The diversified portfolio five different stocks from healthcare and energy sectors including 'UNH' (UnitedHealth Group Inc), 'CAH' (Cardinal Health Inc), 'XOM' (Exxon Mobil Corp), 'LLY' (Eli Lilly And Co), and 'COP' (ConocoPhillips). For 131 monthly observations (T=131), the excess monthly return data for each of the five stocks are presented below:

```
Excess return data of 5 stocks and market risk premium

                 CAH       COP       LLY       UNH       XOM       mrp
Date
2012-03-31  0.042022 -0.007754  0.025549  0.059620  0.001959  0.030632
2012-04-30 -0.020185 -0.058323  0.027112 -0.048036 -0.005196 -0.008197
2012-05-31 -0.021355 -0.032868  0.000819 -0.007068 -0.083324 -0.062951
2012-06-30  0.020257  0.070619  0.047163  0.052128  0.087561  0.038855
2012-07-31  0.025252 -0.015010  0.025401 -0.127367  0.014258  0.011898
...              ...       ...       ...       ...       ...       ...
2022-09-30 -0.078077 -0.079416  0.045532 -0.052365 -0.114520 -0.121296
2022-10-31  0.100972  0.194770  0.082508  0.061920  0.231856  0.042563
2022-11-30  0.015558 -0.057418 -0.013013 -0.054012 -0.027828  0.013053
2022-12-31 -0.080804 -0.084451 -0.059921 -0.074911 -0.055141 -0.104771
2023-01-31 -0.040857 -0.013003 -0.105088 -0.104252  0.005968  0.015953
```

Now, for each stock, I have estimated the full-sample CAPM factor model ($r_i^e = \alpha_i + \beta_i r_m^e + \epsilon$) using ordinary least squares (OLS), least absolute deviations, shrinkage estimator, and Bayesian regression.

CAPM model parameters under Ordinary least squares method

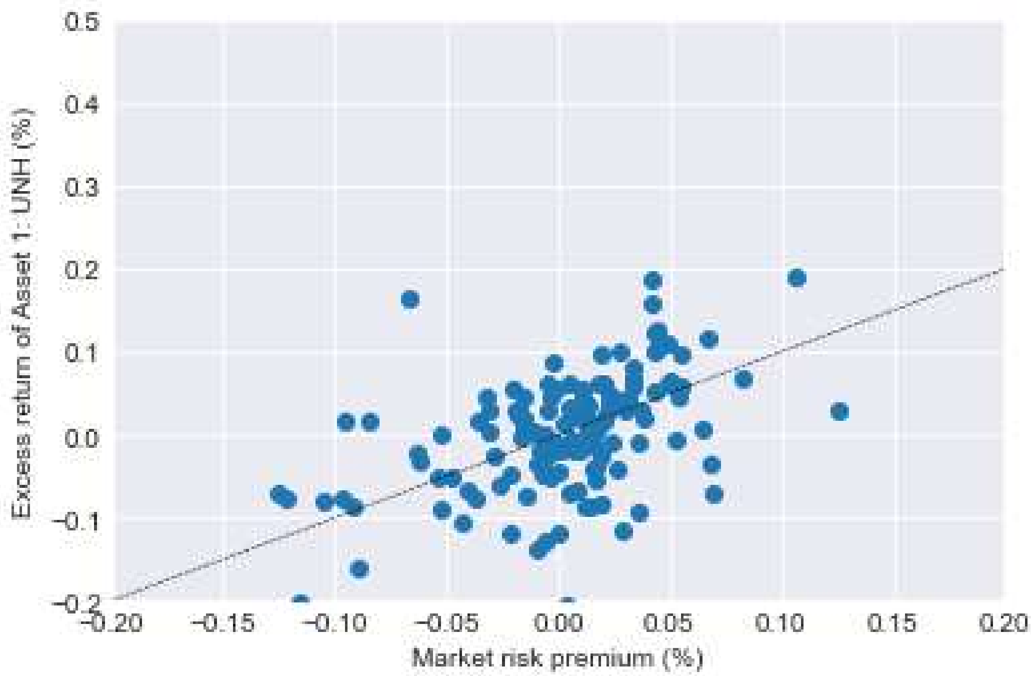
Model Summary for Asset 1:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                    UNH   R-squared:                       0.292
Model:                            OLS   Adj. R-squared:                  0.287
Method:                 Least Squares   F-statistic:                     53.23
Date:                Fri, 10 Feb 2023   Prob (F-statistic):           2.68e-11
Time:                        23:04:33   Log-Likelihood:                 206.43
No. Observations:                 131   AIC:                            -408.9
Df Residuals:                     129   BIC:                            -403.1
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0111      0.004      2.506      0.013       0.002       0.020
mrp            0.7490      0.103      7.296      0.000       0.546       0.952
==============================================================================
Omnibus:                        3.472   Durbin-Watson:                   2.282
Prob(Omnibus):                  0.176   Jarque-Bera (JB):                2.919
Skew:                          -0.304   Prob(JB):                        0.232
Kurtosis:                       3.408   Cond. No.                         23.3
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Based on the OLS model summary, we can generate the following CAPM model for Asset 1 –

