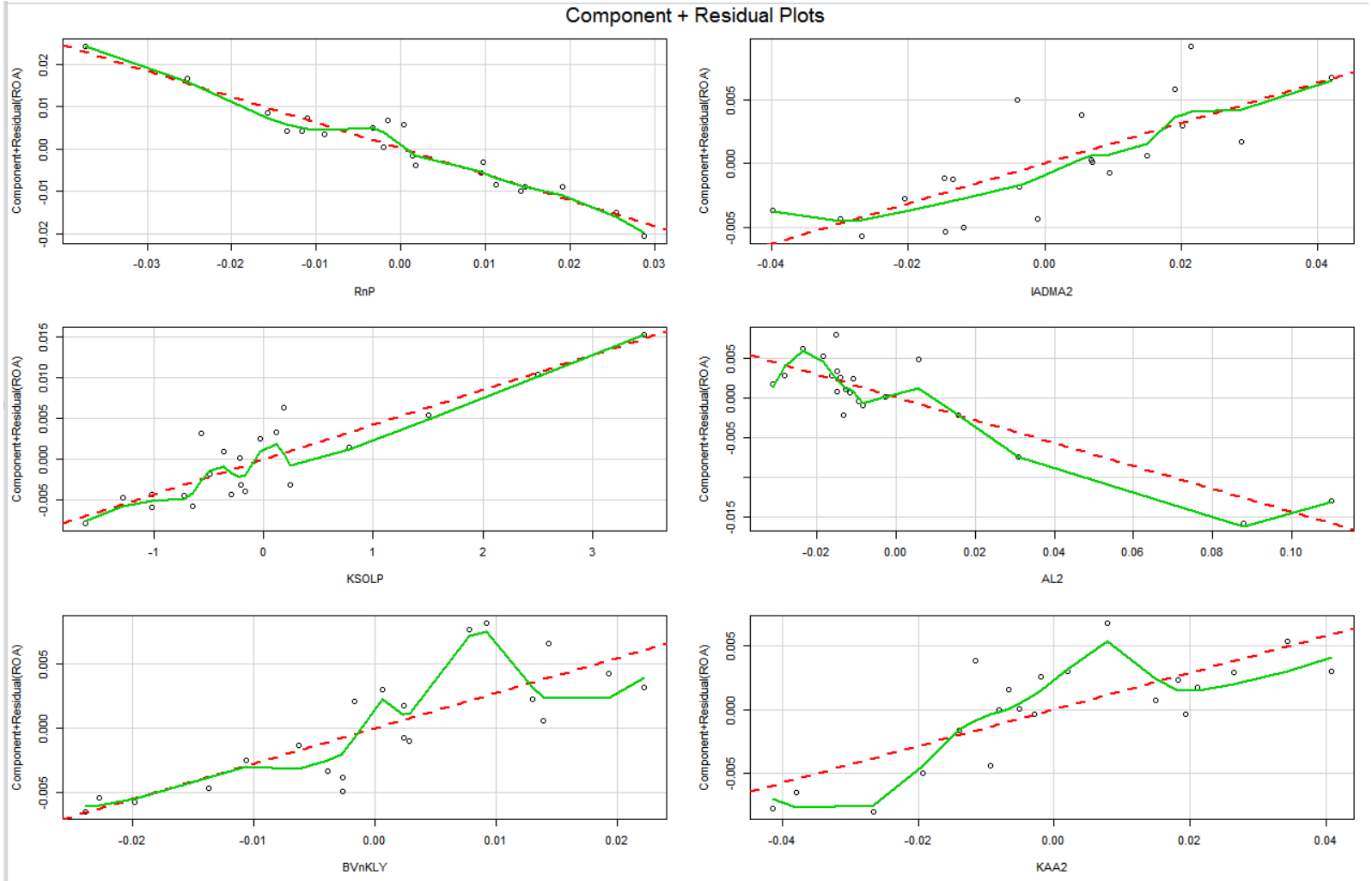
Breusch-Pagan test: Chisquare = 0.02327279, Df = 1, p = 0.8787498

