This problem set asks you to use log dividend-price ratio and book-market ratio to predict the long-horizon 5 year stock market excess returns for United States. Please use the data set we used in Week 12 class, i.e., PredictorData2018part.csv, to answer the following questions. Note that the variables CRSP_SPvw and Rfree in the csv file are already log returns.

1. Please calculate 5-year cumulative stock market excess returns in both simple excess returns and log excess returns. You can do the calculation in the data csv file. Or if you prefer, you can also use R program to calculate. After your calculation, please show the first six and last six observations of the excess returns, so that the TA can check whether your calculation is right or wrong.

Calculate 5-year cumulative log excess return *logexret* by 5yr_CRSP_SPvw - 5yr_Rfree
Calculate 5-year cumulative simple excess return by 5yr_simpleret – 5yr_simpleRf

```
> setwd("~/Desktop/EF4822")
> da = read.csv("PredictorData2018part.csv")
> head(da)
  yyyy Index  D12  E12       b.m    tbl    AAA    BAA    lty cay         ntis  Rfree         infl       eqis          ltr
1 1927 17.66 0.77 1.11 0.3746886 0.0317 0.0446 0.0532 0.0316 NaN  0.076474752 0.0317 -0.022598870 0.26551235  0.089448628
2 1928 24.35 0.85 1.38 0.2596667 0.0426 0.0461 0.0560 0.0340 NaN  0.063068738 0.0426 -0.011560694 0.49742929  0.000827246
3 1929 21.45 0.97 1.61 0.3384578 0.0303 0.0467 0.0595 0.0340 NaN  0.163522172 0.0303  0.005847953 0.72059294  0.034099467
4 1930 15.34 0.98 0.97 0.5547454 0.0148 0.0452 0.0671 0.0330 NaN  0.113885891 0.0148 -0.063953488 0.30784749  0.046429195
5 1931  8.12 0.82 0.61 1.1707317 0.0241 0.0532 0.1042 0.0407 NaN -0.012944196 0.0241 -0.093167702 0.14466470 -0.053157349
6 1932  6.89 0.50 0.41 1.4420843 0.0004 0.0459 0.0842 0.0315 NaN -0.005031571 0.0004 -0.102739726 0.03726708  0.168452113
        corpr        svar csp  ik   CRSP_SPvw CRSP_SPvwx   X5yr_CRSP X5yr_Rf X5yr_logexret X1yr_simpleret X1yr_simpleRf
1  0.07443637 0.009419065 NaN NaN  0.35879164  0.2945602 -0.05939666  0.1435    -0.2028967     0.43159849    0.03220780
2  0.02841156 0.019799325 NaN NaN  0.38844041  0.3331307 -0.50709568  0.1122    -0.6192957     0.47467910    0.04352040
3  0.03273004 0.124614012 NaN NaN -0.08834698 -0.1213454 -0.36302437  0.0725    -0.4355244    -0.08455682    0.03076372
4  0.07975053 0.066648919 NaN NaN -0.26302852 -0.2958606 -0.29951927  0.0445    -0.3440193    -0.23128003    0.01491006
5 -0.01850982 0.159402740 NaN NaN -0.45525321 -0.4892035  0.41983993  0.0312     0.3886399    -0.36571266    0.02439275
6  0.10820224 0.307451657 NaN NaN -0.08890738 -0.1483694  1.20927150  0.0083     1.2009715    -0.08506969    0.00040008
  X5yr_simpleret X5yr_simpleRf X5yr_simexret
1      0.9423329      1.154307    -0.2119739
2      0.6022421      1.118737    -0.5164944
3      0.6955695      1.075193    -0.3796233
4      0.7411744      1.045505    -0.3043305
5      1.5217180      1.031692     0.4900261
6      3.3510425      1.008335     2.3427080
```

