**Research Question: Impact of Inflation and Government Factors on DPI based on level of Federal Debt with Respect to GDP**

The question that is studied by this research question is whether the inflation metrics, government purchases and nominal interest rate significantly impact disposable personal income, and whether or not these impacts change based on the level of federal debt. For the sake of this research question, we consider federal debt which makes up more than 77% of the GDP as harmful as that is a cutoff suggested by economic literature. For the first part of this research question, the null hypothesis is that the inflation metrics, government purchases and nominal interest rate do not significantly impact disposable personal income or Ho: $\boldsymbol{\beta}_1 = 0$, $\boldsymbol{\beta}_2 = 0$, $\boldsymbol{\beta}_{10} = 0$, $\boldsymbol{\beta}_{11} = 0$, $\boldsymbol{\beta}_{12} = 0$, $\boldsymbol{\beta}_{13} = 0$ and the reduced model is $Y = \boldsymbol{\beta}_0$. The alternative hypothesis is that the inflation metrics, government purchases, and nominal interest rate do indeed significantly impact disposable personal income or Ha: $\boldsymbol{\beta}_1 \neq 0$ or $\boldsymbol{\beta}_2 \neq 0$ or $\boldsymbol{\beta}_{10} \neq 0$ or $\boldsymbol{\beta}_{11} \neq 0$ or $\boldsymbol{\beta}_{12} \neq 0$ or $\boldsymbol{\beta}_{13} \neq 0$ and the full model is $Y = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 X1 + \boldsymbol{\beta}_2 X2 + \boldsymbol{\beta}_{10} X10 + \boldsymbol{\beta}_{11} X11 + \boldsymbol{\beta}_{12} X12 + \boldsymbol{\beta}_{13} X13$. For the second part of this research question, the null hypothesis is that the impact the inflation metrics, government purchases and nominal interest rate have on disposable personal income does not change based on the level of federal debt or Ho: $\boldsymbol{\beta}_{0117} = 0$, $\boldsymbol{\beta}_{0217} = 0$, $\boldsymbol{\beta}_{1017} = 0$, $\boldsymbol{\beta}_{1117} = 0$, $\boldsymbol{\beta}_{1217} = 0$, $\boldsymbol{\beta}_{1317} = 0$ and the reduced model is $Y = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 X1 + \boldsymbol{\beta}_2 X2 + \boldsymbol{\beta}_{10} X10 + \boldsymbol{\beta}_{11} X11 + \boldsymbol{\beta}_{12} X12 + \boldsymbol{\beta}_{13} X13 + \boldsymbol{\beta}_{17} X17$. The alternative hypothesis is that the impact the inflation metrics, government purchases and nominal interest rate have on disposable personal income does indeed change based on the level of federal debt or Ha: $\boldsymbol{\beta}_{0117} \neq 0$ or $\boldsymbol{\beta}_{0217} \neq 0$ or $\boldsymbol{\beta}_{1017} \neq 0$ or $\boldsymbol{\beta}_{1117} \neq 0$ or $\boldsymbol{\beta}_{1217} \neq 0$ or $\boldsymbol{\beta}_{1317} \neq 0$ and the full model is $Y = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 X1 + \boldsymbol{\beta}_2 X2 + \boldsymbol{\beta}_{10} X10 + \boldsymbol{\beta}_{11} X11 + \boldsymbol{\beta}_{12} X12 + \boldsymbol{\beta}_{13} X13 + \boldsymbol{\beta}_{17} X17 + \boldsymbol{\beta}_{0117}(X1 * X17) + \boldsymbol{\beta}_{0217}(X2 * X17) + \boldsymbol{\beta}_{1017}(X10 * X17) + \boldsymbol{\beta}_{1117}(X11 * X17) + \boldsymbol{\beta}_{1217}(X12 * X17) + \boldsymbol{\beta}_{1317}(X13 * X17)$.

The independent variables of this model had some issues with multicollinearity. The CPI forecast and deflator forecast had high variance inflation factors (> 10) (figure 1.1). In addition, the added variable plots (figure 1.2) showed that CPI forecast had no effect on the model when all other variables were included in the model. The standard errors of these variables were also higher in the full model than in reduced models when one of the two variables were excluded and there were significant differences in the Anova Type I and II sums of squares. The scatterplot (figures 1.3 and 1.4) between the two variables also showed significant correlation with an R^2 value of 0.95. This multicollinearity is likely due to the fact that both the CPI forecast and deflator forecast are inflation forecasts but CPI forecast measures the change in price of a fixed basket of goods while deflator forecast measures the change in price of all economic goods. Since both metrics provide important information but are collinear, I decided to take the average of the two variables (X18) and replace the existing inflation forecasts (X10 and X11). For the sake of consistency, it is a good idea to also take the true inflation metrics (X12 and X13) and replace them with the average of the metrics (X19). The true inflation metrics also have mild multicollinearity (VIF ~ 5) (figure 1.1) and the scatterplot (figure 1.5) between these two variables also shows some correlation with an R^2 value of 0.77. Thus, it is a good idea to take the average of the true CPI and the true GDP Deflator. After these steps, all significant multicollinearity between independent variables is handled. In terms of outliers, using the bonferroni procedure for studentized deleted residuals, I found that there were no Y outliers, and using the hat matrix (figure 1.10) I found that there were 19 potential terms with high leverage X values. These high leverage values do not seem too extreme as they are very close to the cutoff. In terms of influential points, there are three points with high influence on a single fitted value based on the DFFits metric but these points are not too influential, as they are just slightly higher than the cutoff of 1 for a medium sized dataset. There are also no points with high influence on all fitted values and the regression coefficients as there are no points where the cook's distance (figure 1.10 and 1.11) or DFBetas metrics (1.12) exceed the specified cutoffs. Therefore, there do not seem to be any issues with outliers or influential points in this model. In terms of homoscedasticity, this assumption is violated as the residual plot (figure 1.6) shows non constant variance and the Breusch Pagan test (figure 1.7) returns a statistically significant result (p-value < 0.05). The error terms of the model also do not seem to be normally distributed as the QQplot (figure 1.8) shows that the actual quantiles deviate significantly from the theoretical quantiles and the shapiro test (figure 1.9) returns a statistically significant result (p-value < 0.05).

