EM4

YiTao Hu, Charles Rambo, Junyu(Kevin) Wu, Jin(Jane) Huangfu

31/01/2020

Probelm 1

1

Using CRSP, get monthly market returns ex and cum dividends, as well as the monthly t-bill rate, from 1963 to 2018. Create the market dividend yield by summing the div- idends over the last 12 months and divide by current price (you can do this using information extracted using the ex- and cum-dividend returns). Construct excess re- turns by subtracting the log of the 1-month gross t-bill rate from the 1-month gross cum-dividends returns. Note: to get to gross returns you may have to add 1 to the original data series. From the St. Louis Fed data page (FRED; https://fred.stlouisfed.org/), get monthly data on the term and default spreads for the same sample. For the former, use the "10- Year Treasury Constant Maturity Minus Federal Funds Rate," for the latter subtract "Moodyís Seasoned Aaa Corporate Bond Minus Federal Funds Rate" from "Moodyís Seasoned Baa Corporate Bond Minus Federal Funds Rate."

```
# import data and Library
library(sandwich)
```

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

```
library(readr)
library(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
```

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.5.2
```

```
library(DataAnalytics)
library(lmtest)
```

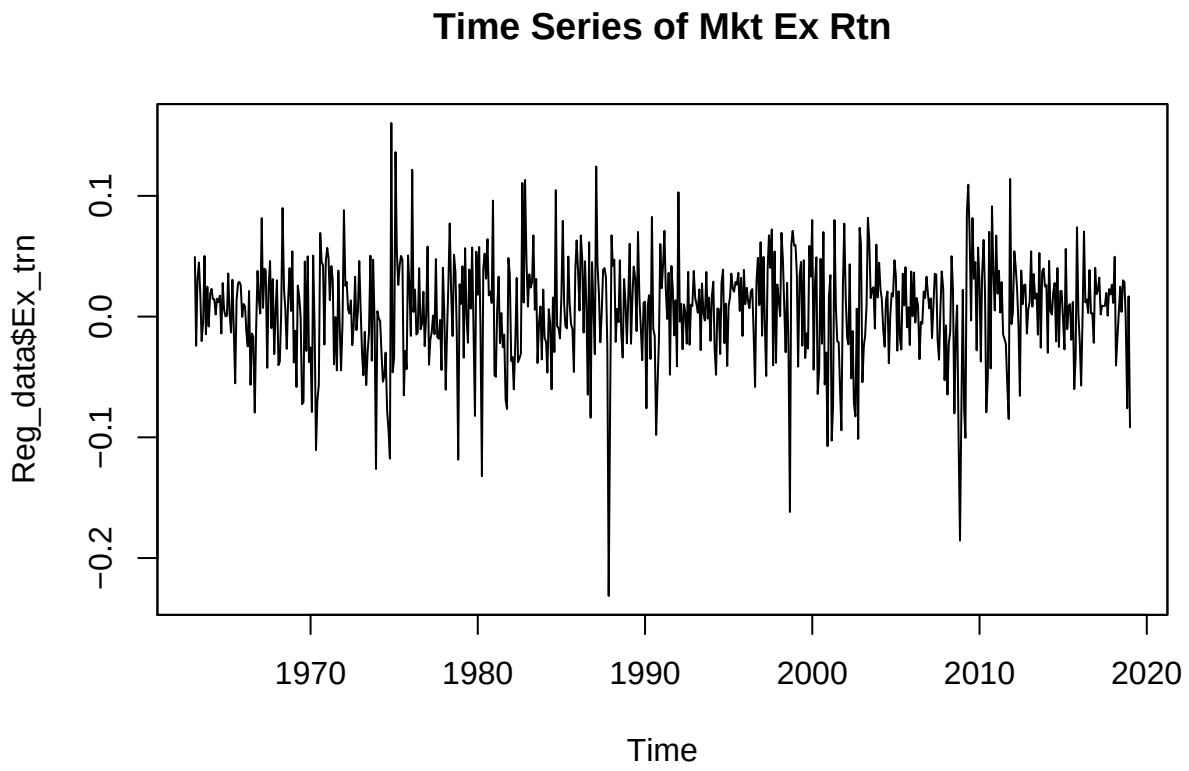
```
## Warning: package 'lmtest' was built under R version 3.5.2
```