Due to generating 5-year cumulative excess return, there is no data in the last four lines of the csv file, so lines 83 to 88 in the file are the last six observations

```
> da[83:88,]
   yyyy   Index    D12    E12       b.m    tbl    AAA    BAA    lty         cay         ntis  Rfree        infl      eqis
83 2009 1115.10 22.405  50.97 0.3255307 0.0005 0.0526 0.0637 0.0458 -0.01687832  0.010533821 0.0005 0.027213311 0.1971352
84 2010 1257.64 22.729  77.35 0.3581003 0.0014 0.0502 0.0610 0.0414 -0.02395380  0.012573506 0.0014 0.014957235 0.1297017
85 2011 1257.60 26.425  86.95 0.3572325 0.0001 0.0393 0.0525 0.0248 -0.01667665 -0.006495091 0.0001 0.029624188 0.1281862
86 2012 1426.19 31.247  86.51 0.3490324 0.0007 0.0365 0.0463 0.0241 -0.02421396 -0.012258026 0.0007 0.017410224 0.1164981
87 2013 1848.36 34.992 100.20 0.3044081 0.0007 0.0462 0.0538 0.0367 -0.01683586  0.012168529 0.0007 0.015017356 0.1286623
88 2014 2058.90 39.443 102.31 0.3237557 0.0003 0.0379 0.0474 0.0240 -0.02241051  0.005656006 0.0003 0.007564933 0.1093213
          ltr       corpr       svar csp         ik CRSP_SPvw   CRSP_SPvwx X5yr_CRSP X5yr_Rf X5yr_logexret X1yr_simpleret
83 -0.14898273  0.03001226 0.07439572 NaN 0.02969122 0.2648018  0.232909808 0.9207777  0.0034     0.9173777     0.30317268
84  0.10143661  0.12438739 0.03253249 NaN 0.03165262 0.1515208  0.126973639 0.7912348  0.0032     0.7880348     0.16360255
85  0.27096245  0.17939641 0.05395472 NaN 0.03384285 0.0182319 -0.002845667 0.6546648  0.0041     0.6505648     0.01839912
86  0.03438618  0.10674847 0.01617660 NaN 0.03444347 0.1609600  0.134092198 0.7543651  0.0091     0.7452651     0.17463803
87 -0.12786088 -0.07079492 0.01248302 NaN 0.03496345 0.3252631  0.297248463 0.8132963  0.0216     0.7916963     0.38439479
88  0.24741715  0.17274067 0.01292945 NaN 0.03568950 0.1352590  0.112101849 0.4430314  0.0446     0.3984314     0.14483322
   X1yr_simpleRf X5yr_simpleret X5yr_simpleRf X5yr_simexret
83   0.000500125       2.511243      1.003406     1.5078367
84   0.001400980       2.206119      1.003205     1.2029138
85   0.000100005       1.924497      1.004108     0.9203888
86   0.000700245       2.126261      1.009142     1.1171197
87   0.000700245       2.255330      1.021835     1.2334950
88   0.000300045       1.557421      1.045610     0.5118116
```

Conclusion:

	5yr_logexret	*5yr_exret*
First six observations	-0.20289666	-0.21197390
	-0.61929568	-0.51649444
	-0.43552437	-0.37962332
	-0.34401927	-0.30433053
	0.38863993	0.49002614
	1.20097150	2.34270798
Last six observations	0.91737766	1.50783674
	0.78803482	1.20291377
	0.65056476	0.92038881
	0.74526513	1.11711966
	0.79169630	1.23349501
	0.39843136	0.51181164