$$Excess\ return\ of UNH(\hat{y}) = 0.0111 + 0.7490 * mrp$$

$$where, \beta_0 = 0.0111, \beta_1 = 0.7490\ and\ x = \ mrp\ (market\ risk\ premium)$$



The scatter plot shows that half of the observations indicates a linear negative correlation between excess return of asset 1 (y) and market risk premium (x), while the other half indicates a linear positive correlation between x and y.

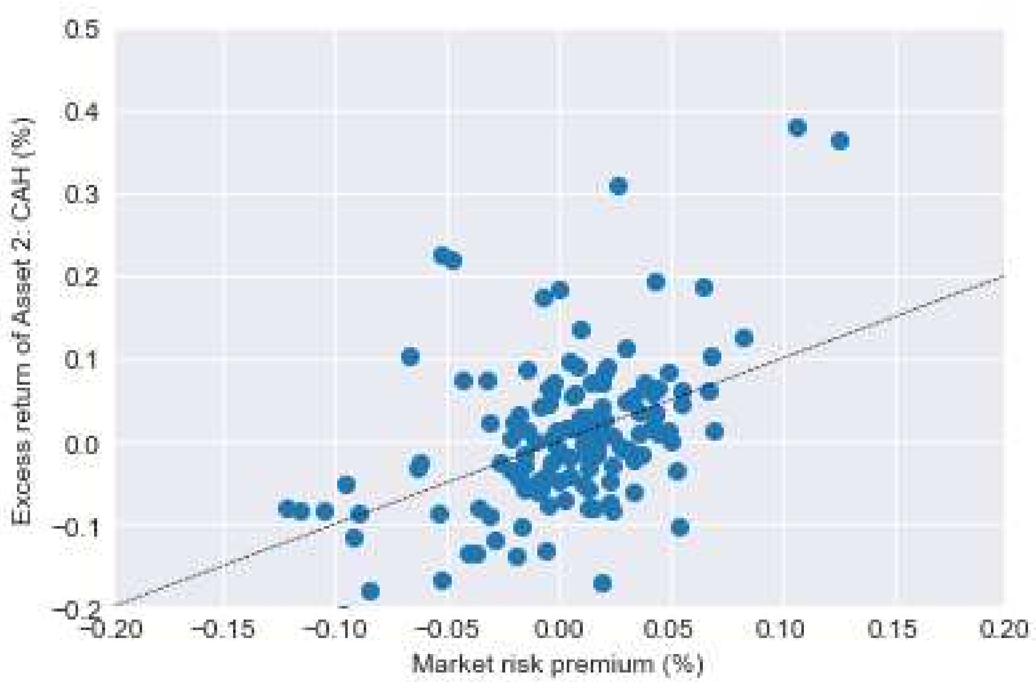
Model Summary for Asset 2:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                    CAH   R-squared:                       0.228
Model:                            OLS   Adj. R-squared:                  0.222
Method:                 Least Squares   F-statistic:                     38.13
Date:                Fri, 10 Feb 2023   Prob (F-statistic):           7.99e-09
Time:                        23:04:32   Log-Likelihood:                 177.48
No. Observations:                 131   AIC:                            -351.0
Df Residuals:                     129   BIC:                            -345.2
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0008      0.006      0.146      0.884      -0.010       0.012
mrp            0.7907      0.128      6.175      0.000       0.537       1.044
==============================================================================
Omnibus:                        5.734   Durbin-Watson:                   2.009
Prob(Omnibus):                  0.057   Jarque-Bera (JB):                7.404
Skew:                          -0.209   Prob(JB):                       0.0247
Kurtosis:                       4.087   Cond. No.                         23.3
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Based on the OLS model summary, we can generate the following CAPM model for Asset 2 –

