Interest Rates & Bond Supply Study

Thomas Vermaelen

6/27/2019

Summary

In this study, we attempt to find an IV to study the relationship between price informativeness and US interest rates. The ratio of US publicly owned debt to GDP is used as an IV.

As an initial step, we regress Annual US Treasury rates (treasury_yield) on Bond supply (Debt/GDP). We then run the diagnostics for the regression and fit the appropriate transformations to the data.

We then look at the significance of adding a linear time trend. The trend variable ends up being significant without affecting the significance of our IV, log(Debt/GDP)

In order to capture the autocorrelation of the residuals from our initial (transformed) regression we evaluate two different techniques:

- The Feasible Generalized Least Squares (FGLS) regression (Pairs-Winsten), where the independent variable log(Debt/GDP) is assumed to be exogenous
- OLS using serial correlation-robust standard error of the coefficient on log(Debt/GDP) (Newey-West's HAC)

In both cases, we find log(Debt/GDP) to be significant

Variables

All variables are stored in the table $bond.data$:

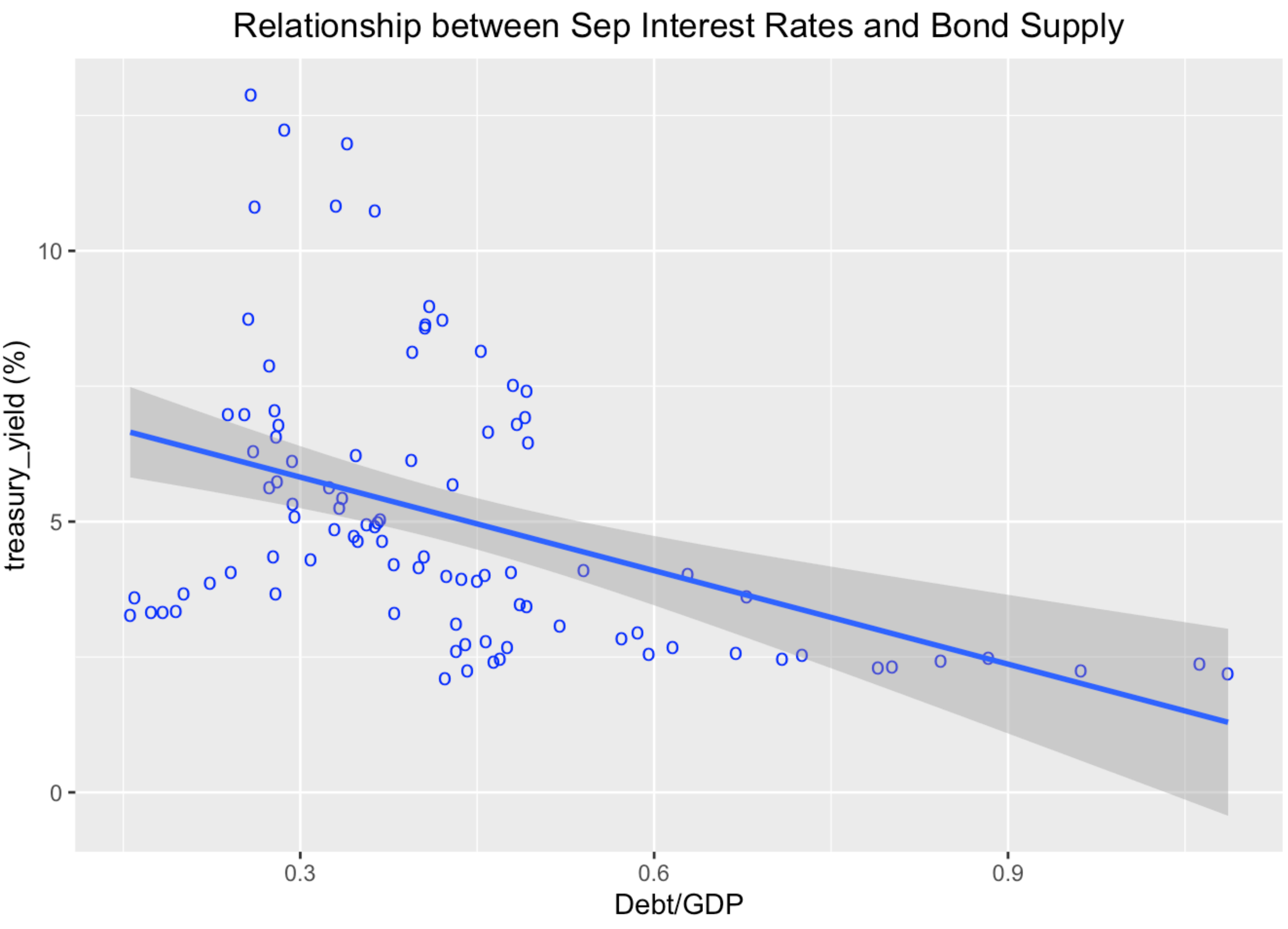
- $FiscalYear$, US fiscal Year from 1919 to 2012, which is assumed to run from June 30 of the previous year up to September 30 of the given year (e.g. 2012 FY runs from June 30, 2011 to September 30, 2012). This is according to Krishnamurthy & Vissing-Jorgensen (2012), Appendix B
- $bond\_supply$ (Debt/GDP) is the supply of Treasuries scaled by GDP, at market value, calculated as: (Debt/GDP)*(Total mkt value of T-Bills/total face value of T-Bills), for the given fiscal year. The series comes from Henning Bohn and has

been annualized.

- *treasury_yield* is the 10-Year Treasury Constant Maturity yield on September 30 (end of FY in the US)

Regression Analysis

1. a) Regressing treasury_yield on Debt/GDP

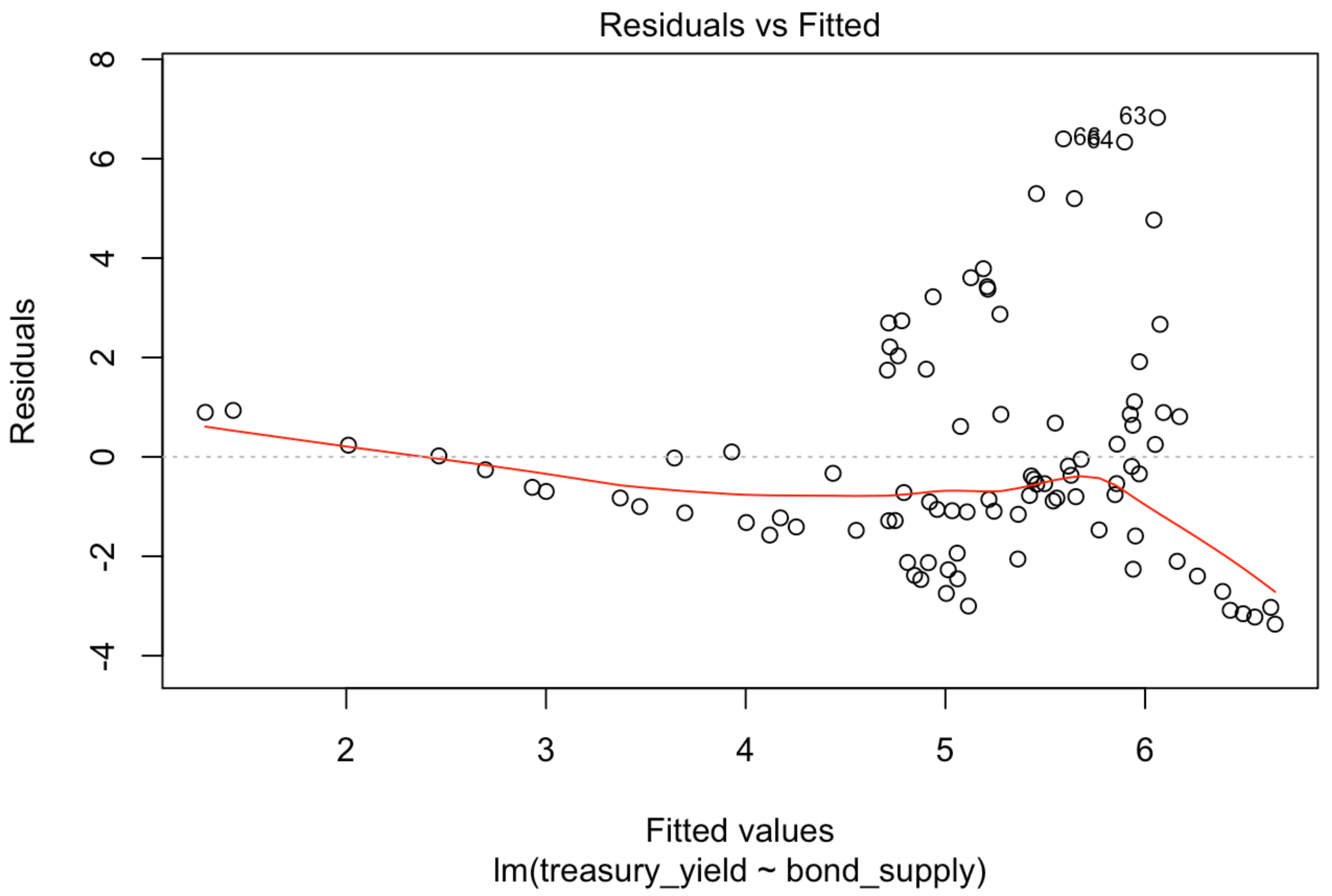


```
##
## Call:
## lm(formula = treasury_yield ~ bond_supply, data = bond.data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -3.365 -1.387 -0.544  0.883  6.827
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    7.548      0.594  12.707  < 2e-16 ***
## bond_supply   -5.756      1.270  -4.531 1.76e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.3 on 92 degrees of freedom
## Multiple R-squared:  0.1824, Adjusted R-squared:  0.1735
## F-statistic: 20.53 on 1 and 92 DF,  p-value: 1.761e-05
```

Regression analysis yields a significant coefficient of -5.756 on Debt/GDP implying that an increase in the ratio of Debt/GDP by 1 leads to a decrease in the interest rate by 5.76 percentage points. This interpretation is not very intuitive, therefore, it would make more sense to take logs of the variables