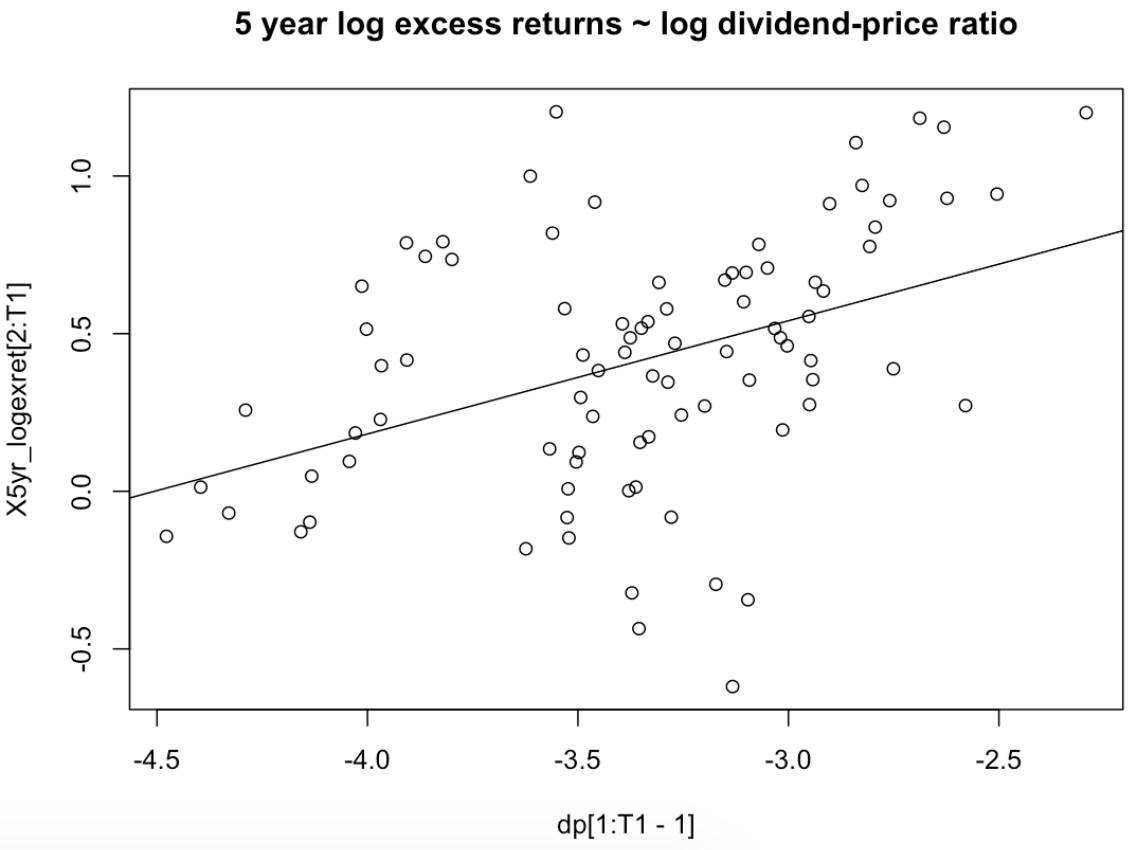
2. Please use log dividend-price ratio to predict the 5-year stock market excess returns (both simple excess returns and log excess returns). Does log dividend-price ratio significantly predict the 5-year excess returns at 5% level? What is the prediction sign, positive or negative? What is the R^2? Which excess return does log dividend-price ratio predict better, simple excess return or log excess return? Please make the plot for describing the linear relationship between predicted excess returns and log dividend-price ratio.

```
> setwd("~/Downloads/CityU/Sem B 2020-2021/EF 4822 Financial Econometrics /PS3 ")
> da=read.csv("PredictorData2018part.csv")
> dp=log(da[,3]/da[,2])
> X5yr_logexret=da$X5yr_logexret
> X5yr_simexret=da$X5yr_simexret
> T1=length(X5yr_logexret)
> lmlgdp=lm(X5yr_logexret[2:T1]~dp[1:T1-1])
> summary(lmlgdp)

Call:
lm(formula = X5yr_logexret[2:T1] ~ dp[1:T1 - 1])

Residuals:
     Min       1Q   Median       3Q      Max
-1.11301 -0.21926  0.01339  0.22291  0.86055

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.61977    0.28903   5.604 2.54e-07 ***
dp[1:T1 - 1]  0.35946    0.08537   4.210 6.30e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3681 on 85 degrees of freedom
  (4 observations deleted due to missingness)
Multiple R-squared:  0.1726,    Adjusted R-squared:  0.1628
F-statistic: 17.73 on 1 and 85 DF,  p-value: 6.304e-05
```