```
HW4data <- read_csv("HW4data.csv",
     col_types = cols(DATE = col_date(format = "%Y%m%d")))

Mkt <- read_csv("Mkt.csv",
     col_types = cols(DATE = col_date(format = "%Y%m%d")))

#turn percentages in to dicimals
HW4data[,5:7]=HW4data[,5:7]/100
#compute the monthly div yield
Mkt$mondiv=(Mkt$vwretd-Mkt$vwretx)*Mkt$totval
#compute rolling annul dividend
```

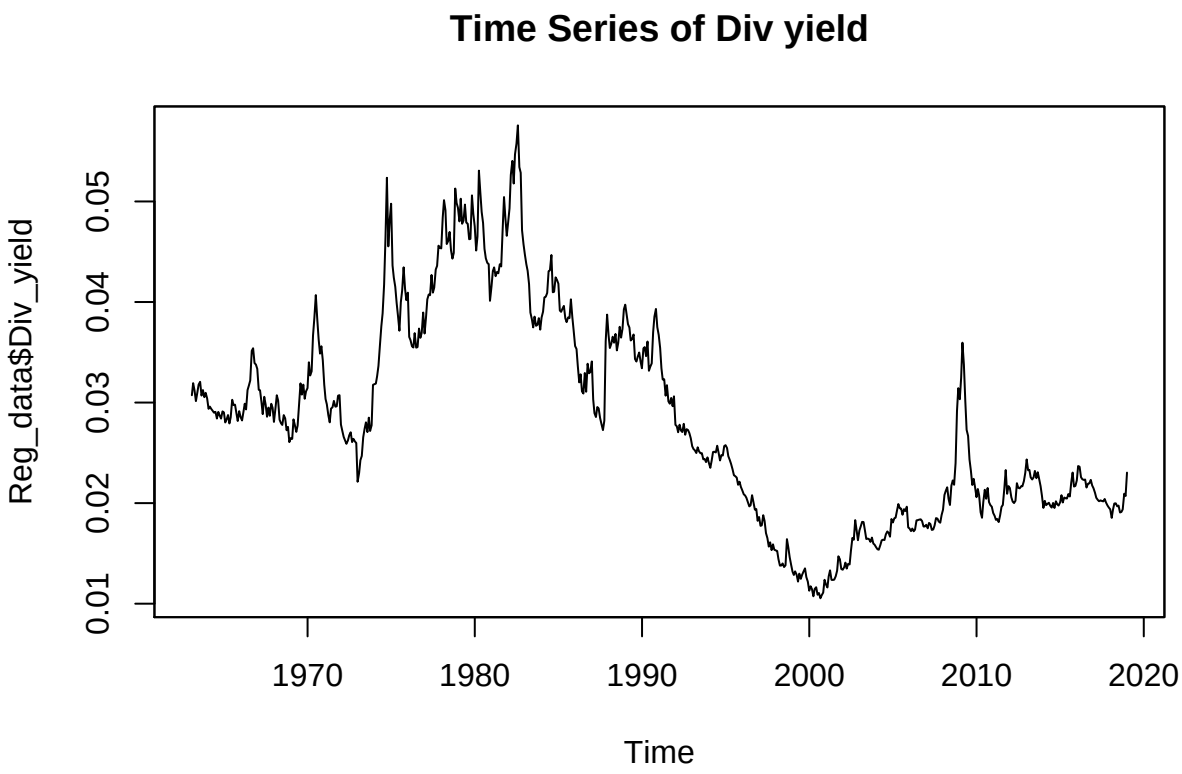
```
HW4data$annual_roll_div=rollapply(data = Mkt$mondiv,FUN=sum,width=12)[-1]
HW4data$Mktval=Mkt$totval[13:nrow(Mkt)]
#compute annaul div yield
HW4data$div_yield=HW4data$annual_roll_div/HW4data$Mktval
#compute monthly excess return
HW4data$exRtn=HW4data$vwretd-log(1+HW4data$t30ret)
#compute monthly default spread
HW4data$dfsp=HW4data$BAAFFM-HW4data$AAAFFM
#convert the data into a zoo obj
HW4data=zoo(HW4data[,-1],order.by = HW4data$DATE)
```

2 Plot your data

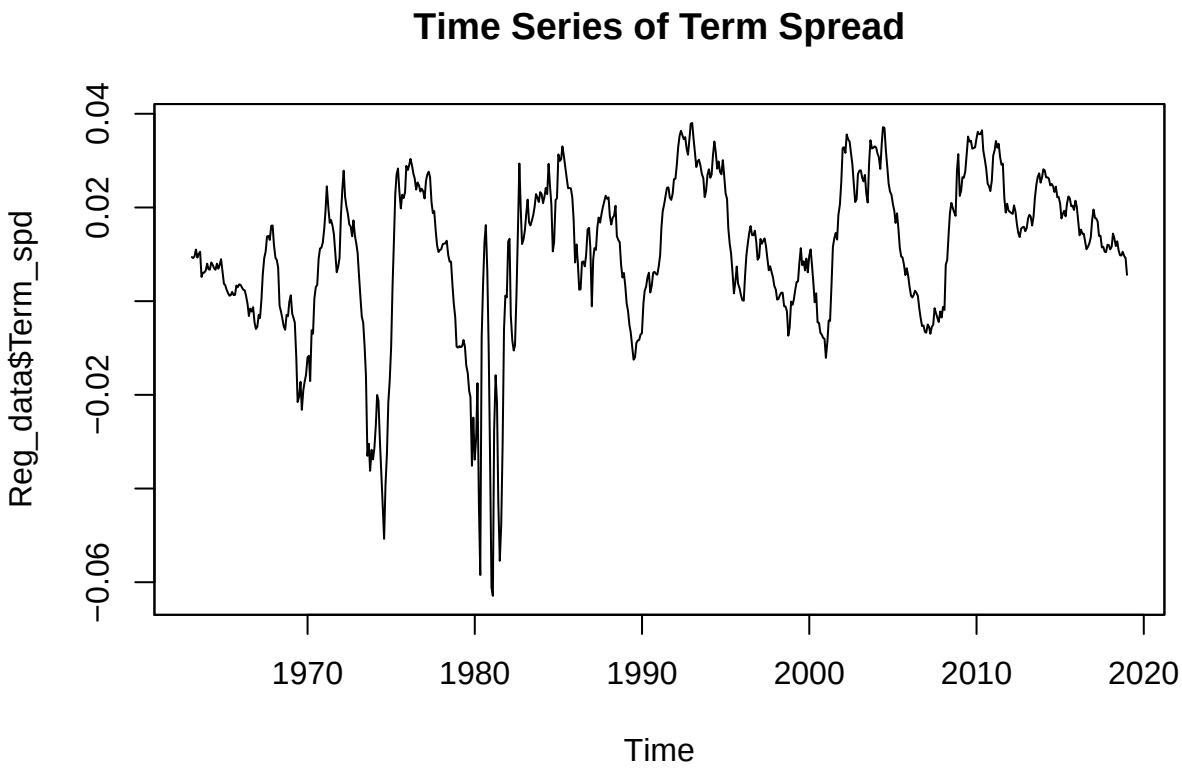
```
Reg_data=HW4data[,c(10,9,4,11)]
colnames(Reg_data)=c('Ex_trn','Div_yield','Term_spd','Creadit_spd')
plot.zoo (Reg_data$Ex_trn,
          xlab = 'Time',
          main='Time Series of Mkt Ex Rtn')
```