$$Excess\ return\ of\ CAH(\hat{y}) = 0.0008 + 0.7907 * mrp$$

$$where, \beta_0 = 0.0008, \beta_1 = 0.7907\ and\ x = mrp\ (market\ risk\ premium)$$



The scatter plot shows that half of the observations indicates a linear negative correlation between excess return of asset 2 (y) and market risk premium (x), while the other half indicates a linear positive correlation between x and y.

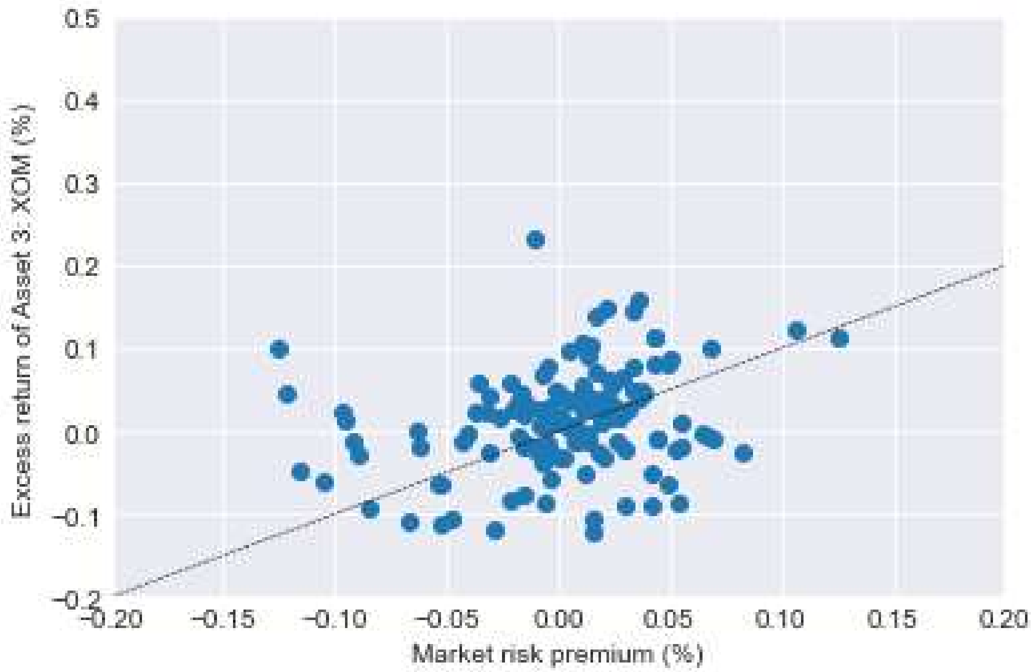
Model Summary for Asset 3:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                    XOM   R-squared:                       0.354
Model:                            OLS   Adj. R-squared:                  0.349
Method:                 Least Squares   F-statistic:                     70.56
Date:                Fri, 10 Feb 2023   Prob (F-statistic):           6.99e-14
Time:                        23:04:33   Log-Likelihood:                 182.66
No. Observations:                 131   AIC:                            -361.3
Df Residuals:                     129   BIC:                            -355.6
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.0010      0.005     -0.183      0.855      -0.011       0.009
mrp            1.0339      0.123      8.400      0.000       0.790       1.277
==============================================================================
Omnibus:                       51.234   Durbin-Watson:                   1.670
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              176.139
Skew:                           1.409   Prob(JB):                     5.65e-39
Kurtosis:                       7.932   Cond. No.                         23.3
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Based on the OLS model summary, we can generate the following CAPM model for Asset 3 –