To remedy the diagnostic assumption violations of heteroscedasticity and non normality, I first started by building a weighted dataset using weighted least squares. This method resolved the issue of heteroscedasticity (figure 1.13 and 1.14); however, there were still violations of normality. To resolve the issue of non normality, I used repeated bootstrapping (figure 1.15) to build an empirical sampling distribution (figure 1.16) of the parameters in order to generate confidence intervals for inference. I used the bonferroni method with a family confidence level of 90% to generate confidence intervals for each of the nine parameters of interest, and for the intercept term. The bonferroni confidence intervals are as follows: Government Purchases (4.22, 5.06), Nominal Interest Rate (-84.46, 254.09), Debt Level (-30834.58, -4744.46), Inflation Forecast (-726.71, -57.14), True Inflation (-221.11, 112.07), Purchases-Debt Interaction (0.45, 8.20), Nominal Interest-Debt Interaction (-877.27, 520.58), Inflation-Debt Interaction (-36.75, 508.51), Forecasted Inflation-Debt Interaction (40.93, 5442.22). The only significant parameters are Government Purchases, Inflation Forecast, Level of Debt, Purchases-Debt interaction and Forecasted Inflation-Debt Interaction. Based on these confidence intervals, with 90% confidence, we can conclude that for both levels of debt, a 1 billion dollar increase in government purchases results in an increase in DPI between 4.22 and 5.06 billion dollars and a 1 percentage increase in forecasted inflation results in a decrease in DPI between 57.14 and 726.71 billion dollars. When the debt reaches a harmful level, the DPI decreases by between 4744.46 and 30834.58 billion dollars; however, every 1 billion dollar increase in government purchases reduces this reduction by between 0.45 and 8.20 billion dollars and every 1 percent increase in average forecasted inflation reduces this reduction by between 40.93 and 5442.22 billion dollars (an increase of one percentage in average forecasted inflation will also cause DPI to decrease by between 57.14 and 726.71 as discussed previously).

In conclusion, we can reject the null hypothesis as government purchases and average forecasted inflation do significantly impact DPI and the impact that these variables have on DPI does change based on whether or not there is a harmful level of federal debt. The variables included in the model explain 98.3% of variance in disposable personal income based on K Fold Cross Validation (figure 1.17) results, meaning that the variables have quite a bit of predictive power when it comes to determining disposable personal income. This experiment leads to the conclusion that government spending and purchases can increase disposable personal income; however, this spending should be managed as to not create a harmful level of federal debt which can cause severe reductions in disposable personal income. This experiment also allows for the conclusion that forecasted inflation is correlated with reductions in disposable personal income; however, this pattern could change if there is a harmful level of federal debt.

# Appendix

```
full_data.Government.Purchases     full_data.nominal.interest     full_data.Deflator_Forecast          full_data.CPI_Forecast                   full_data.CPI
                      5.928007                       5.548152                       22.460459                       27.505966                        4.825690
            full_data.Deflator                   harmful_debt
                      5.173921                       2.832942
```