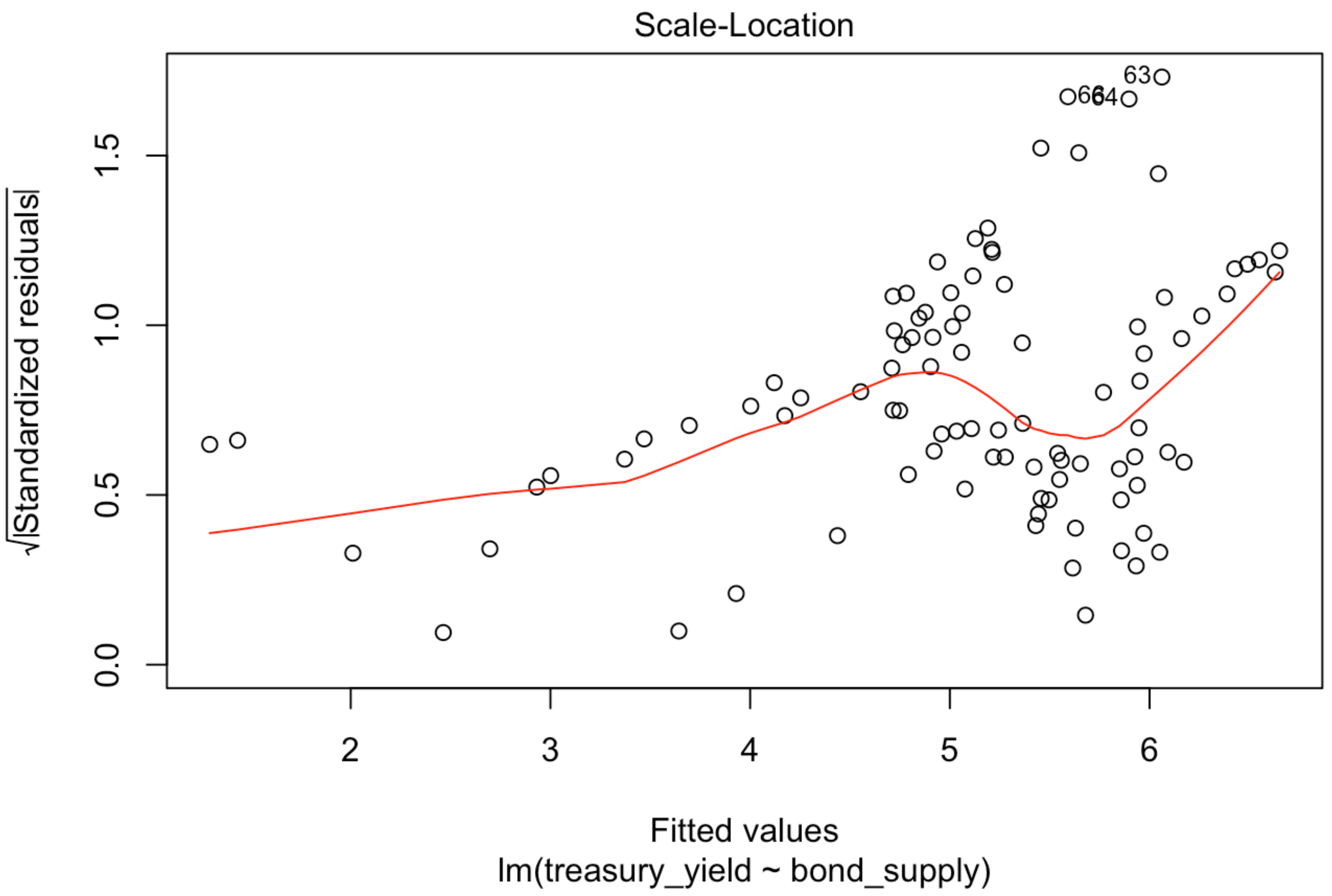
1. b) Regression Diagnostics for treasury_yield ~ Debt/GDP

####Linearity



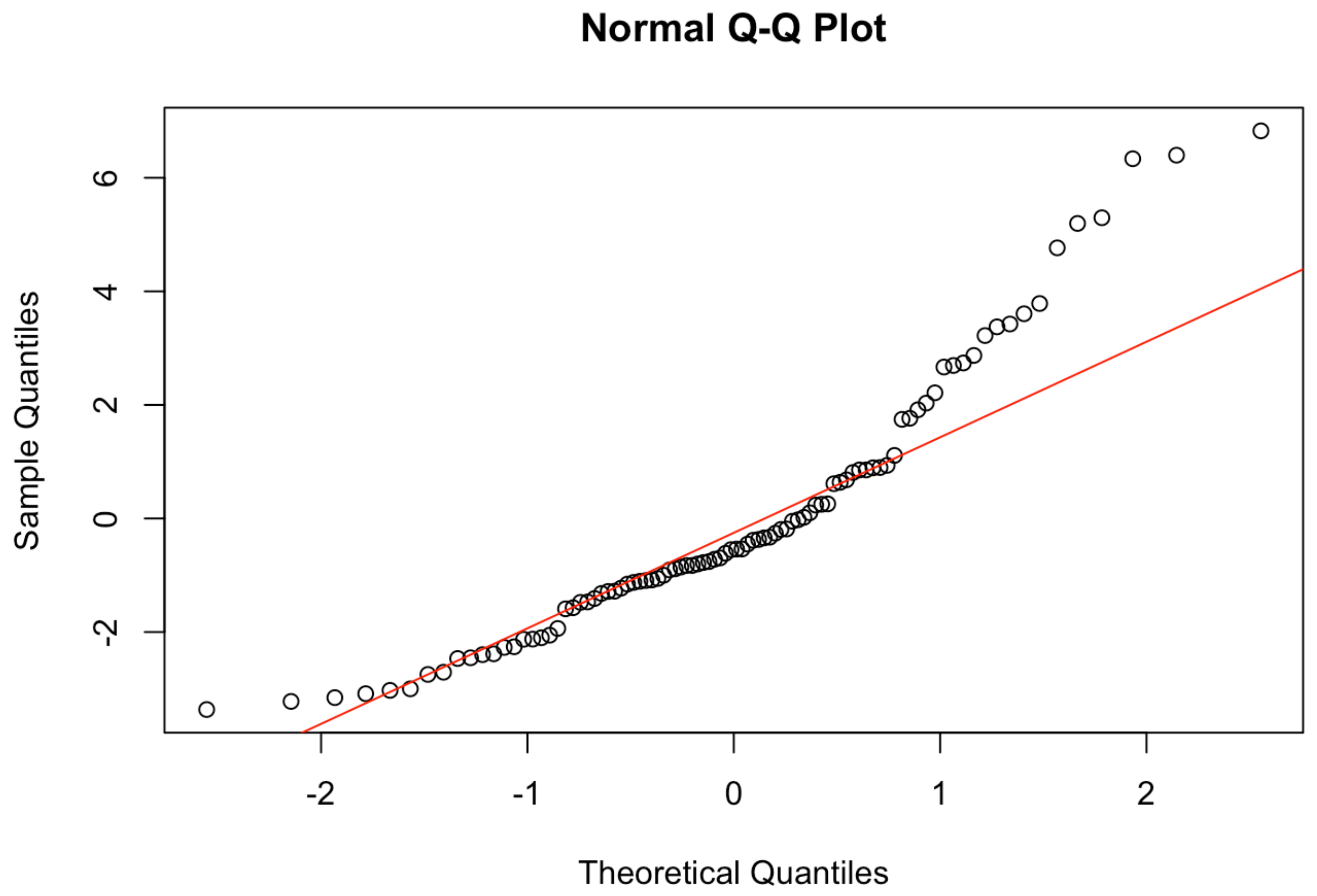
Non-linearity is present. This suggests that a non-linear transformation of the independent variable is more appropriate

Homoskedasticity:



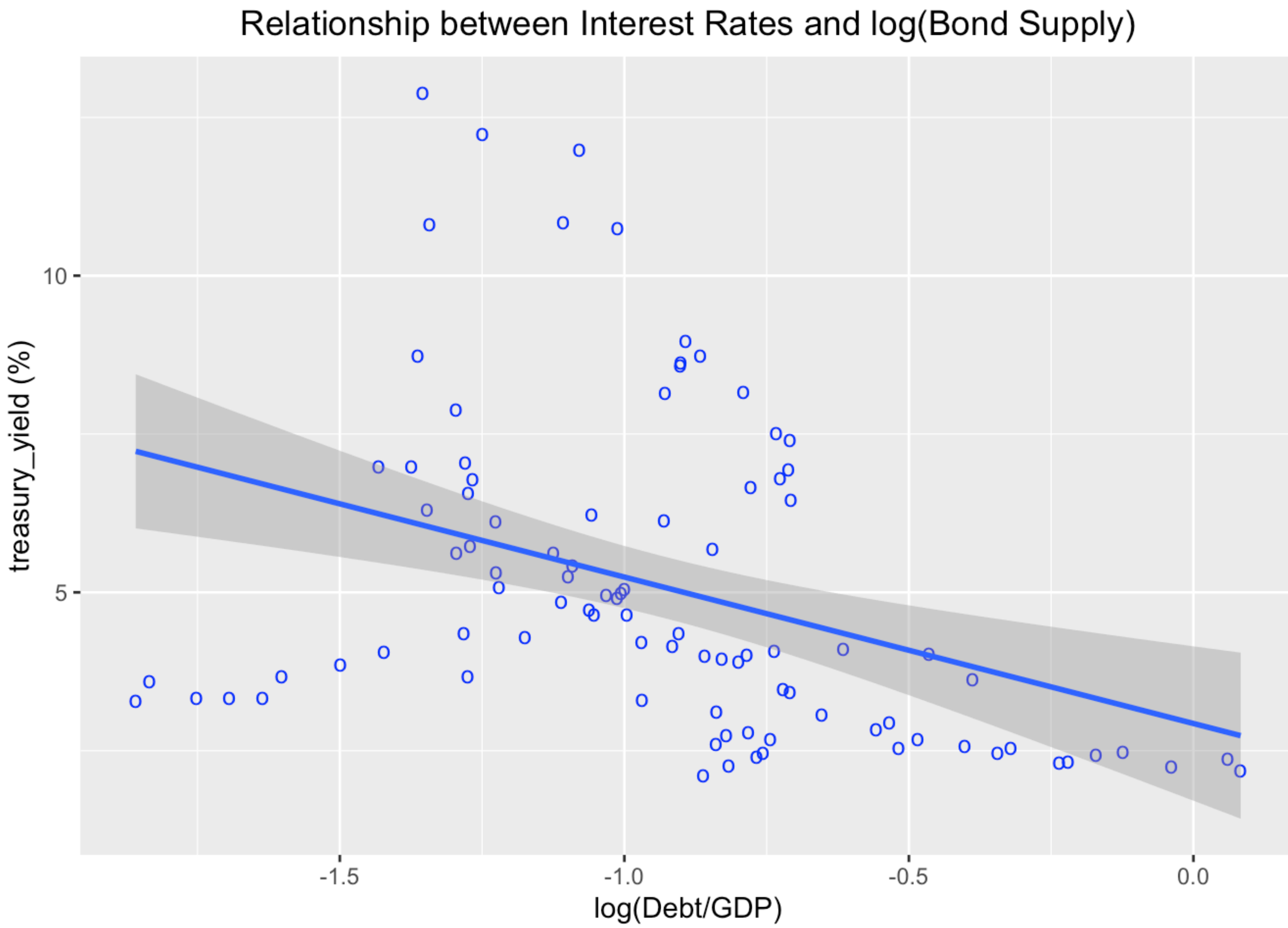
Variance increases for larger values of Y implying that a log-transformation of the dependent variable is needed

Normality of Residuals:



Residuals seem to be non-normally distributed

2. a) Regressing treasury_yield on log(Debt/GDP)