$$Excess\ return\ of XOM(\hat{y}) = -0.001 + 1.0339 * mrp$$

$$where, \beta_0 = -0.001, \beta_1 = 1.0339\ and\ x = \ mrp\ (market\ risk\ premium)$$



The scatter plot shows that half of the observations indicates a linear negative correlation between excess return of asset 3 (y) and market risk premium (x), while the other half indicates a linear positive correlation between x and y.

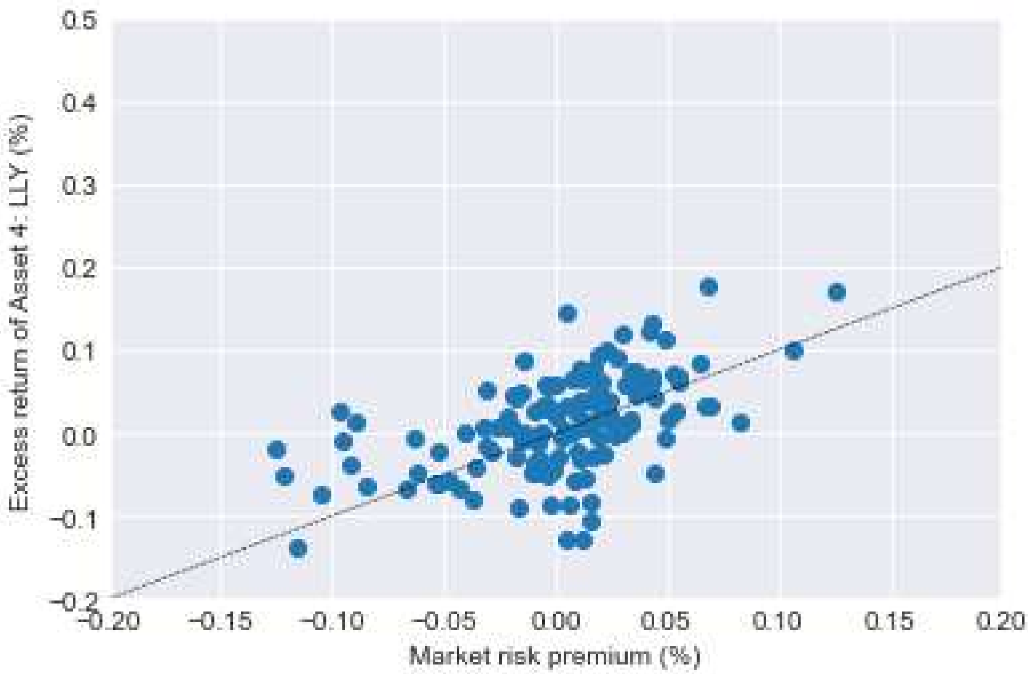
Model Summary for Asset 4:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                    LLY   R-squared:                       0.068
Model:                            OLS   Adj. R-squared:                  0.061
Method:                 Least Squares   F-statistic:                     9.473
Date:                Fri, 10 Feb 2023   Prob (F-statistic):            0.00255
Time:                        23:04:33   Log-Likelihood:                 183.05
No. Observations:                 131   AIC:                            -362.1
Df Residuals:                     129   BIC:                            -356.3
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0127      0.005      2.411      0.017       0.002       0.023
mrp            0.3777      0.123      3.078      0.003       0.135       0.621
==============================================================================
Omnibus:                        5.069   Durbin-Watson:                   1.761
Prob(Omnibus):                  0.079   Jarque-Bera (JB):                5.481
Skew:                           0.260   Prob(JB):                       0.0646
Kurtosis:                       3.857   Cond. No.                         23.3
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Based on the OLS model summary, we can generate the following CAPM model for Asset 4 –

$$Excess\ return\ of LLY(\hat{y}) = 0.0127 + 0.3777 * mrp$$

$$where, \beta_0 = 0.0127, \beta_1 = 0.3777\ and\ x = \ mrp\ (market\ risk\ premium)$$