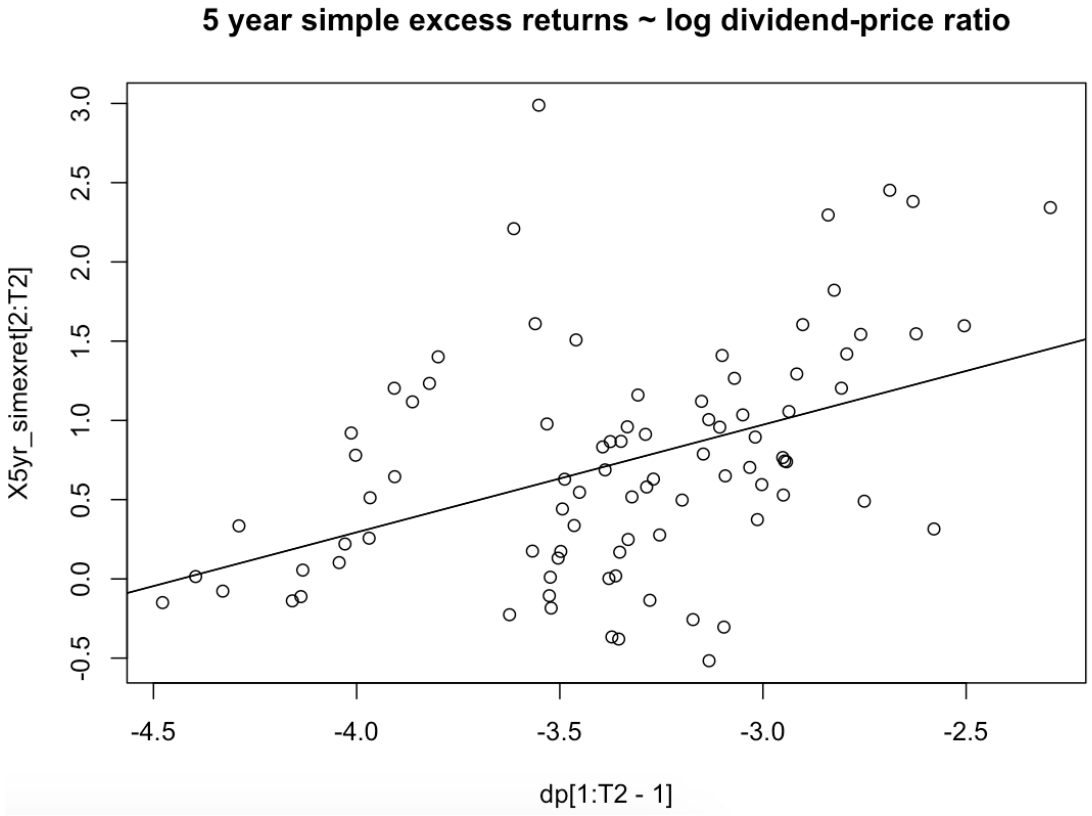
```
> T2=length(X5yr_simexret)
> lmdp=lm(X5yr_simexret[2:T2]~dp[1:T2-1])
> summary(lmdp)
```

```
Call:
lm(formula = X5yr_simexret[2:T2] ~ dp[1:T2 - 1])

Residuals:
    Min      1Q  Median      3Q     Max
-1.3993 -0.3942 -0.0581  0.3121  2.3903

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.0092     0.5177   5.812 1.05e-07 ***
dp[1:T2 - 1]   0.6788     0.1529   4.439 2.70e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6593 on 85 degrees of freedom
  (4 observations deleted due to missingness)
Multiple R-squared:  0.1882,    Adjusted R-squared:  0.1786
F-statistic:  19.7 on 1 and 85 DF,  p-value: 2.701e-05
```

```
> plot(x=dp[1:T1-1],y=X5yr_logexret[2:T1],main="5 year log excess returns ~ log dividend-price ratio")
> abline(lm(X5yr_logexret[2:T1]~dp[1:T1-1]))
```




```
> plot(x=dp[1:T2-1],y=X5yr_simexret[2:T2],main="5 year simple excess returns ~ log dividend-price ratio")
> abline(lm(X5yr_simexret[2:T2]~dp[1:T2-1]))
```



> anova(lmlgdp)

```
Analysis of Variance Table

Response: X5yr_logexret[2:T1]
              Df  Sum Sq Mean Sq F value    Pr(>F)
dp[1:T1 - 1]   1  2.4016 2.40157  17.728 6.304e-05 ***
Residuals     85 11.5146 0.13547
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

> anova(lmdp)

```
Analysis of Variance Table

Response: X5yr_simexret[2:T2]
             Df Sum Sq Mean Sq F value    Pr(>F)
dp[1:T2 - 1]  1  8.564  8.5641  19.702 2.701e-05 ***
Residuals    85 36.948  0.4347
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Conclusion:

Log dividend-price ratio can significantly predict both the 5-year log excess returns and 5-year simple excess returns at 5% level.

Reason: Through the summary(lmlgdp) command, we find that p-value for coefficient of log dividend-price ratio with log return is 6.30e-05<5%. Through the summary(lmdp) command, we find that p-value for coefficient of log dividend-price ratio with simple return is 2.70e-05<5%.

Therefore, both p-value < 5%, so we conclude that log dividend-price ratio significantly predicts the 5-year excess returns at 5% level.

The prediction sign for log dividend-price ratio with 5-year log excess return and 5-year simple excess return are both positive. Through the summary(lmlgdp) command, we find that coefficient is 0.35946. Through the summary(lmdp) command, we find that coefficient is 0.6788.