Figure 1.1

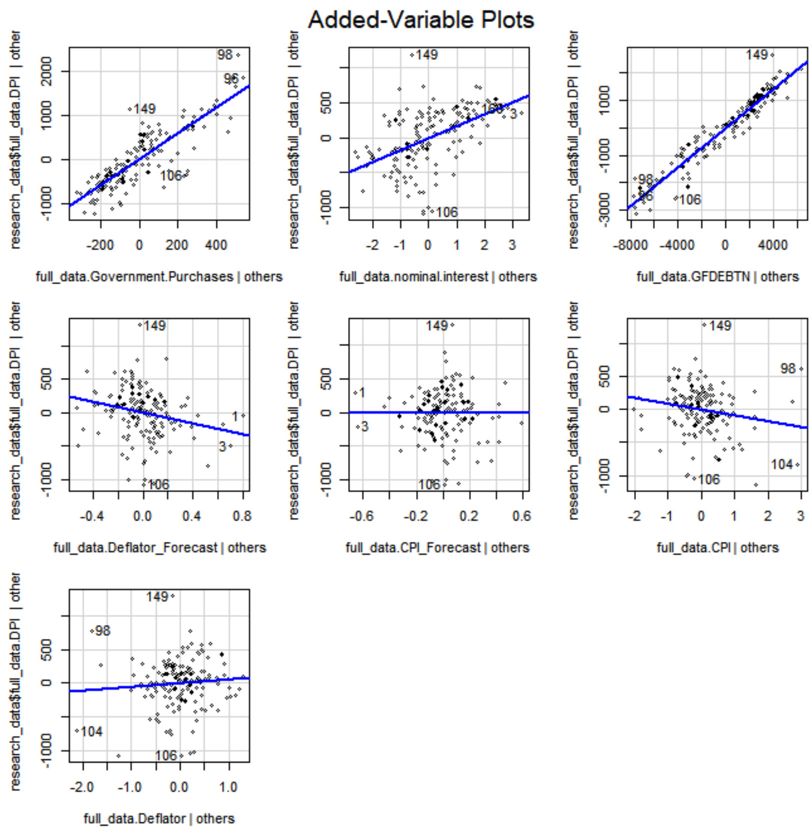



Figure 1.2

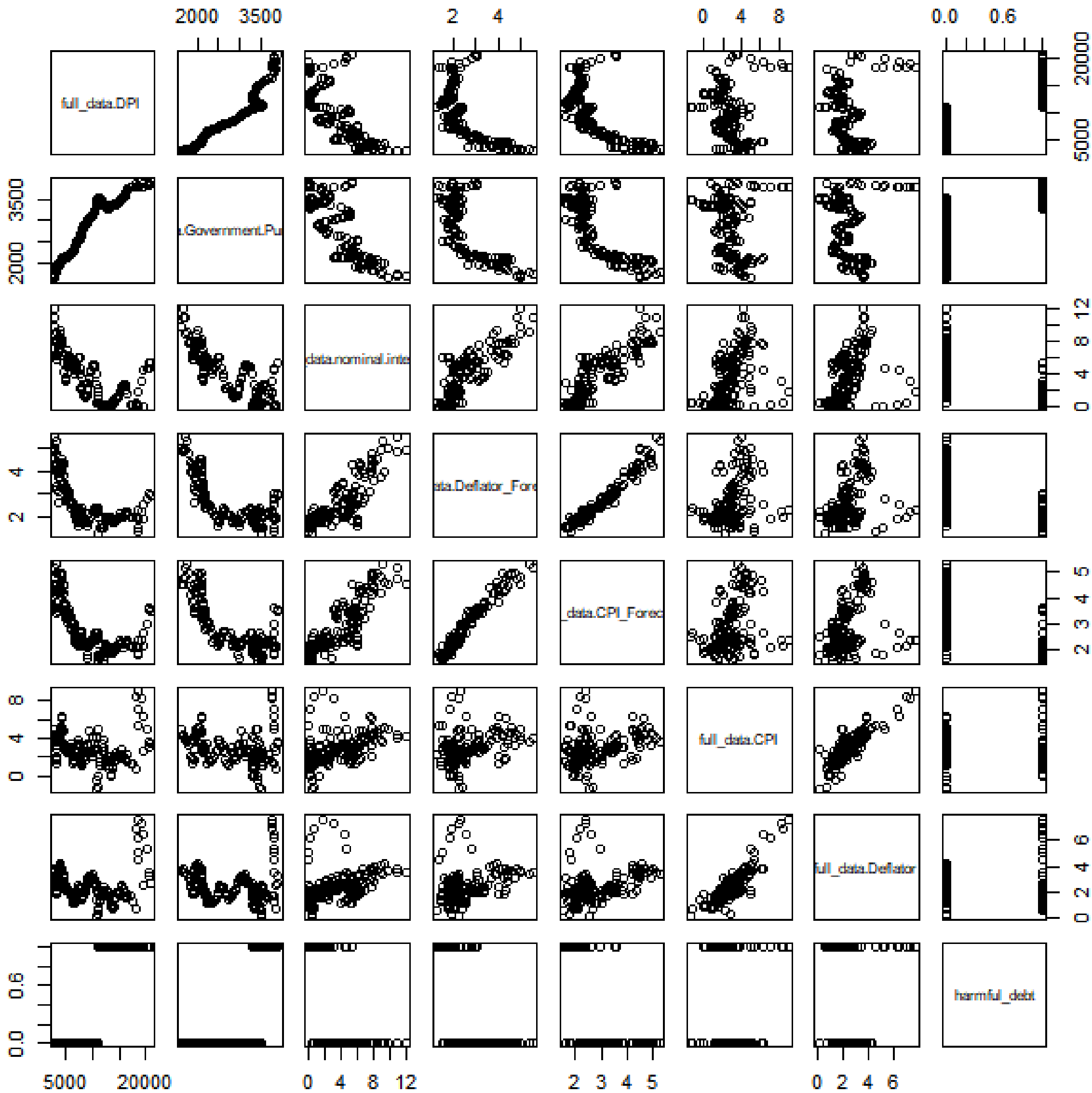