```
plot.zoo (Reg_data$Div_yield,
          xlab = 'Time',
          main='Time Series of Div yield',type='l')
```

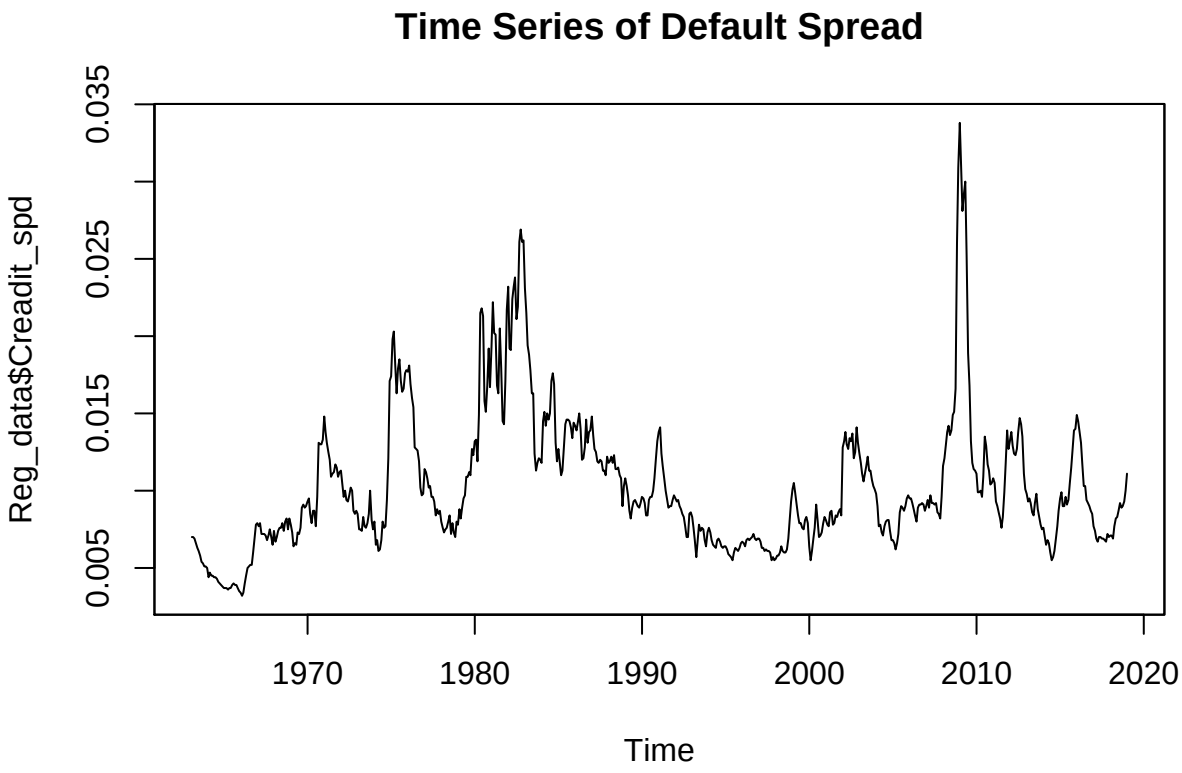