Normality of the error





Linearity check



Predicting Financial Leverage

Using stepwise regression:

```
Residuals:
      Min        1Q    Median        3Q       Max
-0.041000 -0.011089  0.001104  0.009416  0.031370

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.14823    0.04451   3.331 0.004238 **
Zadl2        0.77128    0.04661  16.548 1.74e-11 ***
WzGNSOD     -0.15876    0.03880  -4.091 0.000852 ***
FM          -0.21393    0.03845  -5.563 4.28e-05 ***
DZP          0.10800    0.04171   2.589 0.019772 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.01798 on 16 degrees of freedom
Multiple R-squared:  0.9929, Adjusted R-squared:  0.9912
F-statistic: 562.4 on 4 and 16 DF,  p-value: < 2.2e-16
```

Predictors for model 2

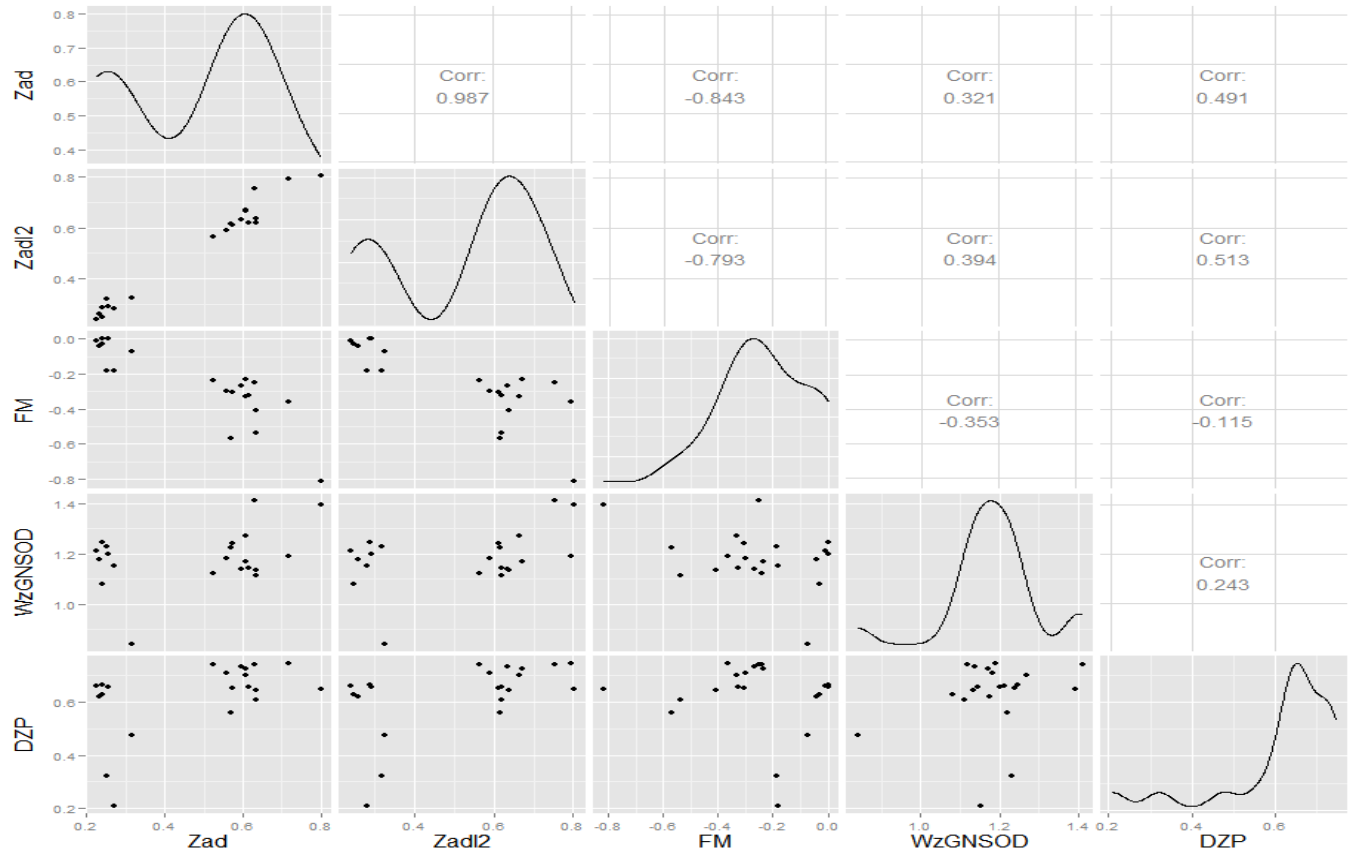
Financial ratios:

$$\boldsymbol{Zadl2} = \frac{average\ liabilities - average\ cash}{average\ liabilities\ and\ equity - average\ cash},$$

$$\boldsymbol{WzGNSOD} = \frac{incoming\ cash\ flow\ from\ operations}{cash\ outflow\ from\ perations},$$

$$\boldsymbol{FM} = \frac{net\ working\ capital}{equity}, \qquad \boldsymbol{DZP} = \frac{long\text{-}term\ debt}{liabilities}$$