Log dividend-price ratio better predict simple excess return.

R^2 of the 5-year simple excess return model is 0.1882, R^2 of the 5-year log excess return model is 0.1726, so R^2 is larger (0.1882>0.1726), and the estimated coefficient is more significant (4.439>4.210 for t-values)

3. Please repeat the same steps and answer the same questions as in Part 2, but use book-market ratio instead of log dividend-price ratio to predict the 5-year excess returns.

```
> T1=length(X5yr_logexret)
> bm=da[,5]
> lmlgbm=lm(X5yr_logexret[2:T1]~bm[1:T1-1])
> summary(lmlgbm)

Call:
lm(formula = X5yr_logexret[2:T1] ~ bm[1:T1 - 1])

Residuals:
     Min       1Q   Median       3Q      Max
-0.92843 -0.26877  0.01237  0.24298  0.93840

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.1113     0.0985   1.130   0.2618
bm[1:T1 - 1]   0.5281     0.1563   3.378   0.0011 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3799 on 85 degrees of freedom
  (4 observations deleted due to missingness)
Multiple R-squared:  0.1183,    Adjusted R-squared:  0.108
F-statistic: 11.41 on 1 and 85 DF,  p-value: 0.001104
```

```
> T2=length(X5yr_simexret)
> lmbm=lm(X5yr_simexret[2:T2]~bm[1:T2-1])
> summary(lmbm)

Call:
lm(formula = X5yr_simexret[2:T2] ~ bm[1:T2 - 1])

Residuals:
    Min      1Q  Median      3Q     Max
-1.1690 -0.4919 -0.1387  0.3714  2.5109

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.2143     0.1794   1.195   0.2355
bm[1:T2 - 1]   0.9035     0.2847   3.173   0.0021 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6919 on 85 degrees of freedom
  (4 observations deleted due to missingness)
Multiple R-squared:  0.1059,    Adjusted R-squared:  0.09541
F-statistic: 10.07 on 1 and 85 DF,  p-value: 0.002097
```

Conclusion:

Book-market ratio can significantly predict both the 5-year log excess returns and 5-year simple excess returns at 5% level.

Reason: Through the summary(lmlgbm) command, we find that p-value for coefficient of book-market ratio is 0.001104<5%. Through the summary(lmbm) command, we find that p-value for coefficient of book-market ratio is 0.002097<5%. Therefore, both p-value < 5%, so we conclude that book-market ratio can significantly predict both the 5-year log excess returns and 5-year simple excess returns at 5% level.

The prediction sign for book-market ratio with 5-year log excess return and 5-year simple excess return are both positive. Through the summary(lmlgbm) command, we find that coefficient is 0.5281. Through the summary(lmbm) command, we find that coefficient is 0.9035.