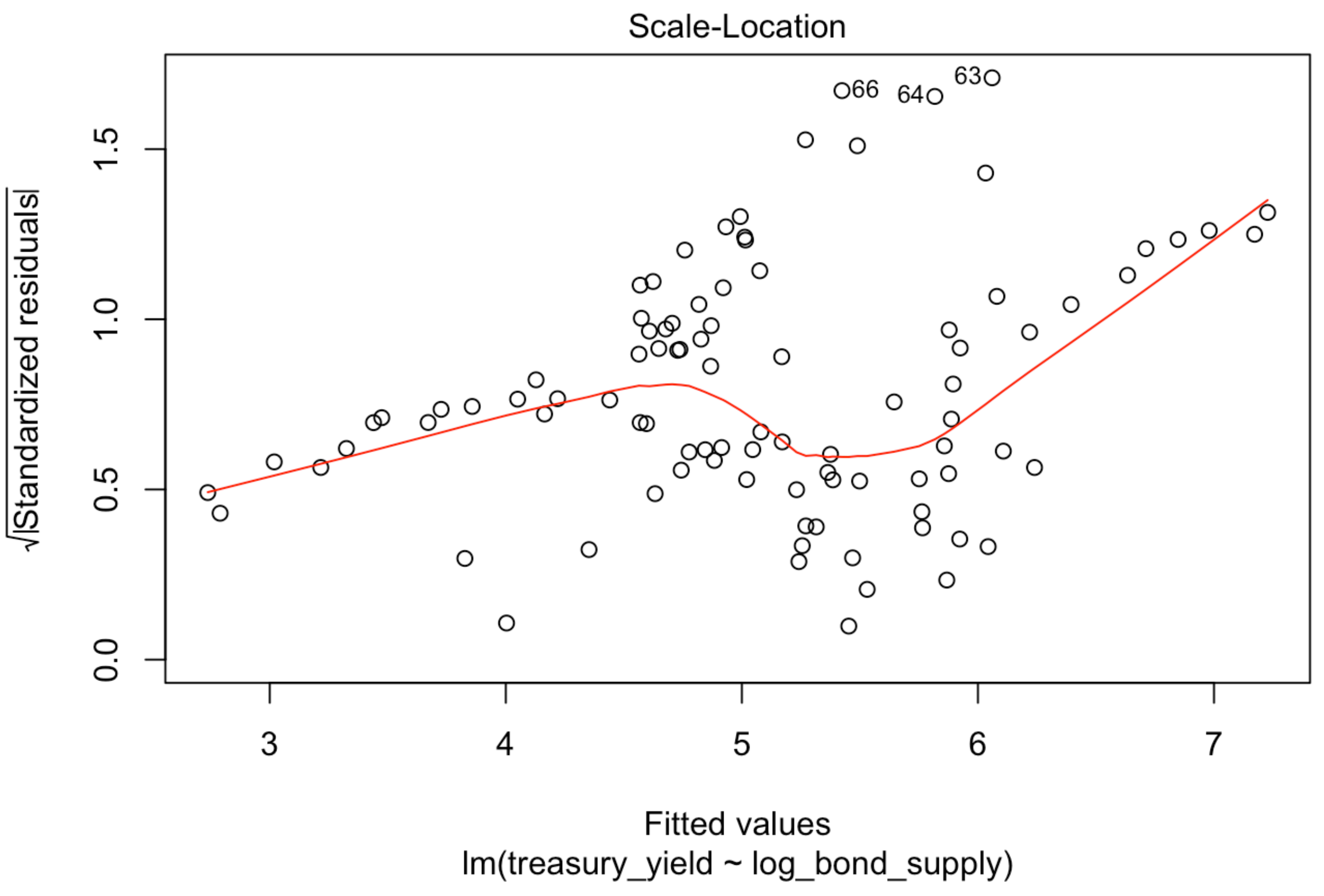



```
##
## Call:
## lm(formula = treasury_yield ~ log_bond_supply, data = bond.dat
a)
##
## Residuals:
##     Min       1Q   Median       3Q      Max
## -3.9413 -1.3596 -0.6160  0.8711  6.8283
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)       2.9293     0.6132   4.777 6.69e-06 ***
## log_bond_supply  -2.3124     0.6048  -3.824 0.000239 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.363 on 92 degrees of freedom
## Multiple R-squared:  0.1371, Adjusted R-squared:  0.1277
## F-statistic: 14.62 on 1 and 92 DF,  p-value: 0.0002391
```

Regression reveals a coefficient of -2.312 for log(bond supply) that is significant at a 5% level signifying that a 1% change in Debt/GDP leads to a decrease in the interest rate by aproximately 0.02 percentage points. R- squared = 0.137 (slightly better than *above)*

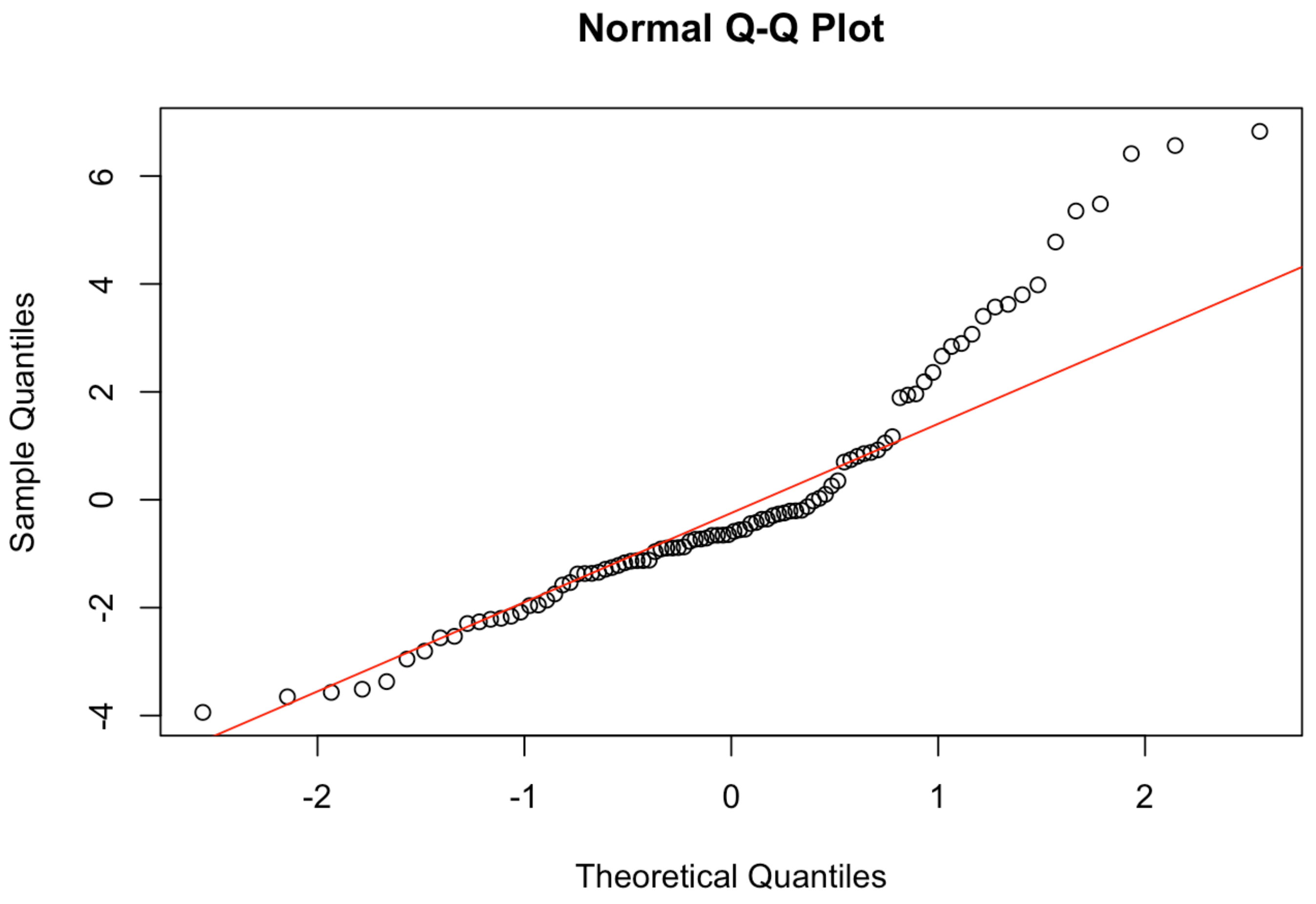
2. b) Regression diagnostics for treasury_yield ~ log(Debt/GDP)

Homoskedasticity:



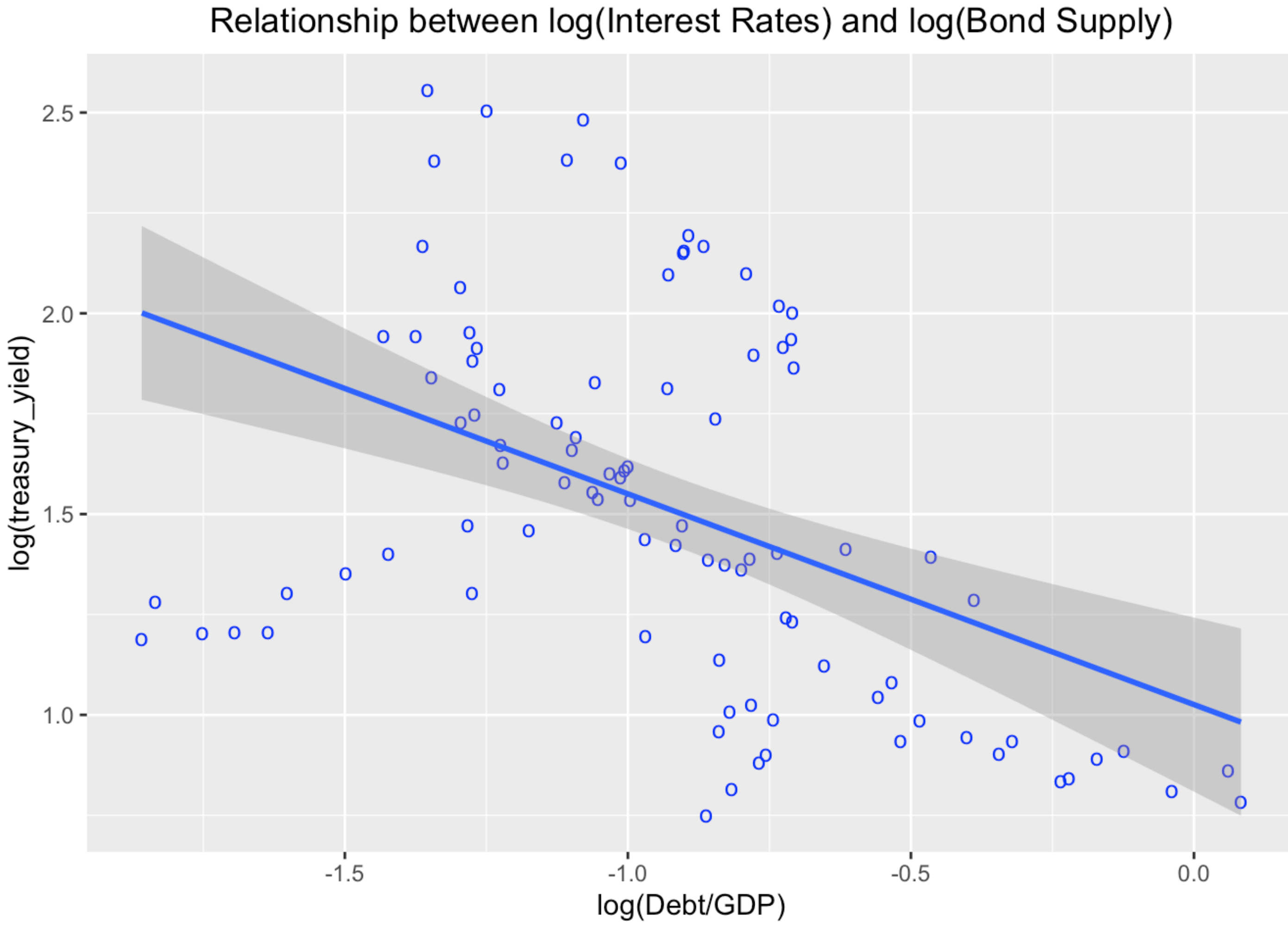
Variance seems to increase at larger values of log(bond supply) still suggesting a log-transformation of treasury yield

Normality of Residuals:



Residuals do not seem to be normally distributed

3 . a) Regressing Log(treasury_yield) ~ Log(Debt/GDP)