The scatter plot shows that just more than half of the observations indicates a linear positive correlation between excess return of asset 4 (y) and market risk premium (x), while the rest indicates a linear negative correlation between x and y.

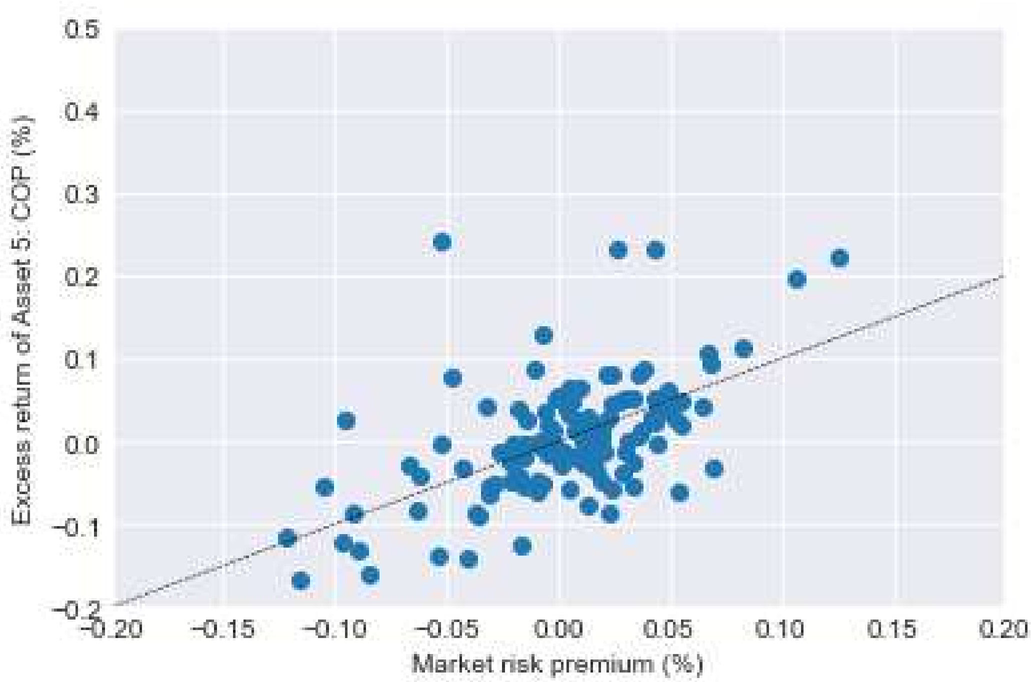
Model Summary for Asset 5:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                    COP   R-squared:                       0.270
Model:                            OLS   Adj. R-squared:                  0.264
Method:                 Least Squares   F-statistic:                     47.71
Date:                Fri, 10 Feb 2023   Prob (F-statistic):           2.03e-10
Time:                        23:04:33   Log-Likelihood:                 135.23
No. Observations:                 131   AIC:                            -266.5
Df Residuals:                     129   BIC:                            -260.7
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0037      0.008      0.492      0.623      -0.011       0.019
mrp            1.2210      0.177      6.907      0.000       0.871       1.571
==============================================================================
Omnibus:                       24.764   Durbin-Watson:                   1.998
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               38.171
Skew:                           0.929   Prob(JB):                     5.14e-09
Kurtosis:                       4.881   Cond. No.                         23.3
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Based on the OLS model summary, we can generate the following CAPM model for Asset 5 –

$$Excess\ return\ of COP(\hat{y}) = 0.0037 + 1.2210 * mrp$$

$$where, \beta_0 = 0.0037, \beta_1 = 1.2210\ and\ x =\ mrp\ (market\ risk\ premium)$$



The scatter plot shows that just more than half of the observations indicates a linear negative correlation between excess return of asset 5 (y) and market risk premium (x), while the rest indicates a linear positive correlation between x and y.