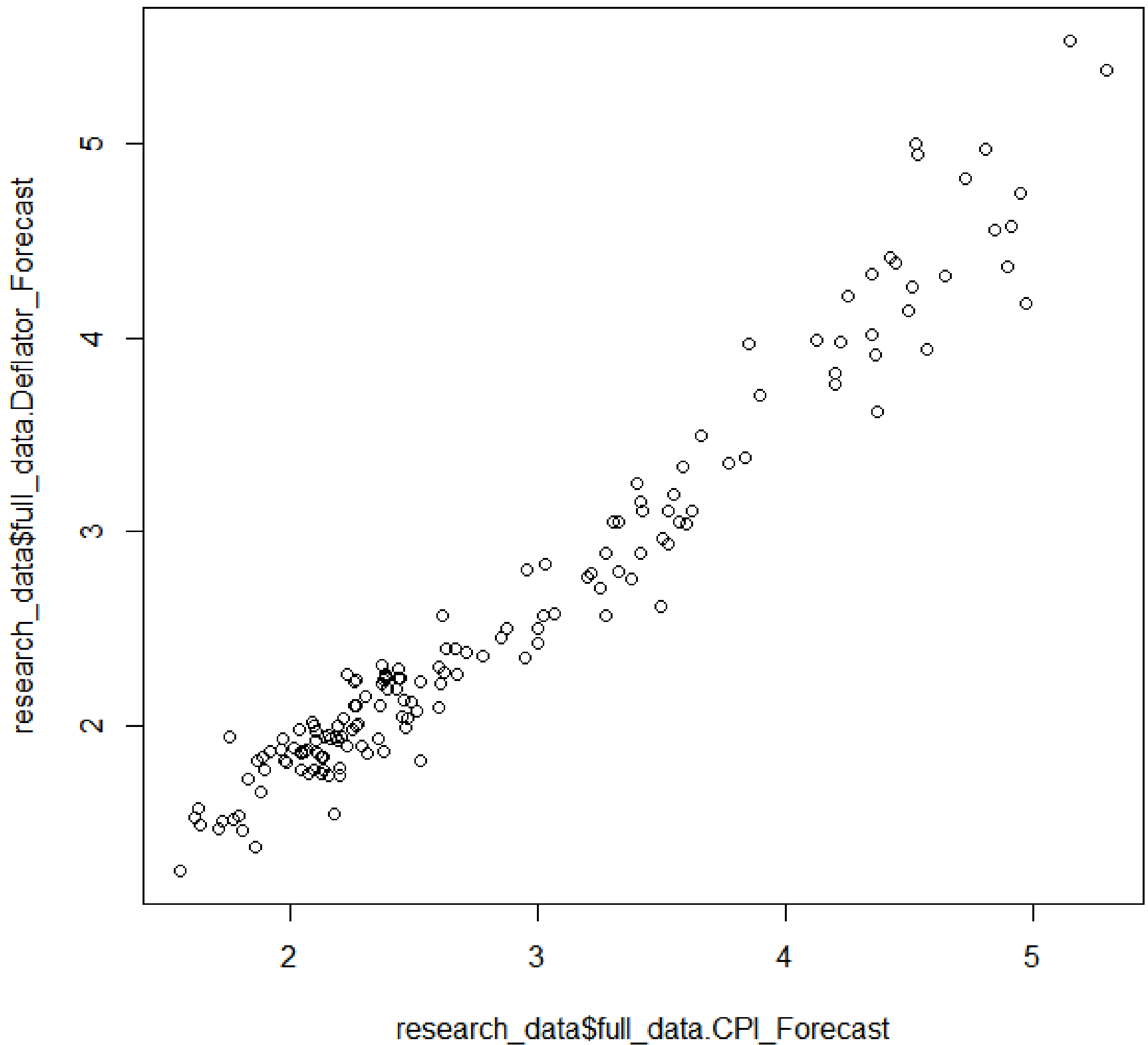
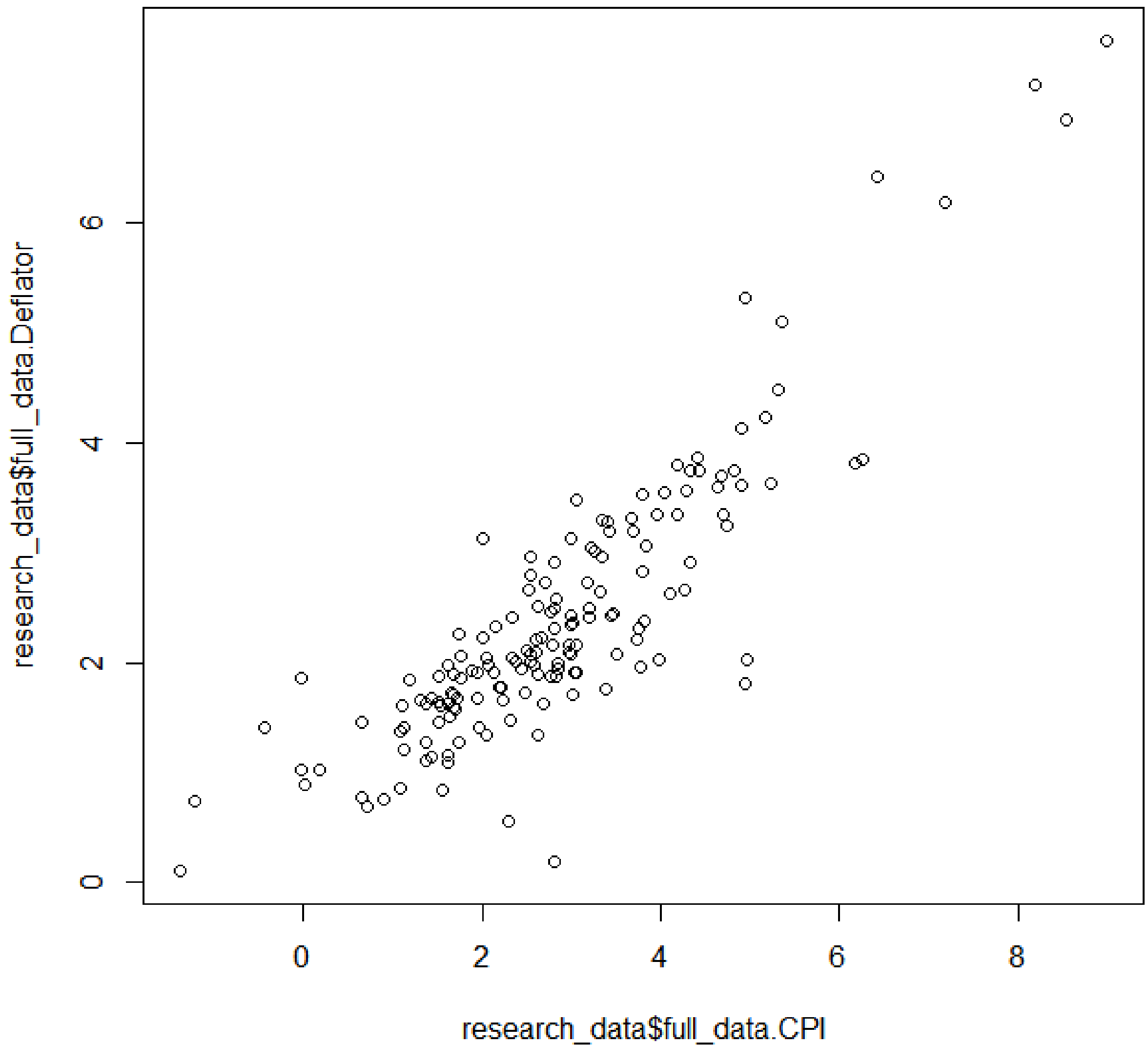