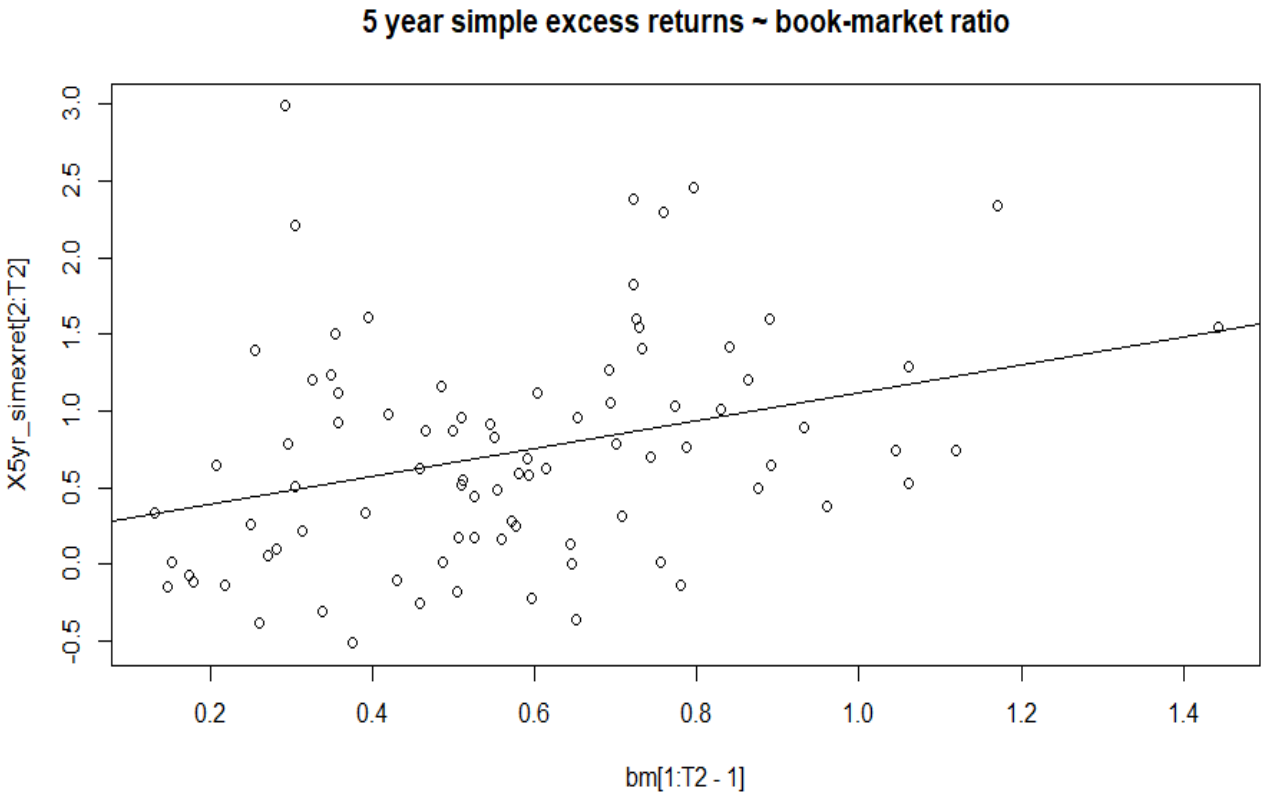
Book-market ratio better predict log excess return.

Reason: R^2 of the 5-year log excess return model is 0.1183, R^2 of the 5-year simple excess return model is 0.1059, so R^2 is larger (0.1183>0.1059), and the estimated coefficient is more significant (3.378 > 3.173 for t-values)

```
> plot(x=bm[1:T1-1],y=X5yr_logexret[2:T1],main="5 year log excess returns ~ book-market ratio")
> abline(lm(X5yr_logexret[2:T1]~bm[1:T1-1]))
```


```
> plot(x=bm[1:T2-1],y=X5yr_simexret[2:T2],main="5 year simple excess returns ~ book-market ratio")
```

> abline(lm(X5yr_simexret[2:T2]~bm[1:T2-1]))



> anova(lmlgbm)

```
Analysis of Variance Table

Response: X5yr_logexret[2:T1]
             Df Sum Sq Mean Sq F value   Pr(>F)
bm[1:T1 - 1]  1  1.647 1.64698   11.41 0.001104 **
Residuals    85 12.269 0.14434
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

> anova(lmbm)

```
Analysis of Variance Table

Response: X5yr_simexret[2:T2]
             Df Sum Sq Mean Sq F value   Pr(>F)
bm[1:T2 - 1]  1  4.821  4.8212  10.071 0.002097 **
Residuals    85 40.691  0.4787
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

4. For 5-year simple excess return, which variable predicts better, log dividend-price ratio or book-market ratio? For 5-year log excess return, which variable predicts better, log dividend-price ratio or book-market ratio?

For 5-year simple excess return:

1) Log dividend-price ratio: R^2 is 0.1882
2) Book-market ratio: R^2 is 0.1059

For 5-year log excess return:

1) Log dividend-price ratio: R^2 is 0.1726
2) Book-market ratio: R^2 is 0.1183

Conclusion:

From the R^2 we can see that both log excess return and simple excess return, the log dividend-price ratio is superior, thus we can conclude for both 5 year simple excess return and 5 year log excess return, log dividend- price ratio predicts better.

5. What are the implications of predictability results in Part 2 and 3 for investment decisions?

From the result of Part 2 and 3, we can easily find that for log dividend-price ratio, it can predict simple excess return better, while book-market ratio predicts log excess return better. Both log dividend-price ratio and book-market ratio can positively predict excess return with all $\beta_1 > 0$ and p-value < 5%, which indicates that the higher the two ratios, the higher the stock excess return. Thus, we can use log dividend-price ratio and book-market ratio as devices for market timing. When the ratios are high, we can increase our allocation in the stock market since the expected returns are high.