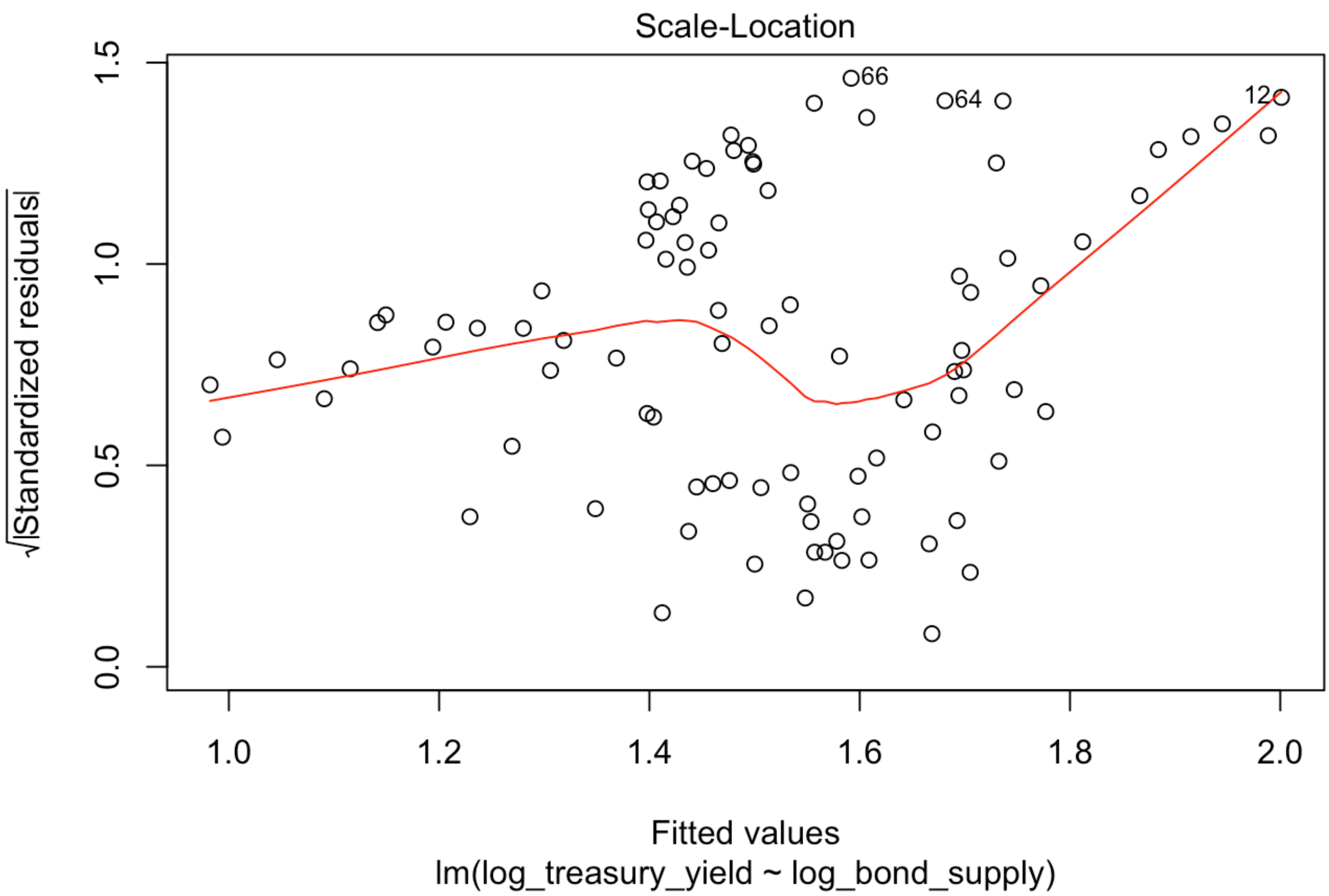



```
##
## Call:
## lm(formula = log_treasury_yield ~ log_bond_supply, data = bond
.data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.81145 -0.29314 -0.02922  0.24238  0.89214
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)       1.0256     0.1091   9.403 4.11e-15 ***
## log_bond_supply  -0.5248     0.1076  -4.879 4.46e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4203 on 92 degrees of freedom
## Multiple R-squared:  0.2055, Adjusted R-squared:  0.1969
## F-statistic:  23.8 on 1 and 92 DF,  p-value: 4.461e-06
```

Coefficient on log(bond supply) is -0.525 and significant. A 1% change in Debt/GDP leads to a 0.53% decrease in interest rates. R-squared is 0.201, which is 46.72% improvement on the previous model.

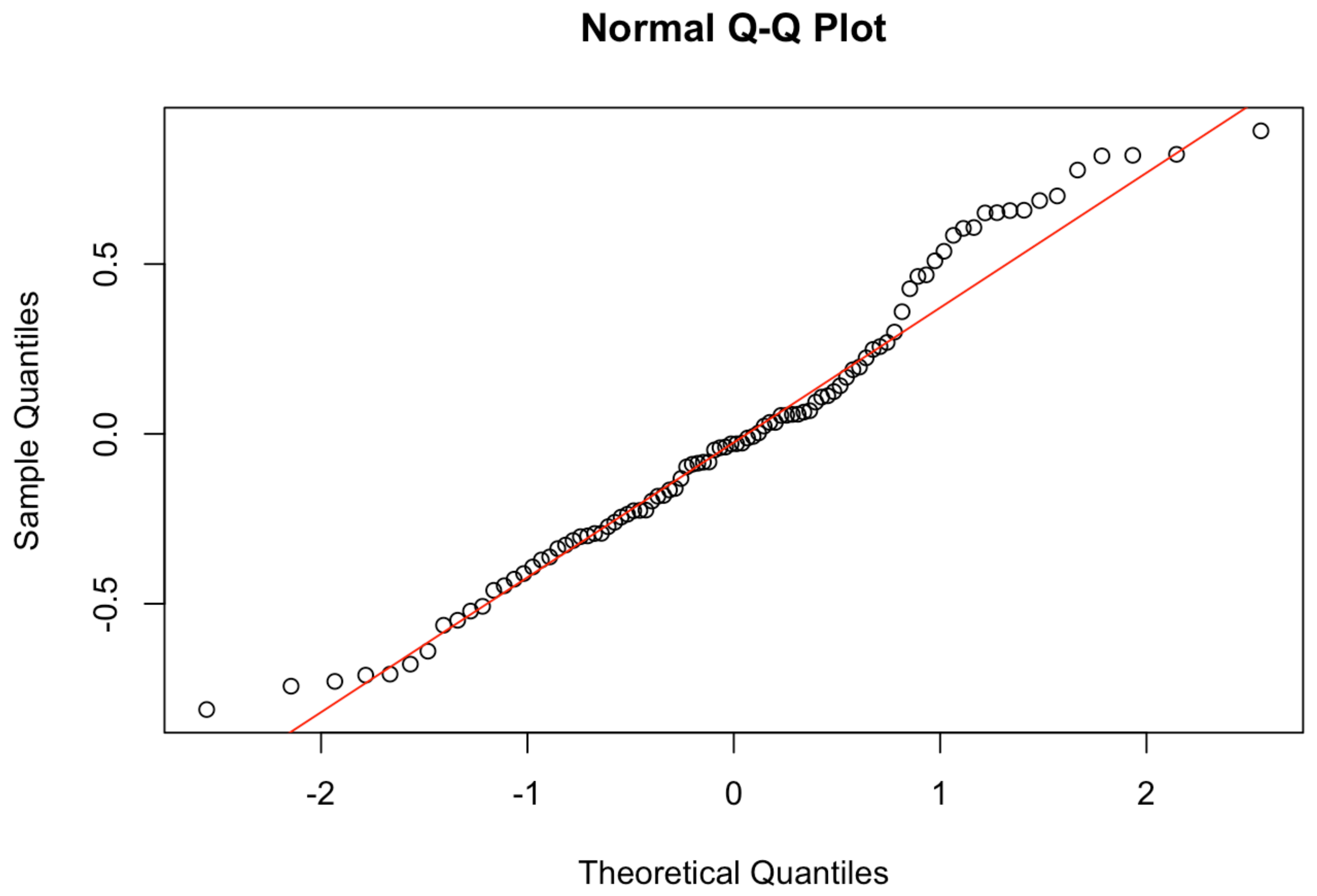
3. b) Regression diagnostics for Log(treasury_yield) ~ Log(Debt/GDP)

Homoskedasticity:



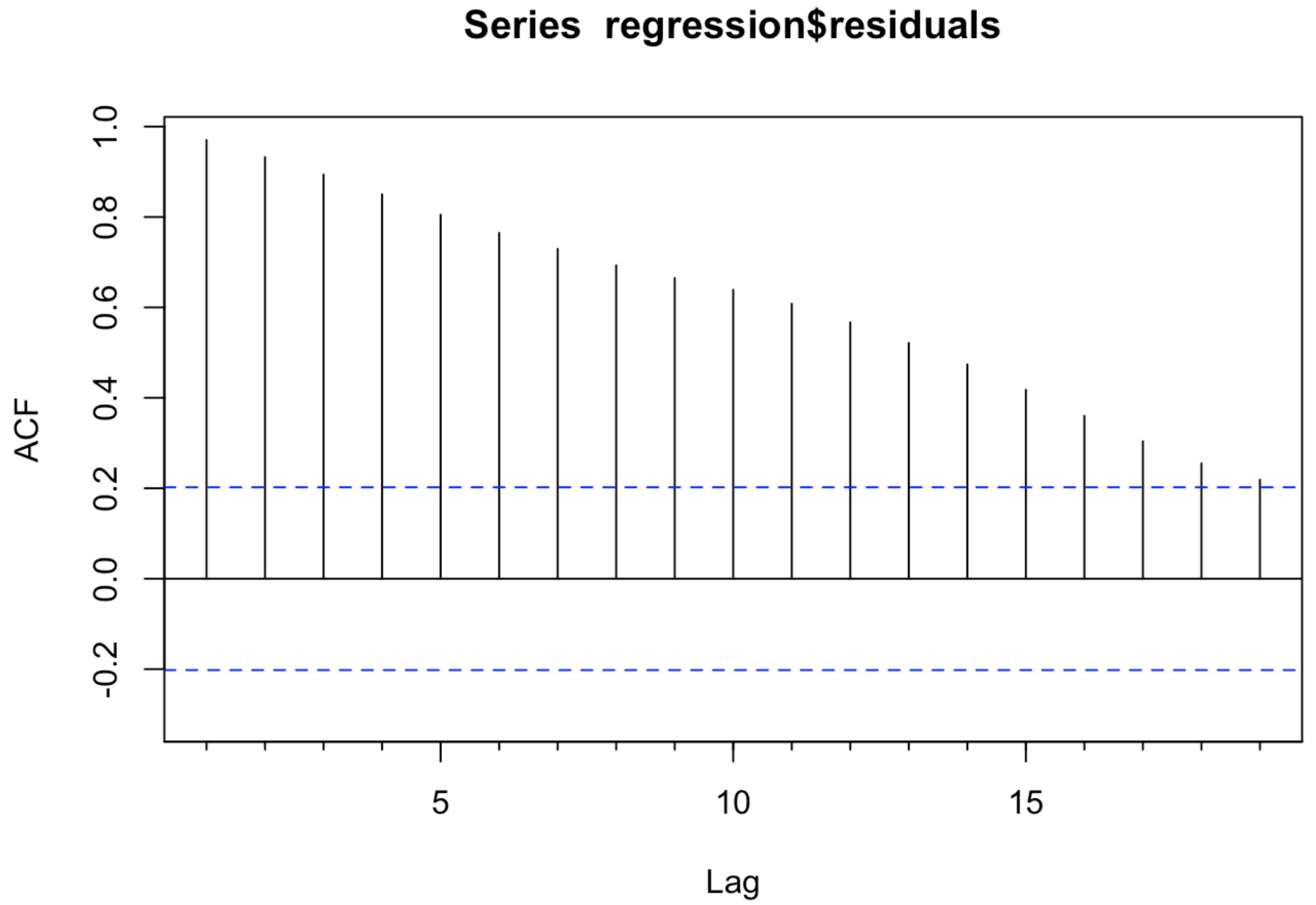
Not sure if Homoskedacity seems to improve