```
plot.zoo (Reg_data$Term_spd,
         xlab = 'Time',
         main='Time Series of Term Spread',type='l')
```




```
plot.zoo (Reg_data$Creadit_spd,
         xlab = 'Time',
```

```
        main='Time Series of Default Spread',type='l')
```



3.

Using your three predictive variables (the lagged dividend yield, term spread, and default spread), forecast excess equity returns at the 1-month, 3-month, 12-month, 24-month, and 60-month horizons. Report your results from each of these regressions (regression coeffcients, standard errors, and R2s). The underlying data is monthly, so make sure to explain your choice of standard errors. For monthly Regression, the model would be:

$$R^E_{t+1} = a + b^{(1)}_{div} DP_t + b^{(1)}_{term} TermSpd_t + b^{(1)}_{credit} CreditSpd_t + e_t$$

```
one_month_model=lm(formula = Reg_data$Ex_trn[-1]~Reg_data$Div_yield[-nrow(Reg_data)]+Reg_data$Term_spd[
summary(one_month_model)
```

```
##
## Call:
## lm(formula = Reg_data$Ex_trn[-1] ~ Reg_data$Div_yield[-nrow(Reg_data)] +
##     Reg_data$Term_spd[-nrow(Reg_data)] + Reg_data$Creadit_spd[-nrow(Reg_data)])
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.240110 -0.025929  0.003685  0.027414  0.161019
##
## Coefficients:
##                                      Estimate Std. Error t value
## (Intercept)                         -0.009868   0.005664  -1.742
## Reg_data$Div_yield[-nrow(Reg_data)]  0.380434   0.194085   1.960
## Reg_data$Term_spd[-nrow(Reg_data)]   0.342422   0.108048   3.169
## Reg_data$Creadit_spd[-nrow(Reg_data)] 0.049361   0.427264   0.116
```

```
##                                      Pr(>|t|)
## (Intercept)                            0.0819 .
## Reg_data$Div_yield[-nrow(Reg_data)]     0.0504 .
## Reg_data$Term_spd[-nrow(Reg_data)]      0.0016 **
## Reg_data$Creadit_spd[-nrow(Reg_data)]   0.9081
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04342 on 667 degrees of freedom
## Multiple R-squared:  0.01893,    Adjusted R-squared:  0.01452
## F-statistic:  4.29 on 3 and 667 DF,  p-value: 0.005195
```

Perform Box-Lijung test on residual.