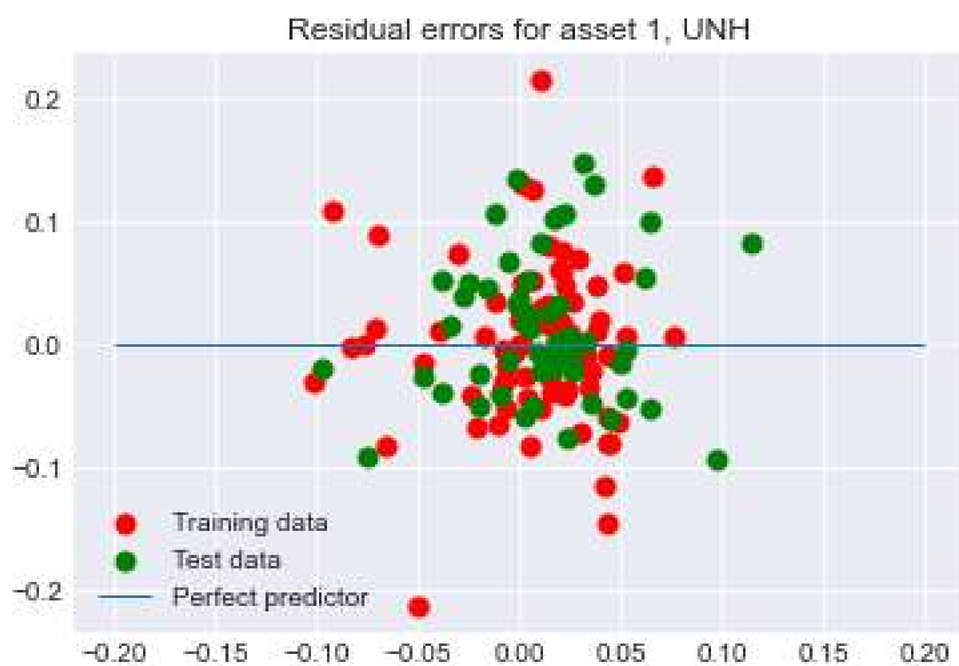
Statistical machine learning approach to test the accuracy of OLS models

In statistical machine learning, Train/Test is a method to measure the accuracy of the model. Training data are the subset of original data that is used to train the machine learning model, whereas testing data are used to check the accuracy of the model. The training dataset is generally larger in size compared to the testing dataset.

Now, let's test our linear model for asset 1 by using the statistical machine learning -

```
Running OLS using the training data:::
Coefficients of asset 1, UNH:  0.854
Variance score with test data of the asset 1, UNH: 0.167
Variance score with train data of the asset 1, UNH: 0.243
Mean Squared Error of the asset 1, UNH: 0.004
Root Mean Squared Error of the asset 1, UNH: 0.064

Test the training model:::
Mean Squared Error of the asset 1 (Test data), UNH: 0.004
Root Mean Squared Error of the asset 1 (Test data), UNH: 0.061
```

The variance of residuals is the same in both training model and test model. The standard deviation of the residuals is 0.064 with the training model and 0.061 with the test model (which is 95.3% of training MSE). Therefore, our training model is 95% accurate in predicting standard deviation of the residuals.