Multicollinearity check for model 2



VIFs:

Zadl2	WzGNSOD	FM	DZP
5.285752	1.204711	3.986270	2.002873

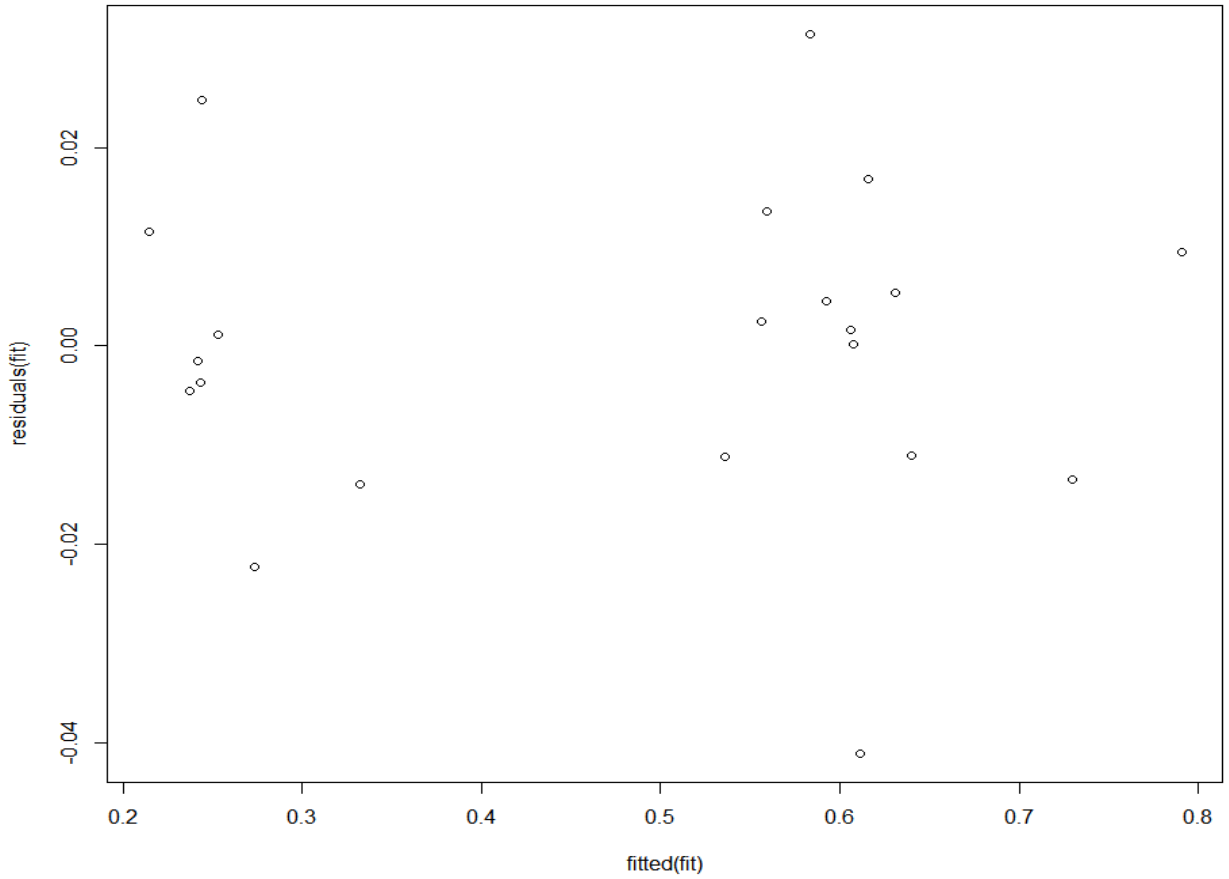
After applying the multicollinearity correction algorithm

```
Residuals:
      Min        1Q    Median        3Q       Max
-0.041000 -0.011089  0.001104  0.009416  0.031370

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.24299    0.04319  -5.625 3.80e-05 ***
Zadl2        0.77128    0.04661  16.548 1.74e-11 ***
WzGNSOD      0.31815    0.03654   8.707 1.82e-07 ***
FM          -0.75246    0.02059 -36.540  < 2e-16 ***
DZP          0.56053    0.03057  18.336 3.64e-12 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.01798 on 16 degrees of freedom
Multiple R-squared:  0.9929, Adjusted R-squared:  0.9912
F-statistic: 562.4 on 4 and 16 DF,  p-value: < 2.2e-16
```