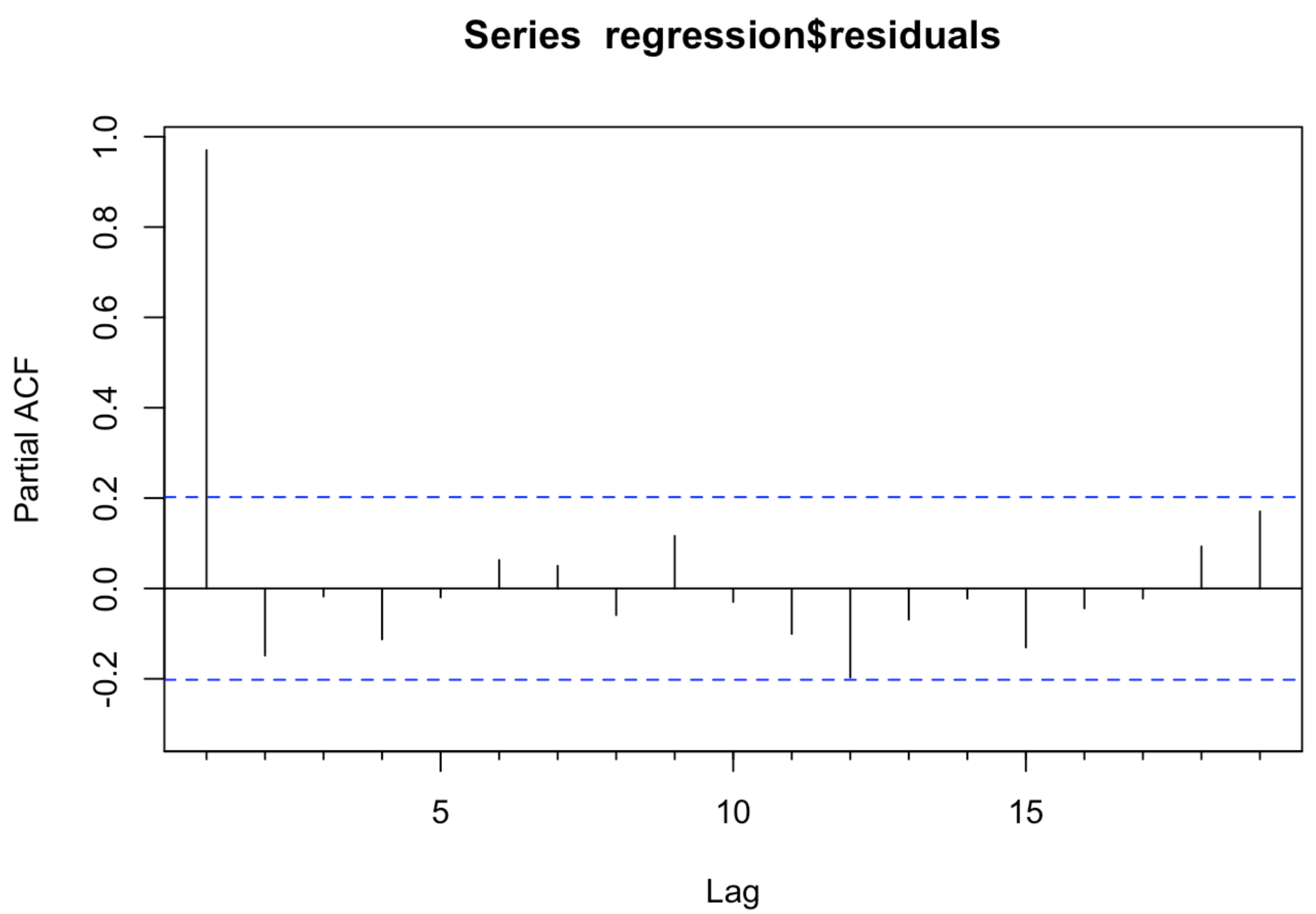
Normality of Residuals:



Normality of residuals looks better

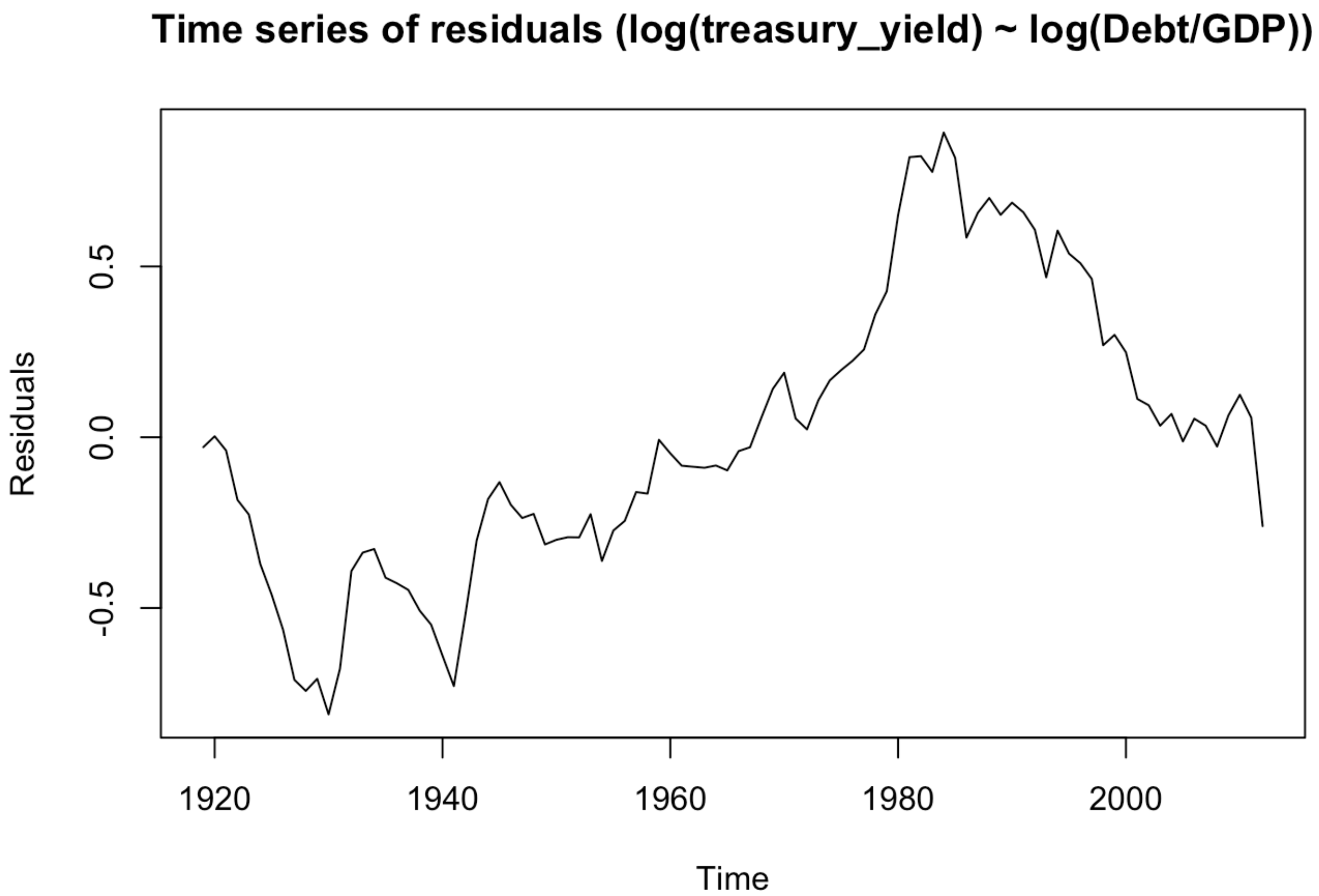
####Autocorrelation of residuals:



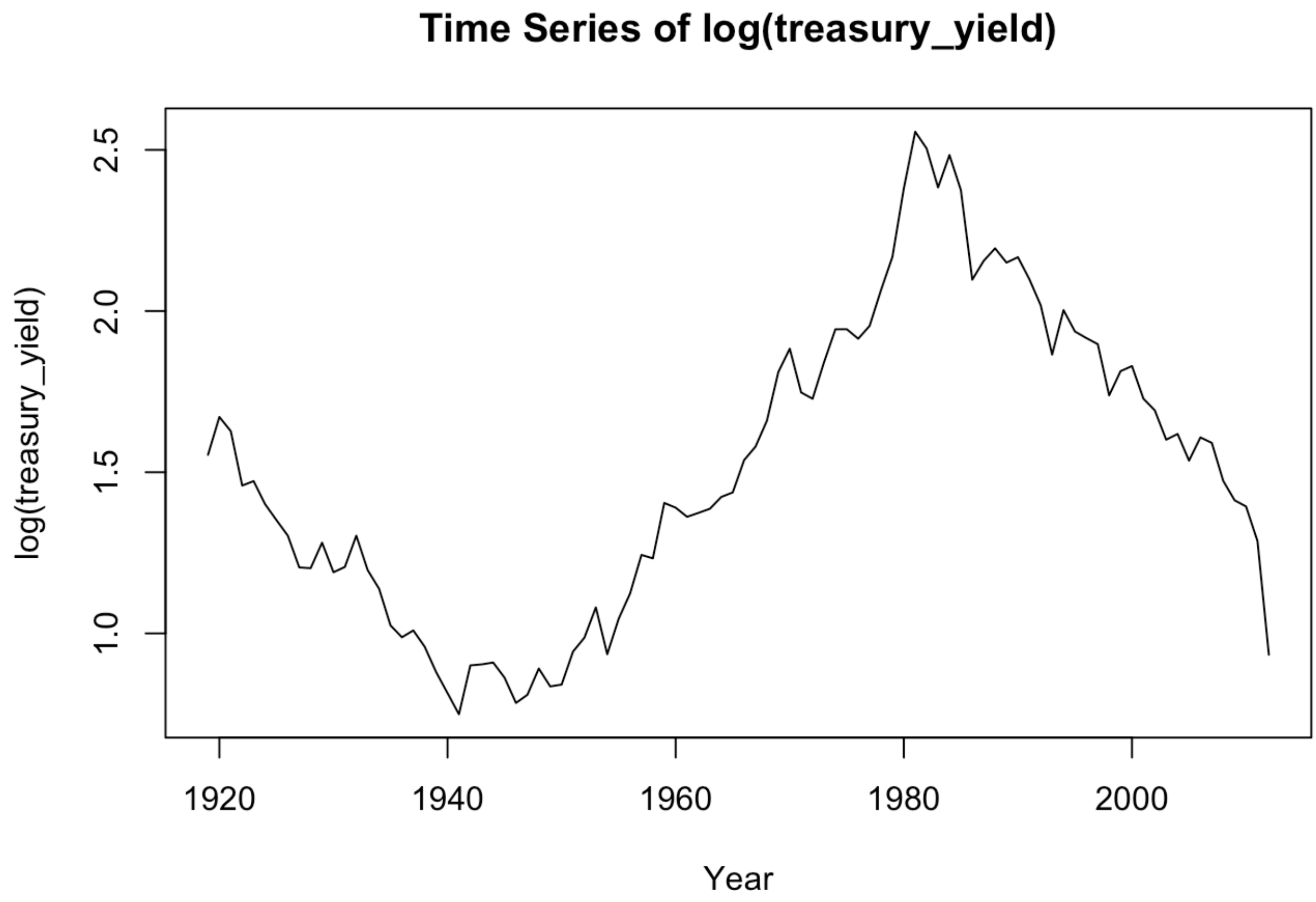


PACF suggests largelly significant lag 1 after controlling for other lags

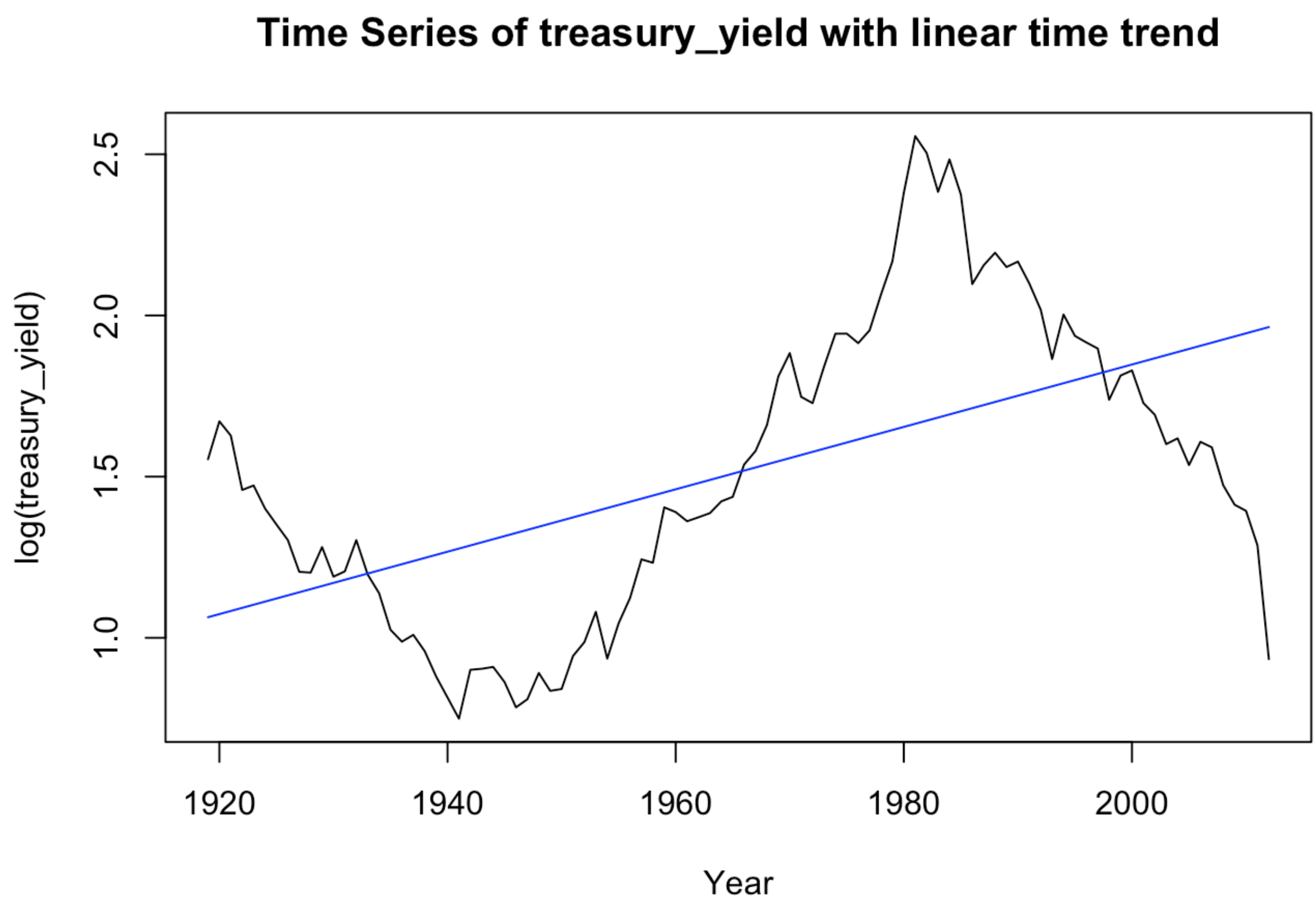
####Time series plot of residuals



Fitting Time Trend



Graph suggests perhaps a linear time trend, let's include it to the model



```
##
## Call:
## lm(formula = bond.data$log_treasury_yield ~ bond.data$trend)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.03016 -0.30879 -0.01081  0.29615  0.89239
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     1.054188   0.081053  13.006  < 2e-16 ***
## bond.data$trend 0.009678   0.001482   6.532 3.52e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3898 on 92 degrees of freedom
## Multiple R-squared:  0.3168, Adjusted R-squared:  0.3094
## F-statistic: 42.67 on 1 and 92 DF,  p-value: 3.524e-09
```

Coefficient of log(bond supply) is still significant after adding linear trend, it seems as if the relationship between log(treasury yield) and log(Debt/GDP) is not spurious. R-squred = 0.603

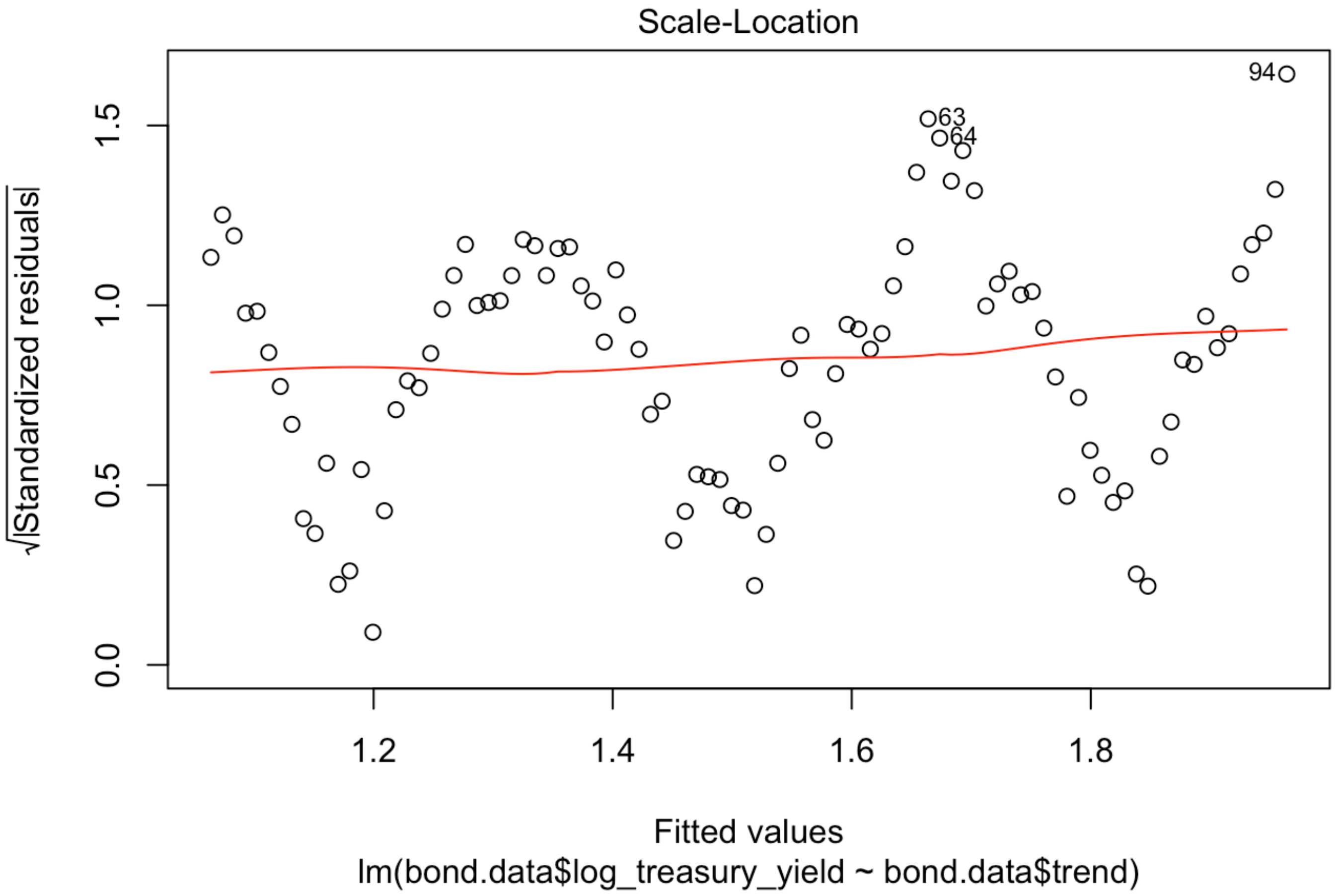
Adding back log(Debt/GDP) to the regression: log(treasury_yield) ~ Trend

```
##
## Call:
## lm(formula = bond.data$log_treasury_yield ~ bond.data$log_bond
_supply +
##     bond.data$trend)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.70786 -0.18194 -0.04292  0.18898  0.67506
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|
)
## (Intercept)                0.413584   0.100698   4.107 8.74e-0
5 ***
## bond.data$log_bond_supply -0.624457   0.077221  -8.087 2.53e-1
2 ***
## bond.data$trend            0.010934   0.001147   9.533 2.42e-1
5 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.299 on 91 degrees of freedom
## Multiple R-squared:  0.6025, Adjusted R-squared:  0.5938
## F-statistic: 68.96 on 2 and 91 DF,  p-value: < 2.2e-16
```

Coefficient of log(bond supply) is not significant, it seems as if the relationship between log(treasury yield) and log(Debt/GDP) might be spurious

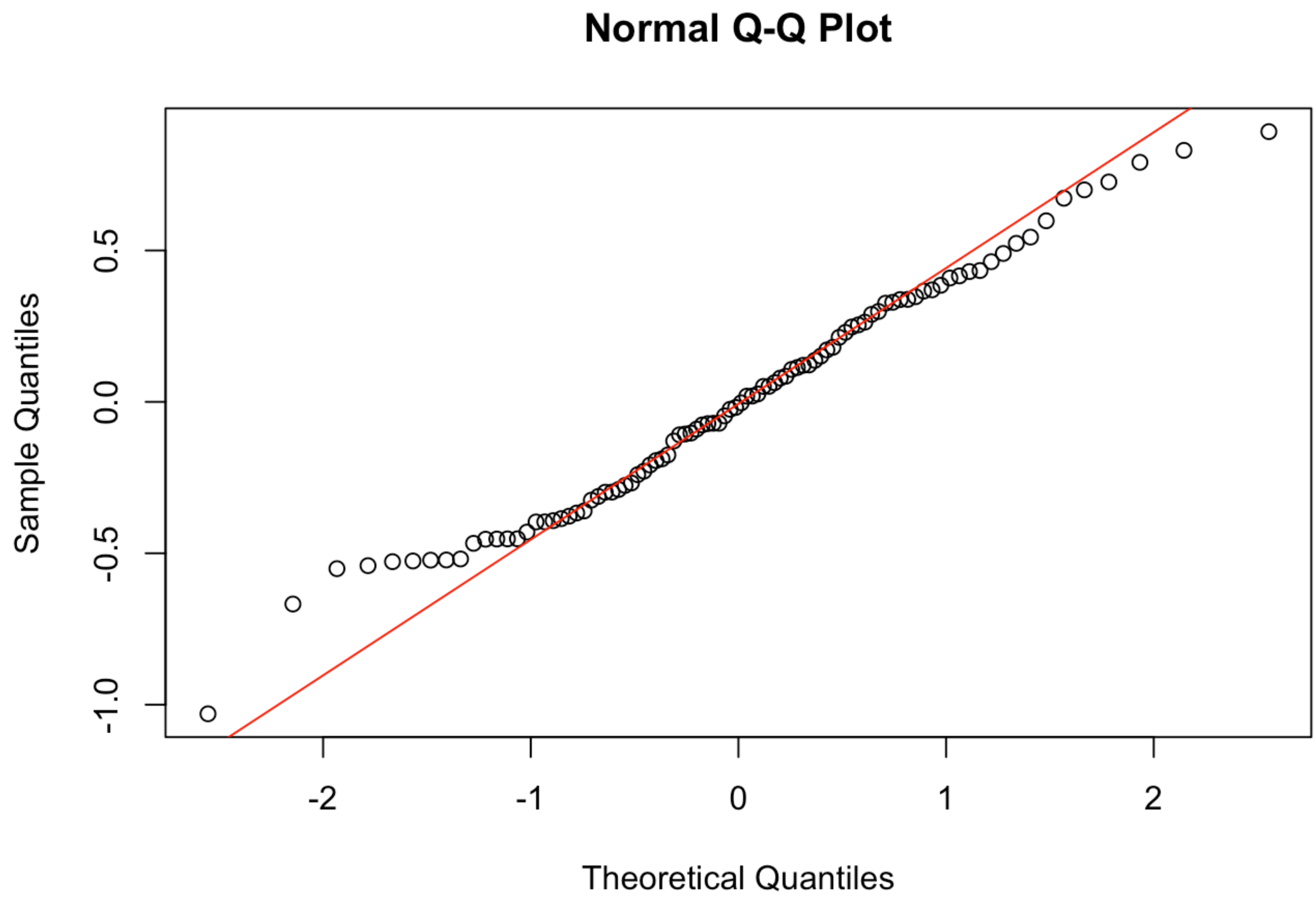
Quadratic trend model diagnostics : log(treasury_yield) ~ log(Debt/GDP) + Trend