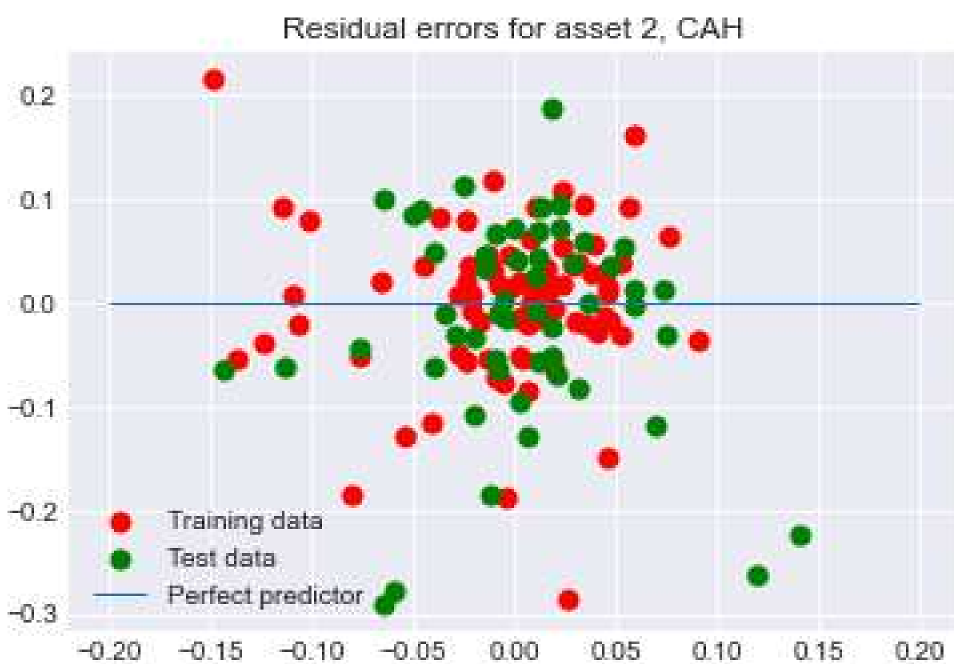


It is observed that the test data and training data are quite similarly spread out around the perfect predictor line (when there is no difference between the predicted model and the true model or when the error term is 0). Here, the training data are highly accurate to estimate residual standard error.

Now, let's test our linear model for asset 2 by using the statistical machine learning –

```
Running OLS using the training data:::
Coefficients of asset 2, CAH:  1.15
Variance score with test data of the asset 2, CAH: 0.214
Variance score with train data of the asset 2, CAH: 0.296
Mean Squared Error of the asset 2, CAH: 0.006
Root Mean Squared Error of the asset 2, CAH: 0.075

Test the training model:::
Mean Squared Error of the asset 2 (Test data), CAH: 0.01
Root Mean Squared Error of the asset 2 (Test data), CAH: 0.101
```

The variance of residuals (0.006) in the training model is much smaller than that (0.01) in the test model. The standard deviation of the residuals is 0.075 with the training model and 0.101 with the test model (which is 1.35x of training MSE). Therefore, our training model is accurate in predicting standard deviation of the residuals.