```
Box.test(as.vector(one_month_model$residuals),lag = round(log(length(one_month_model$residuals))),type =
```

```
##
##  Box-Ljung test
##
## data:  as.vector(one_month_model$residuals)
## X-squared = 8.0015, df = 7, p-value = 0.3325
```

We do not perform NW standard errors, because there is no significant autocorrelation in residuals.

Define a function for horiaton regression

```
horiton_estimate=function(dataframe=Reg_data,Lag=3){
  #constract the new dataframe for new regression
  Data_adj=cbind(as.vector(rollapply(dataframe[,1],width = Lag,FUN = sum,fill = NA,align = 'right')[-1:-
  Data_adj=data.frame(Data_adj)
  colnames(Data_adj)=c('Sum_EX_rtn','Lagged_Div_yield','Lagged_Term_Spd','Lagged_Credit_Spd')

  Model=lm(formula = Data_adj$Sum_EX_rtn~Data_adj$Lagged_Div_yield+Data_adj$Lagged_Term_Spd+Data_adj$Lag
lmSumm(Model)
print(Box.test(as.vector(Model$residuals),lag = round(log(length(Model$residuals))),type = 'Ljung-Box'))
print('The NW standard error would be:')
coeftest(x=Model,vcov. = NeweyWest(x = Model,lag = (Lag-1)*1.5,adjust = T))
return (Model)
}
```

For three month model:

$$\sum_{i=1}^{3} R_i^E = a + b_{div}^{(3)} DP_t + b_{term}^{(3)} TermSpd_t + b_{credit}^{(3)} CreditSpd_t + e_{t+3}$$

```
horiton_estimate(dataframe = Reg_data,Lag = 3)
```

```
## Multiple Regression Analysis:
##     4 regressors(including intercept) and 669 observations
##
## lm(formula = Data_adj$Sum_EX_rtn ~ Data_adj$Lagged_Div_yield +
##     Data_adj$Lagged_Term_Spd + Data_adj$Lagged_Credit_Spd)
##
## Coefficients:
##                             Estimate Std Error t value p value
## (Intercept)                -0.02786   0.01005   -2.77   0.006
## Data_adj$Lagged_Div_yield   1.08200   0.34430    3.14   0.002
```

```
## Data_adj$Lagged_Term_Spd    0.91870   0.19150    4.80   0.000
## Data_adj$Lagged_Credit_Spd  0.28090   0.75730    0.37   0.711
## ---
## Standard Error of the Regression:  0.07695
## Multiple R-squared:  0.045  Adjusted R-squared:  0.041
## Overall F stat: 10.5 on 3 and 665 DF, pvalue= 0
##
##  Box-Ljung test
##
## data:  as.vector(Model$residuals)
## X-squared = 393.59, df = 7, p-value < 2.2e-16
##
## [1] "The NW standard error would be:"
```

```
##
## Call:
## lm(formula = Data_adj$Sum_EX_rtn ~ Data_adj$Lagged_Div_yield +
##     Data_adj$Lagged_Term_Spd + Data_adj$Lagged_Credit_Spd)
##
## Coefficients:
##                (Intercept)   Data_adj$Lagged_Div_yield
##                   -0.02786                     1.08231
##   Data_adj$Lagged_Term_Spd  Data_adj$Lagged_Credit_Spd
##                    0.91875                     0.28090
```

For 12 month model:

$$\sum_{i=1}^{12} R_i^E = a + b_{div}^{(12)} DP_t + b_{term}^{(12)} TermSpd_t + b_{credit}^{(12)} CreditSpd_t + e_{t+12}$$