####Homoskedasticity



Homoskedasdicity seems to improve compared to the model without linear trend

Normality



Normality still looks somewhat good

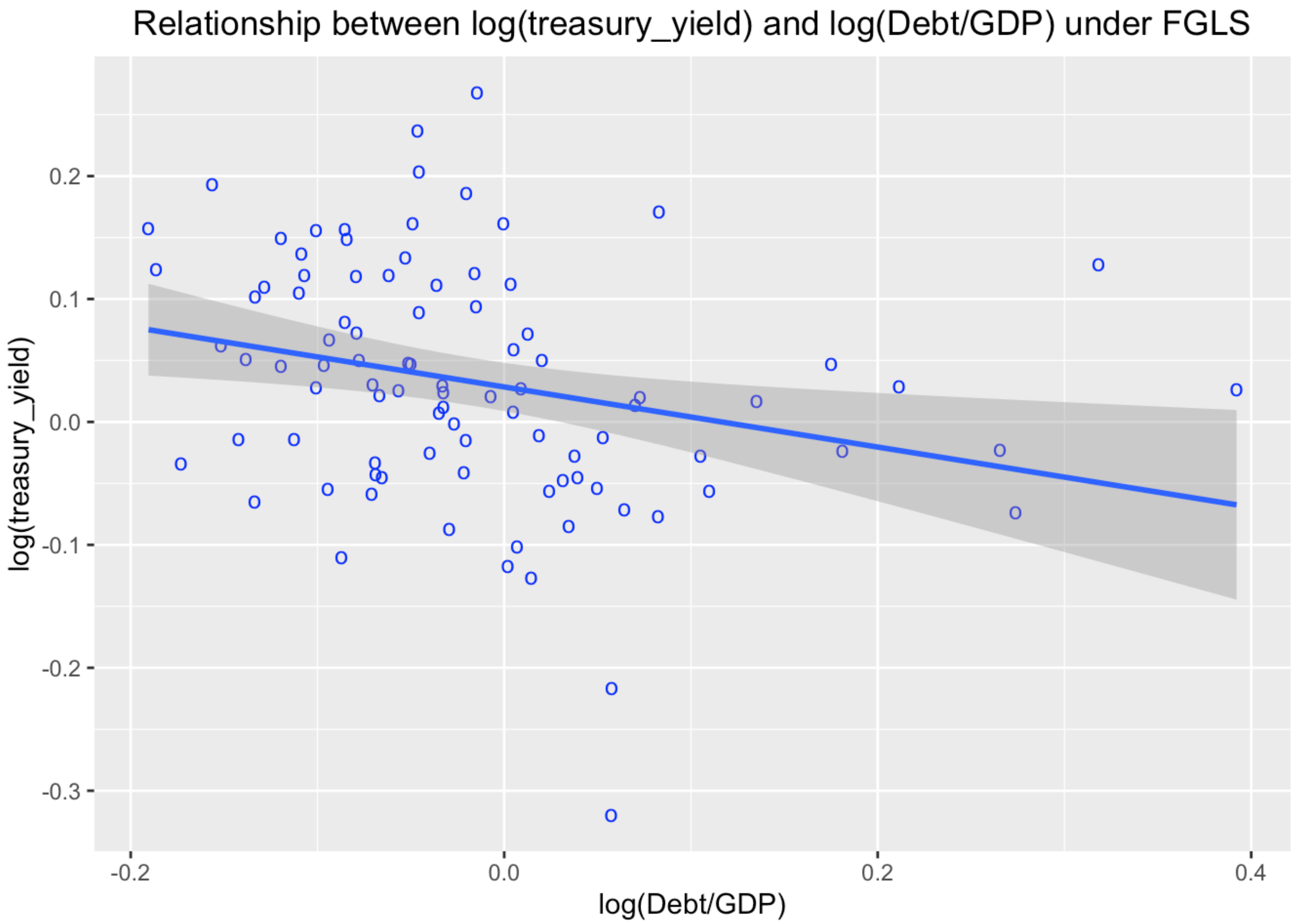
Accounting for Autocorrelation

1. Feasible Generalized Least Squares (Prais-Winsten), assuming strict exogeneity of X

Method:

- Run OLS on log(treasury_yield) and log(Debt/GDP) and obtain residuals
- Run AR(1) on residuals, save coefficient rho
- Run OLS on $log(treasury_yield)[t] - rho * log(treasury_yield)[t-1]$ and $log(Debt/GDP)[t] - rho * log(Debt/GDP)[t-1]$ (provided that the intercept coefficient is divided by $1 - rho$)

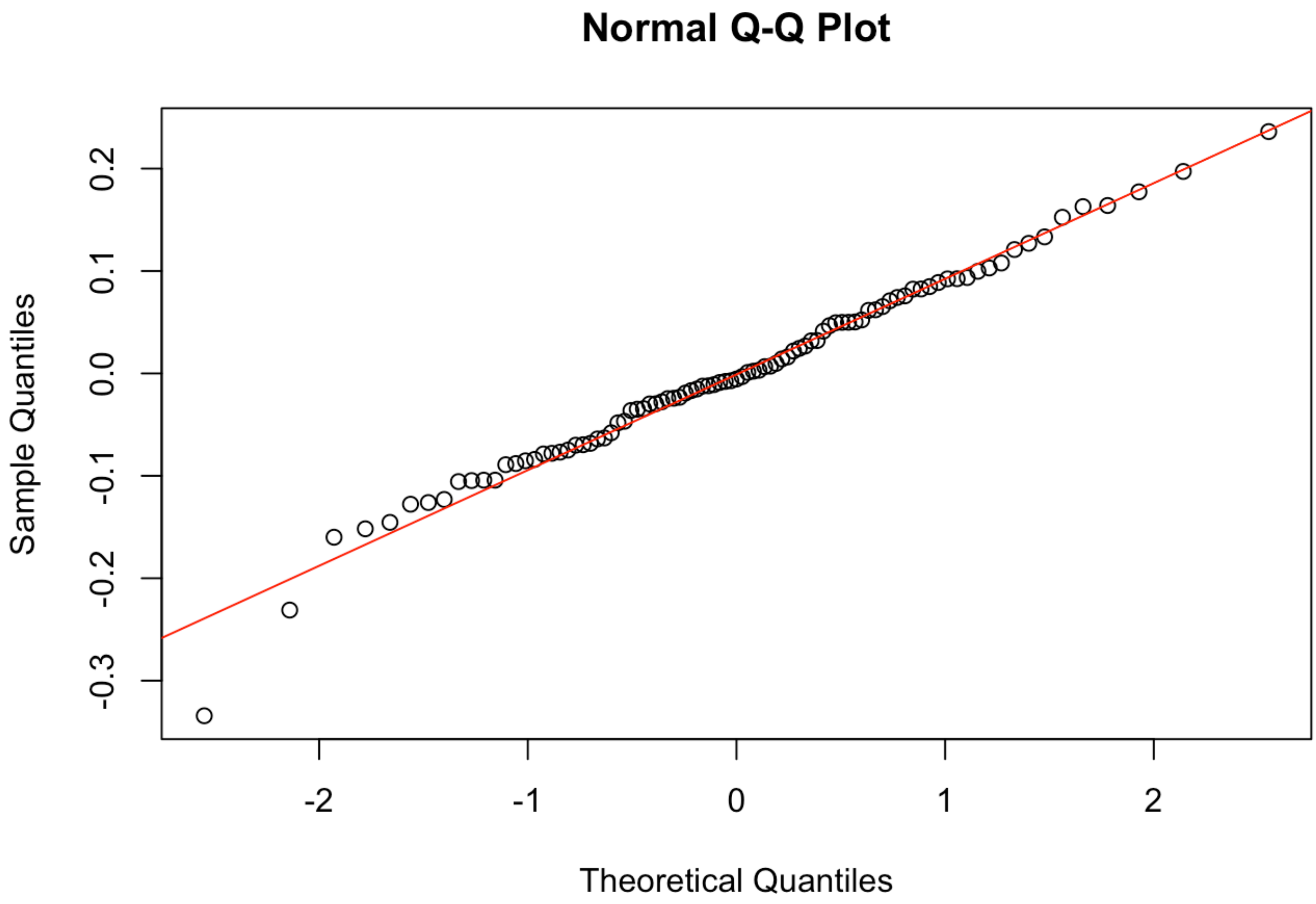
```
## Warning: Unknown or uninitialised column: 'year'.
```



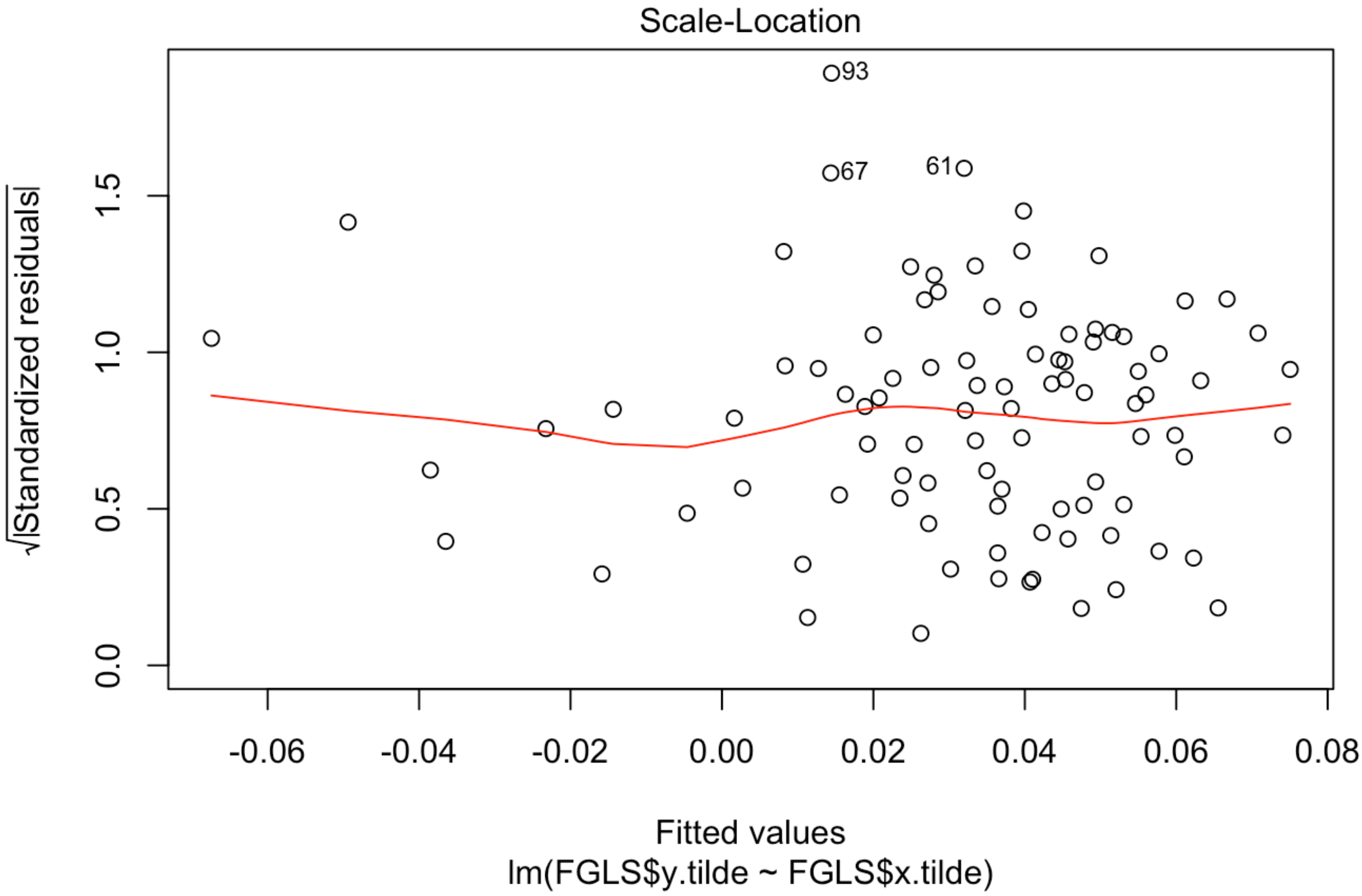

```
##
## Call:
## lm(formula = FGLS$y.tilde ~ FGLS$x.tilde)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.33429 -0.06396 -0.00546  0.06210  0.23613
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.028475   0.009877   2.883  0.00491 **
## FGLS$x.tilde -0.244530   0.092014  -2.658  0.00930 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09417 on 91 degrees of freedom
## Multiple R-squared:  0.07202,    Adjusted R-squared:  0.06182
## F-statistic: 7.063 on 1 and 91 DF,  p-value: 0.009297
```