full_data.DPI
.Government.Pu
data.nominal.inte
ata.Deflator_Fore
_data.CPI_Forec
full_data.CPI
full_data.Deflator
harmful_debt

Figure 1.3



research_data$full_data.Deflator_Forecast
research_data$full_data.CPI_Forecast
2
3
4
5


Figure 1.4



research_data$full_data.Deflator
research_data$full_data.CPI
0
2
4
6
8


Figure 1.5

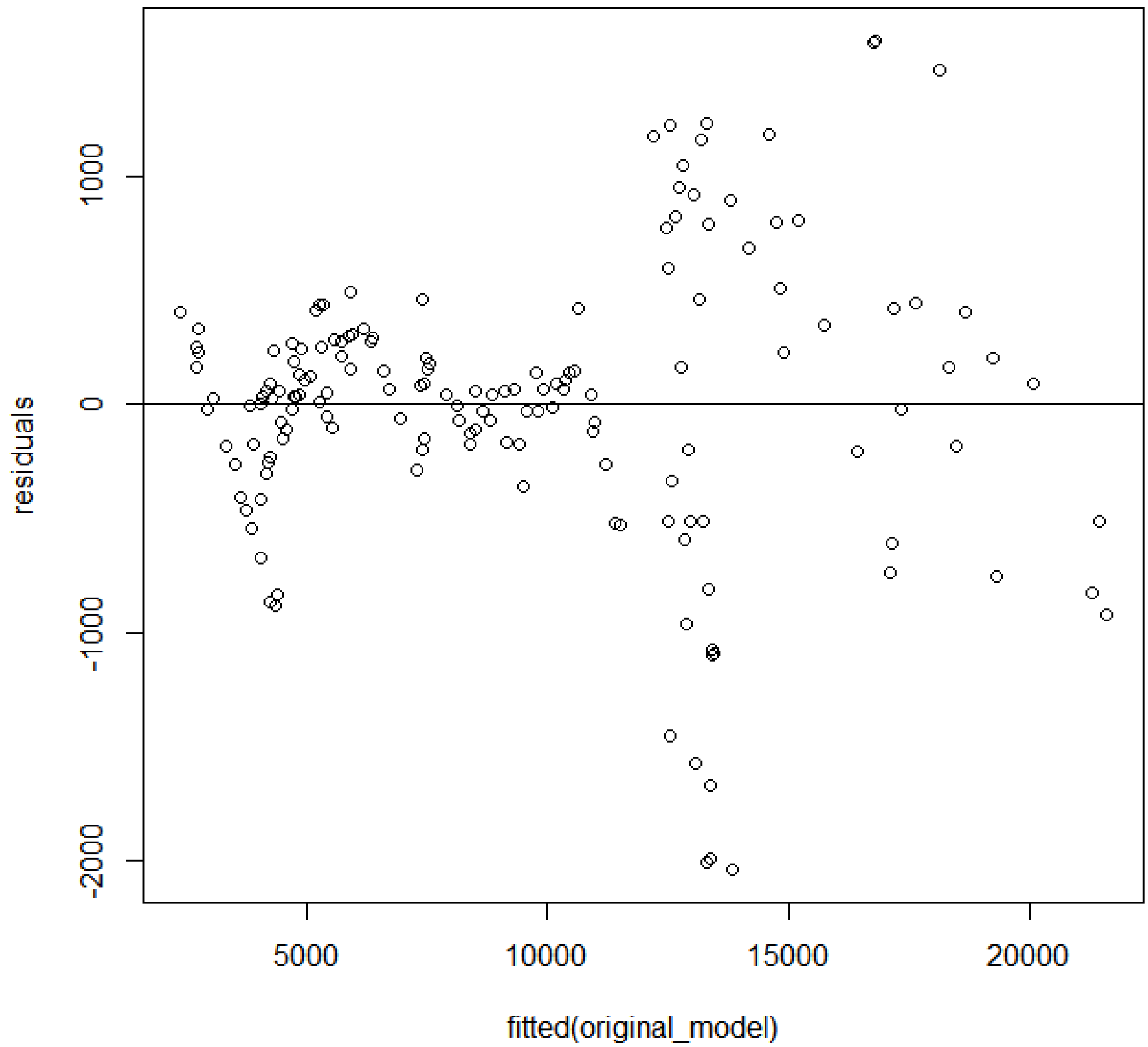



Figure 1.6

```
        studentized Breusch-Pagan test

data:  original_model
BP = 80.217, df = 9, p-value = 1.463e-13
```

Figure 1.7

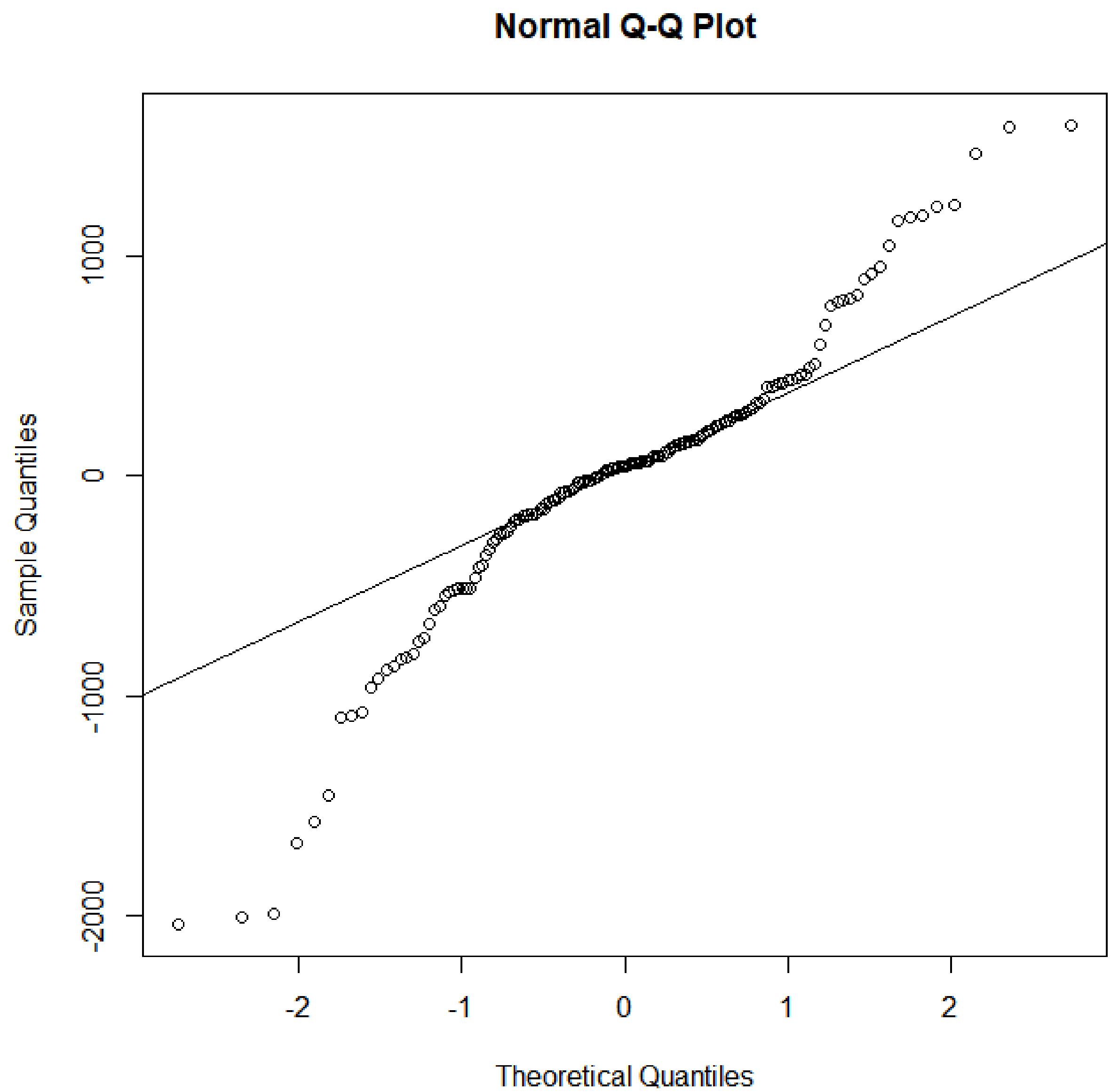



Figure 1.8

```
        Shapiro-Wilk normality test

data:  residuals
W = 0.94042, p-value = 2.944e-06
```