VIFs:

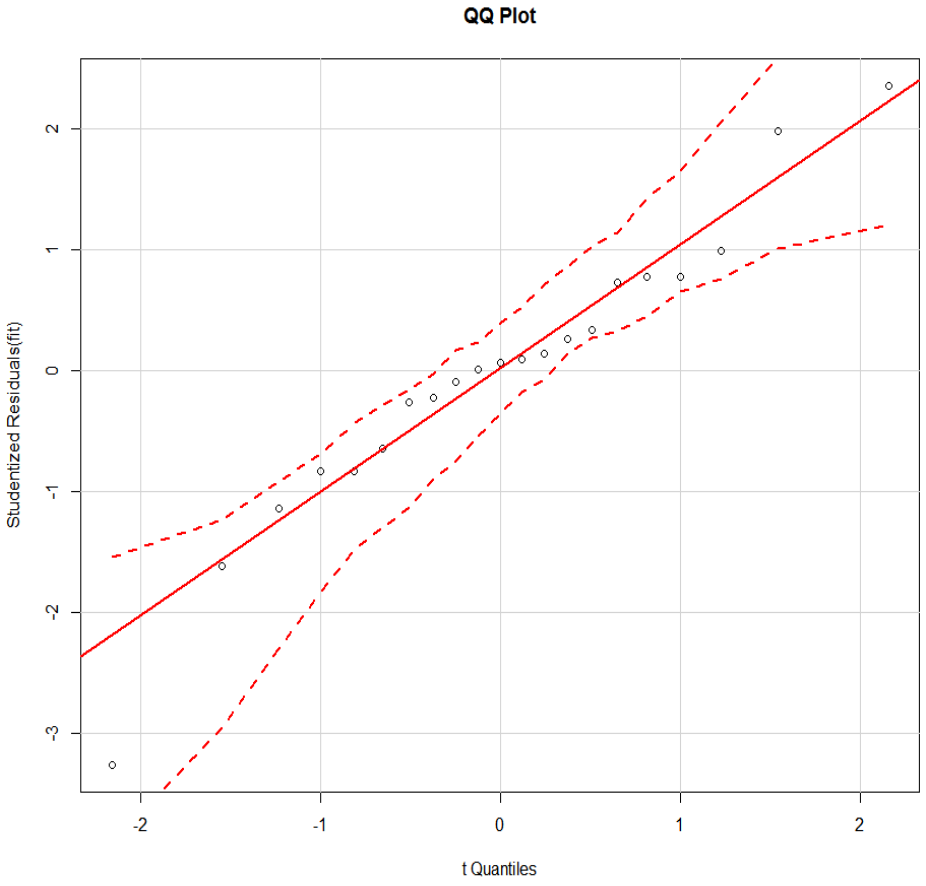
Zadl2	WzGNSOD	FM	DZP
1.022277	1.068233	1.000011	1.075705