Using FGLS (Pairs-Winsten), the coefficient on log(Debt/GDP) is -0.245 and signficant, suggesting that a 1% increase in bond supply leads to a 0.25% decrease in interest rates. However, R-squared is very low with a value of = 0.072

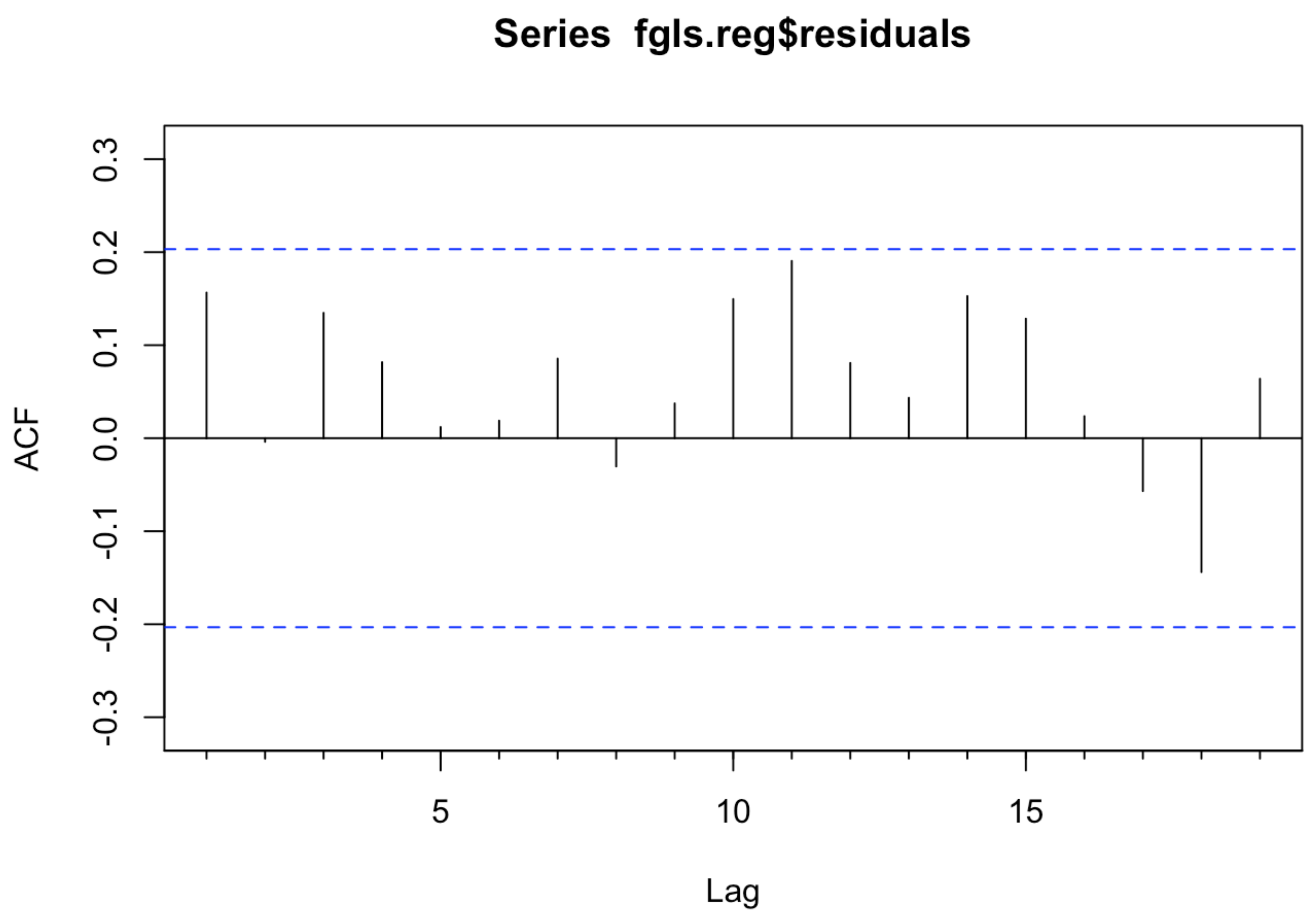
2. Diagnostics of FBLS

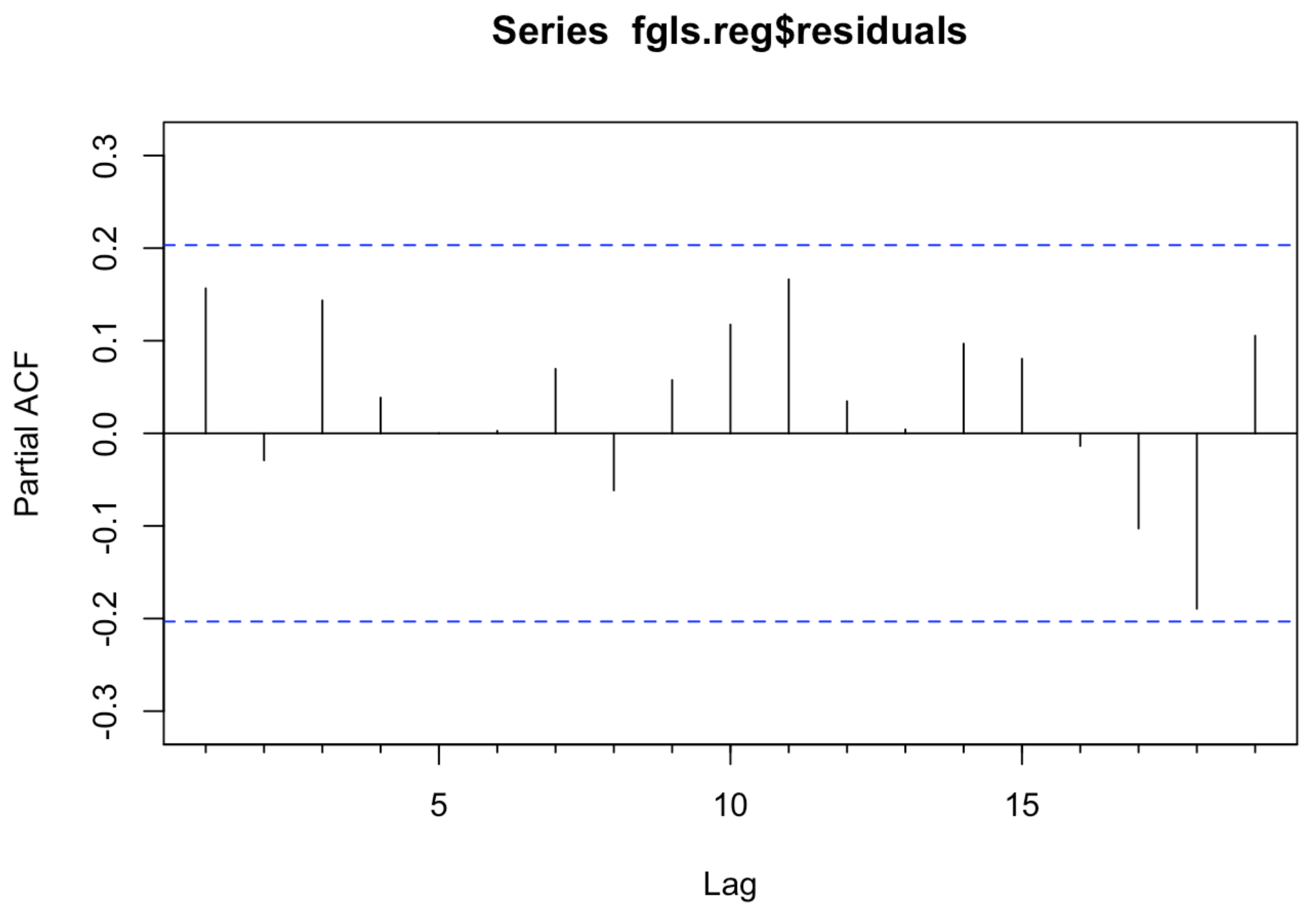


Normality of residuals improves slightly



Homoskedasticity looks good for larger fitted values.





Autocorrelation has been accounted for

Using OLS and Newey-West Heteroskedastic and Autocorrelation Consistent (HAC) standard errors

####Method: - run OLS on Log(treasury_yield) ~ log(Debt/GDP) - compute the correction factor that adjusts for serially correlated errors using the estimated autocorrelation coefficient and a chosen truncation parameter - derive the Newey-West variance estimator using correction factor and use it to compute the HAC standard error for the estimated coefficient of log(Debt/GDP)

```
##
## t test of coefficients:
##
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  1.02563    0.16758  6.1202 2.26e-08
***
## bond.data$log_bond_supply -0.52482    0.19233 -2.7288 0.007616
**
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

OLS regression with serial correlation-robust standard errors yields a significant coefficient of -0.525 for log(Debt/GDP)

Sources

http://www.sthda.com/english/articles/39-regression-model-diagnostics/161-linear-regression-assumptions-and-diagnostics-in-r-essentials/#linearity-of-the-data (http://www.sthda.com/english/articles/39-regression-model-diagnostics/161-linear-regression-assumptions-and-diagnostics-in-r-essentials/#linearity-of-the-data)

OLS assumptions: https://statisticsbyjim.com/regression/ols-linear-regression-assumptions/ (https://statisticsbyjim.com/regression/ols-linear-regression-assumptions/)

GLS and autocorrelation : https://pdfs.semanticscholar.org/7c1b/5ba81a4fcde9abe7efe49228244a98d5ce11.pdf (https://pdfs.semanticscholar.org/7c1b/5ba81a4fcde9abe7efe49228244a98d5ce11.pdf)

FBLS regression: https://www.youtube.com/watch?v=GsLacO_ZJ-s (https://www.youtube.com/watch?v=GsLacO_ZJ-s)

Newey-West HAC : https://www.econometrics-with-r.org/15-4-hac-standard-errors.html (https://www.econometrics-with-r.org/15-4-hac-standard-errors.html)