Figure 1.9

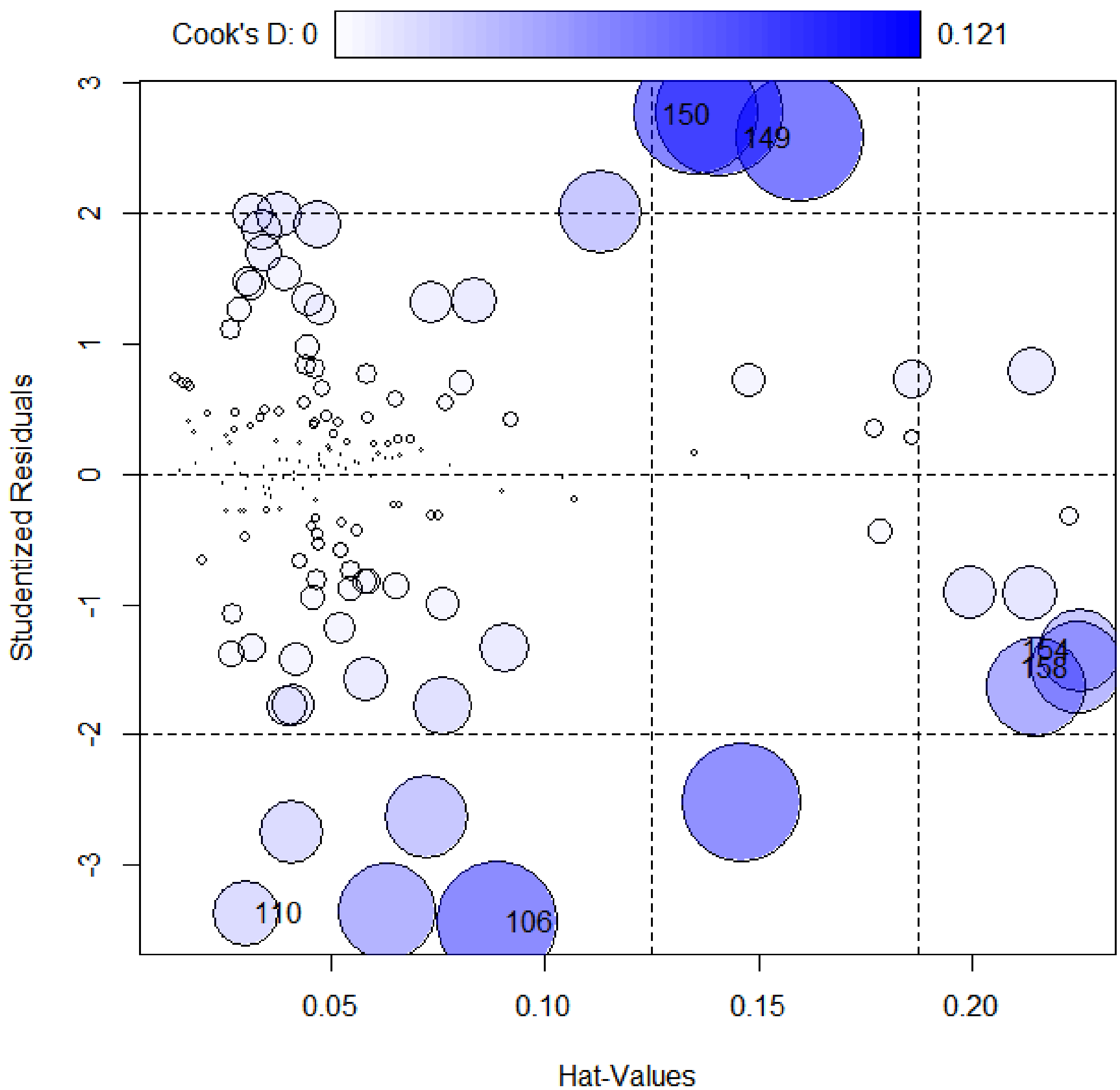



Figure 1.10

```
       StudRes        Hat      CookD
106 -3.425746 0.08887678 0.10683176
110 -3.365196 0.03003447 0.03280768
149  2.570114 0.15967178 0.12099023
150  2.756004 0.14095524 0.11938151
154 -1.346464 0.22514857 0.05239533
158 -1.481696 0.22484960 0.06317962
> qf(0.5,10,150)
[1] 0.9383724
```