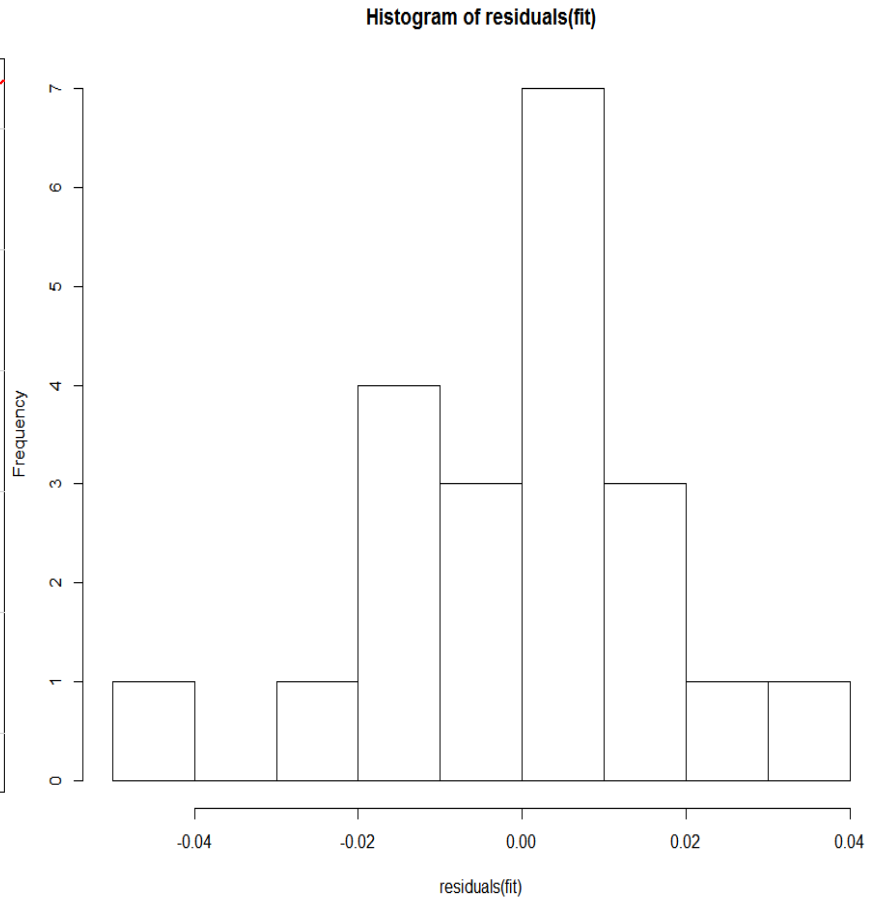
Heteroscedasticity for model 2



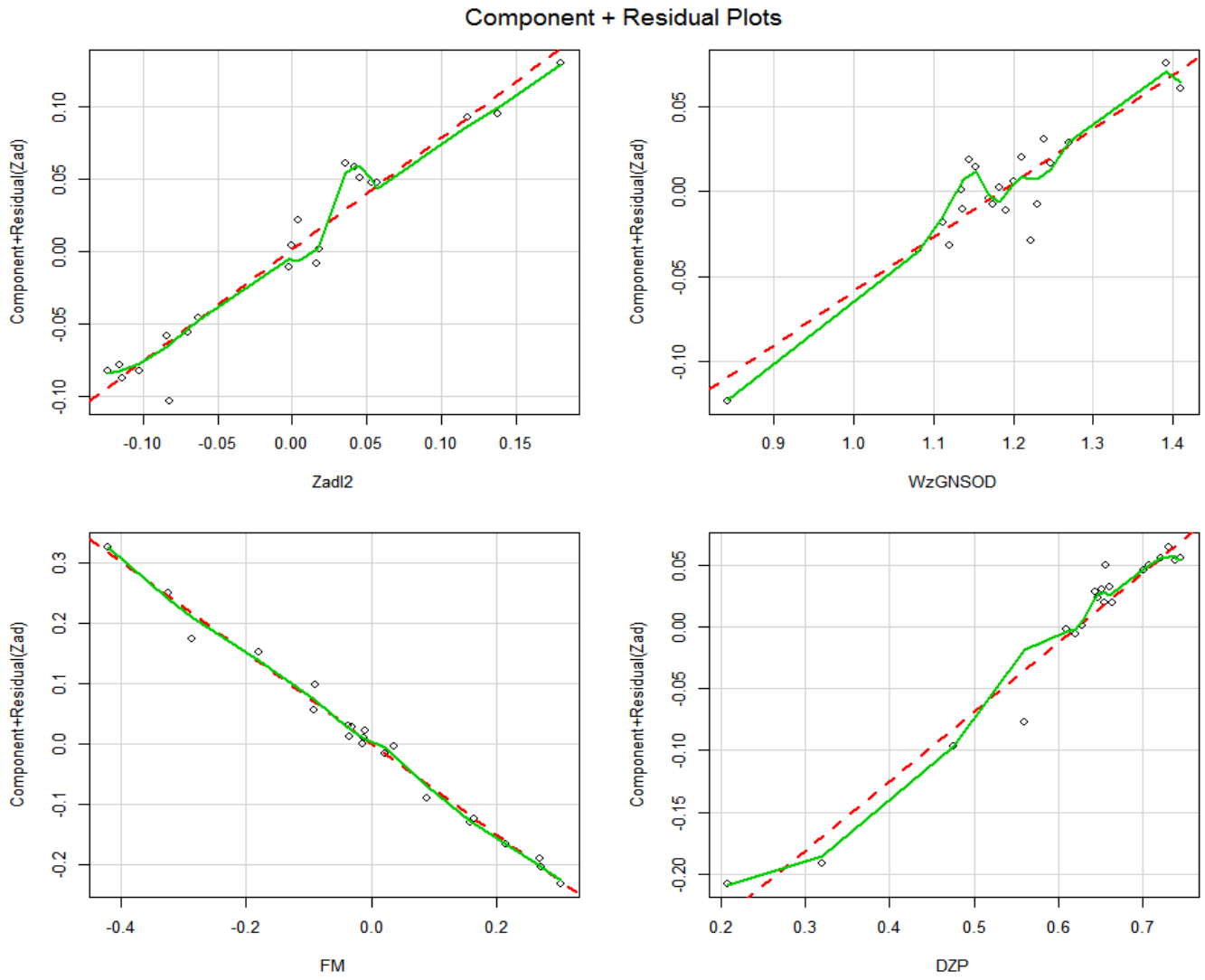
Breusch-Pagan test: Chisquare = 0.2815658, Df = 1, p = 0.5956768