```
Model_12=horiton_estimate(dataframe = Reg_data,Lag = 12)
```

```
## Multiple Regression Analysis:
##     4 regressors(including intercept) and 660 observations
##
## lm(formula = Data_adj$Sum_EX_rtn ~ Data_adj$Lagged_Div_yield +
##     Data_adj$Lagged_Term_Spd + Data_adj$Lagged_Credit_Spd)
##
## Coefficients:
##                             Estimate Std Error t value p value
## (Intercept)                -0.09479   0.01943   -4.88   0.000
## Data_adj$Lagged_Div_yield   3.81800   0.66290    5.76   0.000
## Data_adj$Lagged_Term_Spd    3.36500   0.36740    9.16   0.000
## Data_adj$Lagged_Credit_Spd  1.21600   1.45300    0.84   0.403
## ---
## Standard Error of the Regression:  0.1476
## Multiple R-squared:  0.147  Adjusted R-squared:  0.143
## Overall F stat: 37.8 on 3 and 656 DF, pvalue= 0
##
##  Box-Ljung test
##
## data:  as.vector(Model$residuals)
## X-squared = 1933.2, df = 6, p-value < 2.2e-16
##
## [1] "The NW standard error would be:"
```

For 24 month model:

$$\sum_{i=1}^{24} R_i^E = a + b_{div}^{(24)} DP_t + b_{term}^{(24)} TermSpd_t + b_{credit}^{(24)} CreditSpd_t + e_{t+12}$$

```
horiton_estimate(dataframe = Reg_data,Lag = 24)
```

```
## Multiple Regression Analysis:
##     4 regressors(including intercept) and 648 observations
##
## lm(formula = Data_adj$Sum_EX_rtn ~ Data_adj$Lagged_Div_yield +
##     Data_adj$Lagged_Term_Spd + Data_adj$Lagged_Credit_Spd)
##
## Coefficients:
##                             Estimate Std Error t value p value
## (Intercept)                -0.09497   0.02664   -3.56   0.000
## Data_adj$Lagged_Div_yield   5.27200   0.90110    5.85   0.000
## Data_adj$Lagged_Term_Spd    4.37300   0.49850    8.77   0.000
## Data_adj$Lagged_Credit_Spd  1.89200   1.97400    0.96   0.338
## ---
## Standard Error of the Regression:  0.2001
## Multiple R-squared:  0.145  Adjusted R-squared:  0.141
## Overall F stat: 36.35 on 3 and 644 DF, pvalue= 0
##
##  Box-Ljung test
##
## data:  as.vector(Model$residuals)
## X-squared = 2495.2, df = 6, p-value < 2.2e-16
##
## [1] "The NW standard error would be:"
##
## Call:
## lm(formula = Data_adj$Sum_EX_rtn ~ Data_adj$Lagged_Div_yield +
##     Data_adj$Lagged_Term_Spd + Data_adj$Lagged_Credit_Spd)
##
## Coefficients:
##                (Intercept)   Data_adj$Lagged_Div_yield
##                   -0.09497                     5.27203
##   Data_adj$Lagged_Term_Spd  Data_adj$Lagged_Credit_Spd
##                    4.37303                     1.89203
```

For 60 month model:

$$\sum_{i=1}^{60} R_i^E = a + b_{div}^{(60)} DP_t + b_{term}^{(60)} TermSpd_t + b_{credit}^{(60)} CreditSpd_t + e_{t+12}$$

