Assignment 3

FINA 5250 Empirical Methods in Finance

LIU, Weiyang

wliuax@connect.ust.hk

Question 1

```
##
## Call:
## lm(formula = rGoog_ex ~ rM_ex, data = goog)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.157239 -0.034300 -0.003206  0.029985  0.206897
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.007536   0.004485   1.680   0.0947 .
## rM_ex       0.947432   0.095583   9.912   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05753 on 173 degrees of freedom
## Multiple R-squared:  0.3622, Adjusted R-squared:  0.3585
## F-statistic: 98.25 on 1 and 173 DF,  p-value: < 2.2e-16
```

Question 2

2.1

Yes. β is significantly different from 0 at 0.1% level.

2.2

No. It is significant at 10% level but not at 5% level.

2.3

No. Let's find the 95% confidence interval of β.

```
## Lower bound: 0.756266091111042
```

```
## Upper bound: 1.13859826906802
```

```
## The confidence interval includes 1
```

Therefore, we cannot say beta is significantly different from 1.

Question 3

```
##
## Call:
## lm(formula = rGoog_ex ~ rM_ex + rSmB + rHmL, data = goog)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
```

```
## -0.157640 -0.033382 -0.001348  0.028055  0.202758
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.005524   0.004297   1.286 0.200300
## rM_ex        1.086984   0.097636  11.133  < 2e-16 ***
## rSmB        -0.435487   0.173922  -2.504 0.013220 *
## rHmL        -0.438143   0.121347  -3.611 0.000401 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0548 on 171 degrees of freedom
## Multiple R-squared:  0.4279, Adjusted R-squared:  0.4179
## F-statistic: 42.64 on 3 and 171 DF,  p-value: < 2.2e-16
```

Question 4

As a whole, the factors are significant, because the p-value of F-statistic is smaller than 0.05.

Individually, all factors are significant as well ($p < 0.05$).

Question 5

```
## The single factor can explain 36.22% of the variation
```

```
## The three factors can explain 42.79% of the variation
```

Question 6

```
## Analysis of Variance Table
##
## Model 1: rGoog_ex ~ rM_ex
## Model 2: rGoog_ex ~ rM_ex + rSmB + rHmL
##   Res.Df     RSS Df Sum of Sq      F    Pr(>F)
## 1    173 0.57255
## 2    171 0.51355  2  0.058995 9.8219 9.165e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Yes, the 3-factor model explains significantly more variation ($p < 0.05$).

Question 7

```
##
## Call:
## lm(formula = rGoog_ex ~ rM_ex, data = goog_train)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.153125 -0.039512  0.000441  0.030307  0.201468
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.009934   0.006481   1.533    0.129
## rM_ex       0.818480   0.151493   5.403  4.6e-07 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.06195 on 98 degrees of freedom
## Multiple R-squared:  0.2295, Adjusted R-squared:  0.2216
## F-statistic: 29.19 on 1 and 98 DF,  p-value: 4.605e-07
```

```
##
## Call:
## lm(formula = rGoog_ex ~ rM_ex + rSmB + rHmL, data = goog_train)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.16229 -0.03938  0.00186  0.03448  0.19905
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.006249   0.006167   1.013  0.31348
## rM_ex        1.141430   0.165495   6.897 5.64e-10 ***
## rSmB        -0.535994   0.269725  -1.987  0.04975 *
## rHmL        -0.770732   0.232086  -3.321  0.00127 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05819 on 96 degrees of freedom
## Multiple R-squared:  0.3341, Adjusted R-squared:  0.3133
## F-statistic: 16.06 on 3 and 96 DF,  p-value: 1.547e-08
```

```
## Out-of-sample R-squared of the single-factor model is 0.506000462630713
```

```
## Out-of-sample R-squared of the three-factor model is 0.48096378900801
```

The three-factor model explains even less variation than the single-factor model in the testing set, so the single-factor model is better in this case.