Figure 1.11

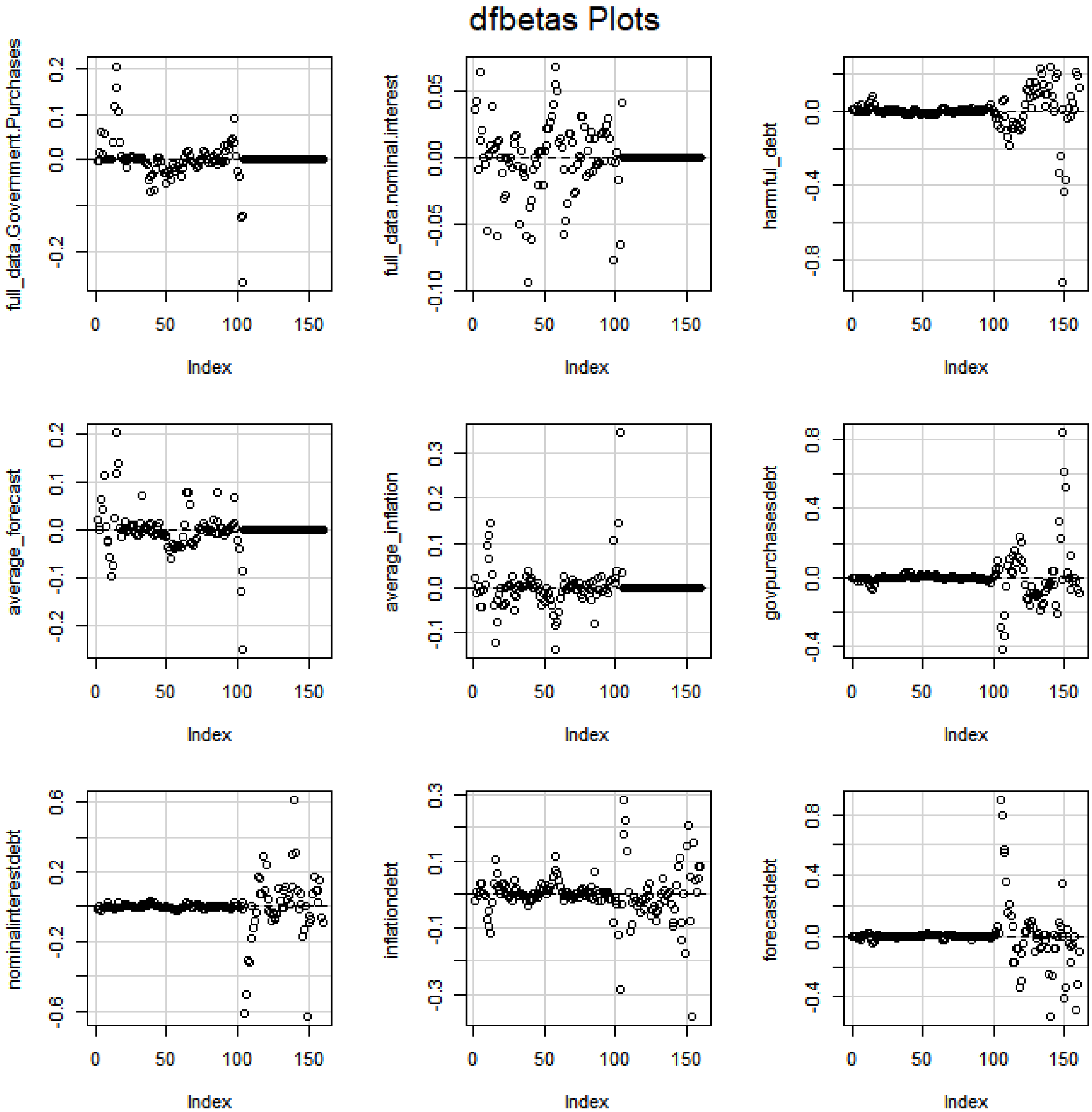



Figure 1.12

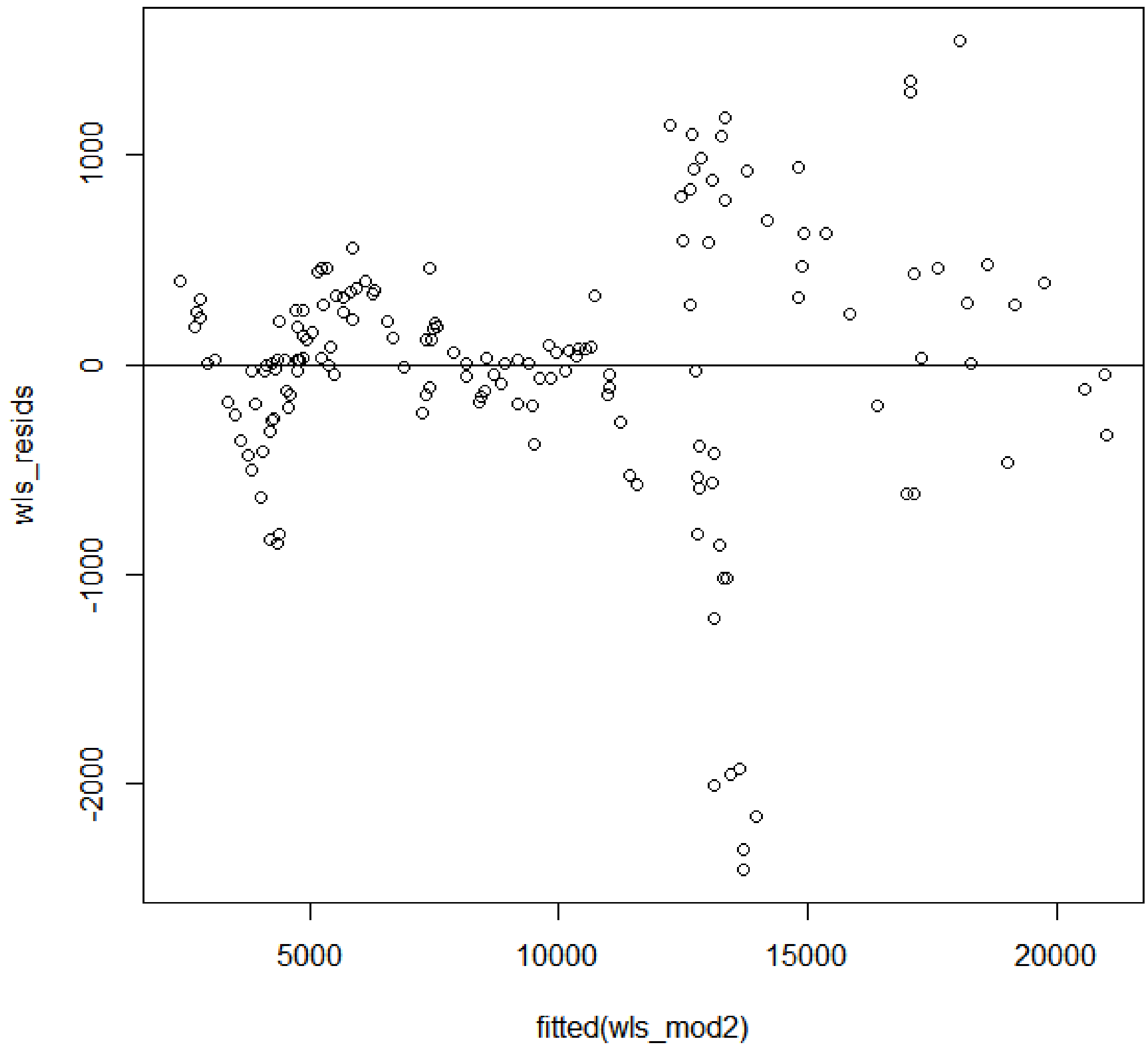
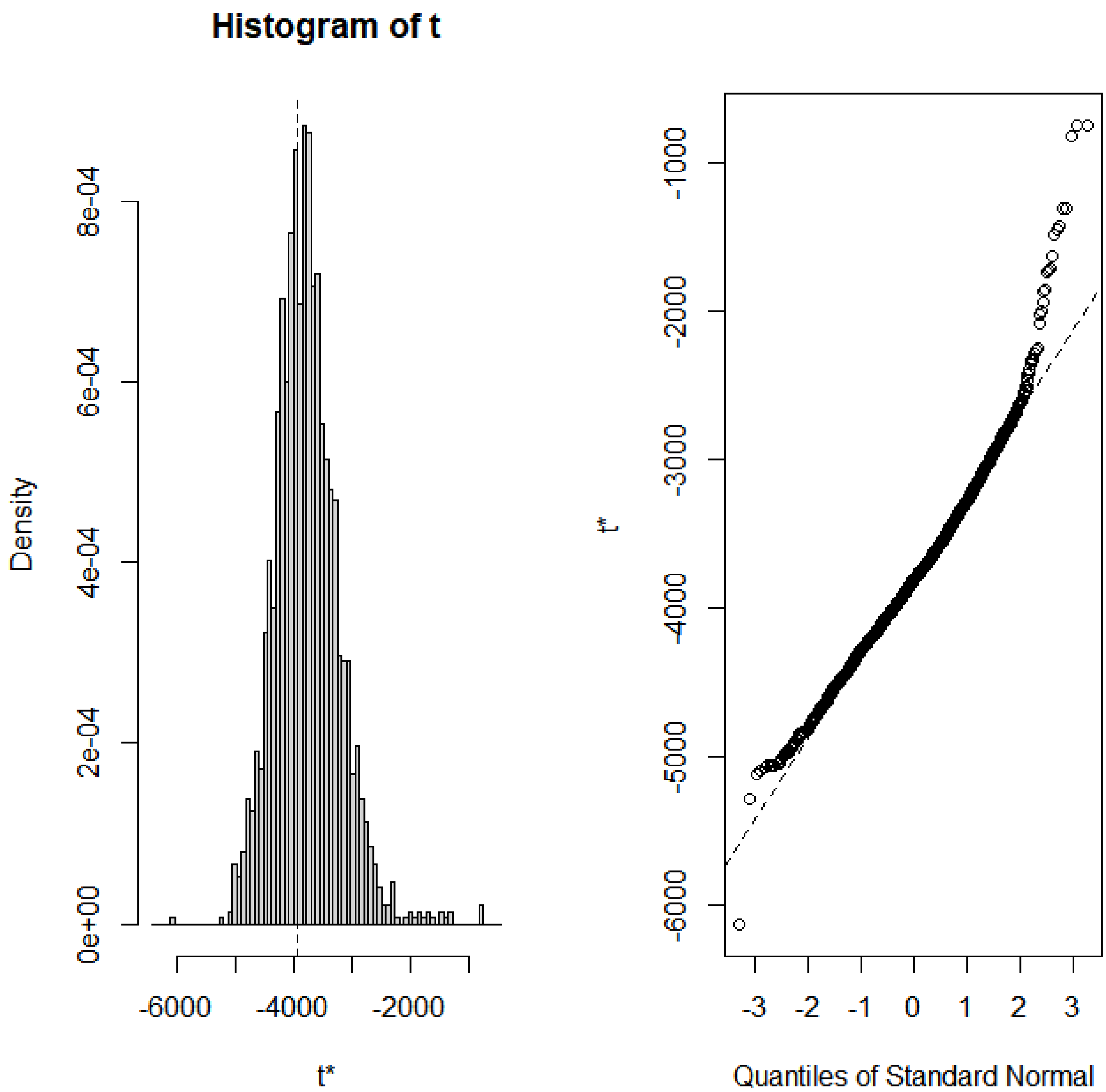



Figure 1.13

```
        studentized Breusch-Pagan test

data:  wls_mod2
BP = 0.076704, df = 9, p-value = 1
```

Figure 1.14



Histogram of t
Density
0e+00
2e-04
4e-04
6e-04
8e-04
-6000
-4000
-2000
t*
-6000
-5000
-4000
-3000
-2000
-1000
t*
-3
-2
-1
0
1
2
3
Quantiles of Standard Normal


Figure 1.15

```
Call:
boot(data = research_data, statistic = boot.wlscoef, R = 2000,
    maxit = 100)


Bootstrap Statistics :
          original        bias     std. error
t1*   -3935.171198  166.63310094  532.7176622
t2*       4.642768   -0.03703762    0.1608229
t3*      84.816438    9.92430820   64.8826319
t4*  -17789.521604 1158.82450725 5000.0192132
t5*    -391.922657  -14.96687952  128.3201291
t6*     -54.518588  -25.84574046   63.8511796
t7*       4.326737   -0.49424707    1.4857999
t8*    -178.345001    3.52480001  267.8882155
t9*     235.882882   41.08158309  104.4954110
t10*   2741.574519  248.20411032 1035.1240937
```

Figure 1.16

```
Linear Regression

160 samples
  9 predictor

No pre-processing
Resampling: Cross-Validated (20 fold)
Summary of sample sizes: 152, 152, 152, 152, 152, 152, ...
Resampling results:

  RMSE      Rsquared   MAE
  653.8214  0.9832874  460.1555

Tuning parameter 'intercept' was held constant at a value of TRUE
```

Figure 1.17