Normality of the error for model 2





Linearity test for model 2



The completed models

- **ROA model**:

$$\boldsymbol{ROA} = -0.6106919 * \boldsymbol{Residuals_{RnP}} + 0.1575532 * \boldsymbol{Residuals_{IADMA2}} + 0.0042944 * \boldsymbol{Residuals_{KSOLP}} - 0.1437924 * \boldsymbol{Residuals_{AL2}} + 0.2734652 * \boldsymbol{Residuals_{BVnKLY}} + 0.1446839 * \boldsymbol{Residuas_{KAA2}} + 0.0557282$$

$$\boldsymbol{ROA} = -0{,}6106919 * \boldsymbol{RnP} + 0{,}75303 * \boldsymbol{IADMA2} + 0{,}0114157 * \boldsymbol{KSOLP} + 0{,}226665 * \boldsymbol{AL2} + 1{,}2273515 * \boldsymbol{BVnKLY} + 0{,}395315 * \boldsymbol{KAA2} - 0{,}0088392 * \boldsymbol{InPoP} + 0{,}0309$$

- **FL model:**

$$\boldsymbol{FL} = 0{,}77128 * \boldsymbol{Zadl2} - 0{,}1587667 * \boldsymbol{WzGNSOD} - -0{,}2139291 * \boldsymbol{FM} + 0{,}1079993434 * \boldsymbol{DZP} + 0{,}1482290032$$

Conclusions

Both models are economically reasonable and can be used from economists, managers and other stakeholders for their specific purposes. Each coefficient and it’s sign is an important source of information showing how much the variable which is related to it is contributing for the final outcome. Low dispersions are achieved for each of the coefficients which is increasing the chance for reliable conclusions.

Future research

- A technique allowing to overcome the differences between the firm's accounting politics;
- Based on this technique much more profound models predicting large economic sectors and each of the firms in them can be made;
- Regression models can be unbiased but in the same time are not consistent estimators. Is there a mathematical model that has both of this qualities? Hazard models?

References

- SIMINICĂ MARIAN, CIRCIUMARU D, SIMION D, **“**Statistical correlations between the return and the indicators of financial balance. Case study: The romanian companies listed on BSE”, University of Craiova, 2011
- Abdul Ghafoor Khan, “The relationship of capital structure decisions with firm performance: A study of the engineering sector of Pakistan”, *International Journal of Accounting and Financial Reporting* ISSN 2162-3082 2012, Vol. 2, No. 1
- San, O.T. and Heng, T.B. (2011), “Capital Structure and Corporate Performance of Malaysian Construction Sector”, *International Journal of Humanities and Social Science*, Voi.1 No.2. pp.28-36.
- *Rooh Ollah Arab, Seyed S. M .and Azadeh B, “*Financial Performance of the Steel Industry in India: A Critical Analysis”, *Middle-East Journal of Scientific Research* 23 (6): 1085-1090, 2015

Thank you for the attention!

"All models are wrong but some are useful"

George E.P. Box