```
horiton_estimate(Lag = 60)
```

```
## Multiple Regression Analysis:
##     4 regressors(including intercept) and 612 observations
##
## lm(formula = Data_adj$Sum_EX_rtn ~ Data_adj$Lagged_Div_yield +
##     Data_adj$Lagged_Term_Spd + Data_adj$Lagged_Credit_Spd)
##
```

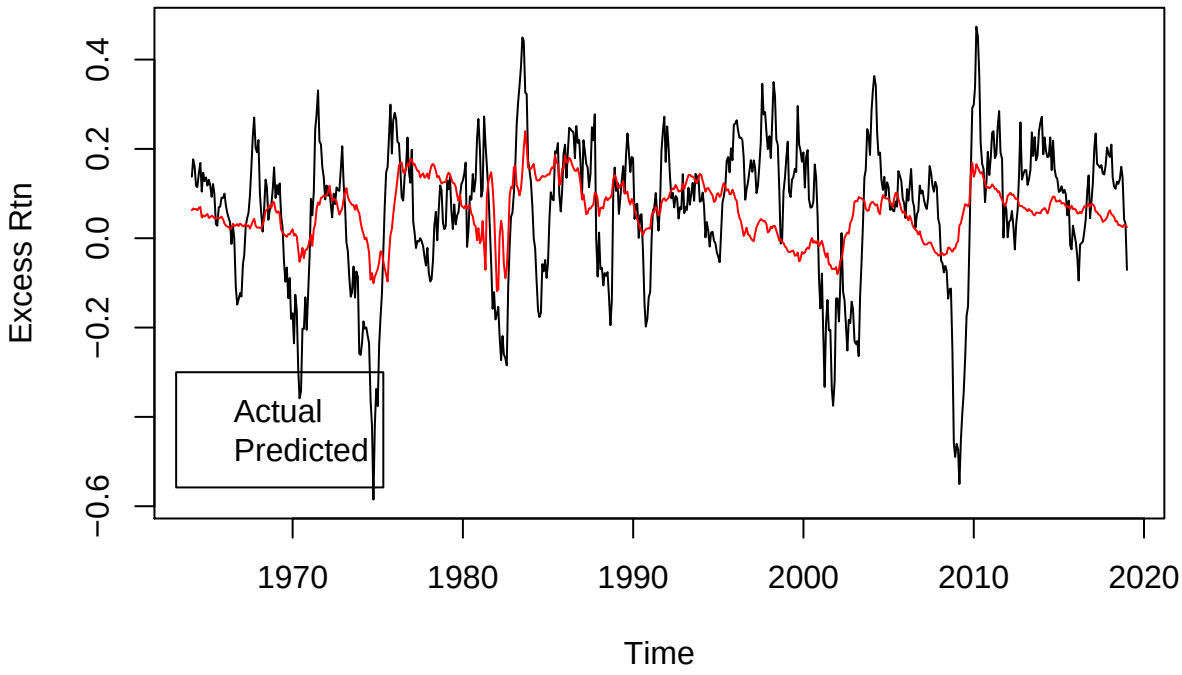
```
## Coefficients:
##                             Estimate Std Error t value p value
## (Intercept)                  -0.1183   0.03465   -3.41   0.001
## Data_adj$Lagged_Div_yield     6.1810   1.16600    5.30   0.000
## Data_adj$Lagged_Term_Spd      7.2850   0.64250   11.34   0.000
## Data_adj$Lagged_Credit_Spd   15.4800   2.55700    6.05   0.000
## ---
## Standard Error of the Regression:  0.2562
## Multiple R-squared:  0.287  Adjusted R-squared:  0.284
## Overall F stat: 81.78 on 3 and 608 DF, pvalue= 0
##
##  Box-Ljung test
##
## data:  as.vector(Model$residuals)
## X-squared = 2649.9, df = 6, p-value < 2.2e-16
##
## [1] "The NW standard error would be:"
```

```
##
## Call:
## lm(formula = Data_adj$Sum_EX_rtn ~ Data_adj$Lagged_Div_yield +
##     Data_adj$Lagged_Term_Spd + Data_adj$Lagged_Credit_Spd)
##
## Coefficients:
##                 (Intercept)   Data_adj$Lagged_Div_yield
##                     -0.1183                      6.1810
##   Data_adj$Lagged_Term_Spd  Data_adj$Lagged_Credit_Spd
##                      7.2854                     15.4773
```

4

Plot the fitted values for 12 months

```
plot(y=rollapply(Reg_data$Ex_trn,width = 12,FUN = sum)[-1],x=index(Reg_data[-1:-12,]),type='l',xlab = '
lines(y = Model_12$fitted.values,col='red',x = index(Reg_data[-1:-12,]))
legend(x = as.Date('1963/02/28'),y = -0.3,legend = c('Actual','Predicted'),col = c('black','red'))
```


Excess Rtn
0.4
0.2
0.0
−0.2
−0.6
Actual
Predicted
1970
1980
1990
2000
2010
2020
Time