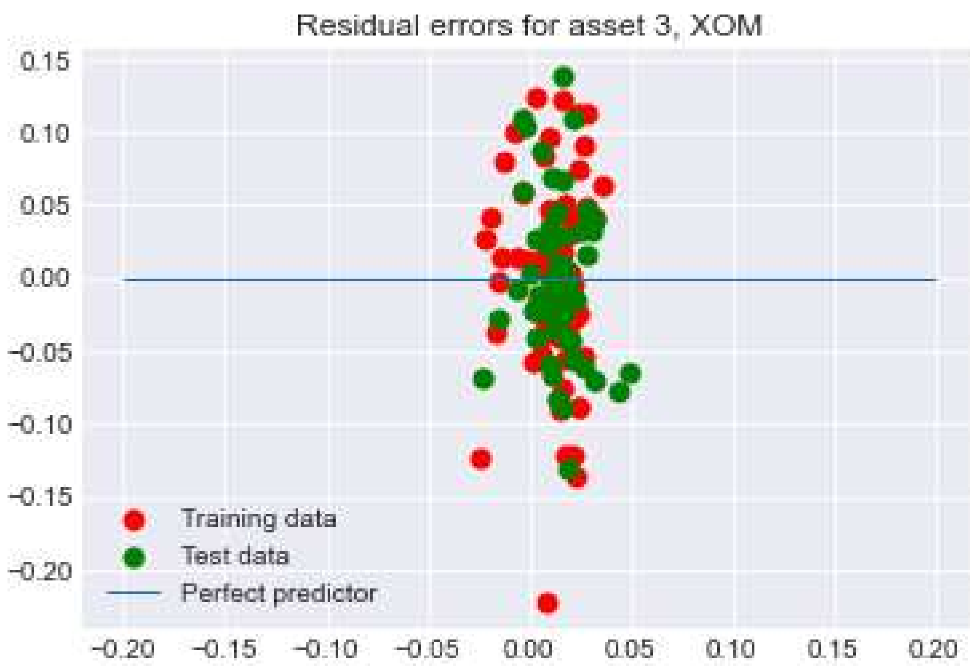


It is observed that the test data and training data are quite similarly spread out around the perfect predictor line (when there is no difference between the predicted model and the true model or when the error term is 0). Here, the training data are highly accurate to estimate residual standard error.

Now, let's test our linear model for asset 3 by using the statistical machine learning –

```
Running OLS using the training data:::
Coefficients of asset 3, XOM:  0.294
Variance score with test data of the asset 3, XOM: 0.109
Variance score with train data of the asset 3, XOM: 0.038
Mean Squared Error of the asset 3, XOM: 0.004
Root Mean Squared Error of the asset 3, XOM: 0.063

Test the training model:::
Mean Squared Error of the asset 3 (Test data), XOM: 0.003
Root Mean Squared Error of the asset 3 (Test data), XOM: 0.056
```

The variance of residuals (0.004) in the training model is slightly higher than that (0.003) in the test model. The standard deviation of the residuals is 0.063 with the training model and 0.056 with the test model (which is 89.0% of training MSE). Therefore, our training model is 89% accurate in predicting standard deviation of the residuals.

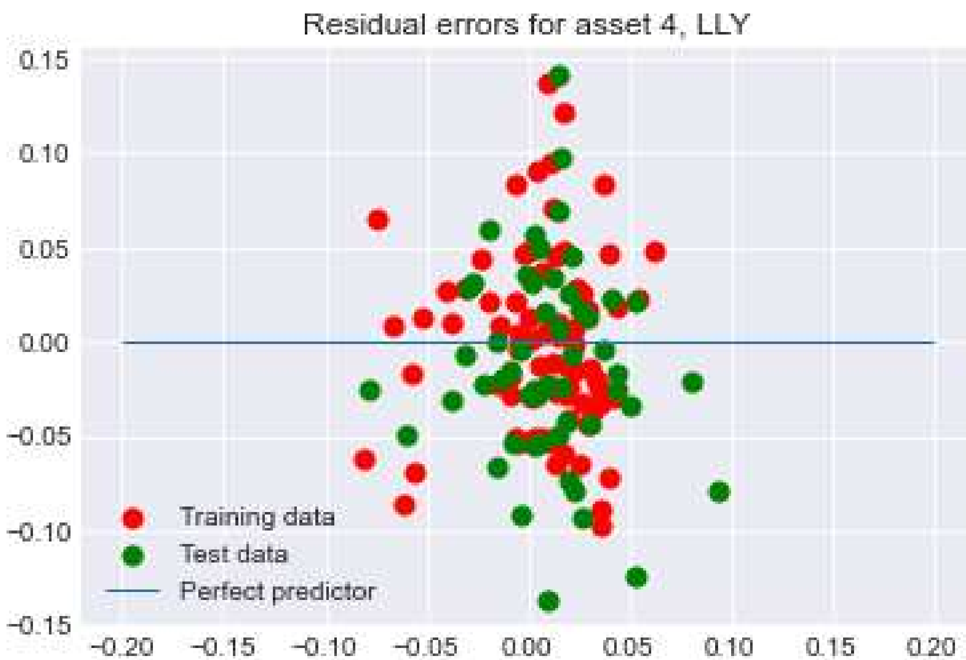


It is observed that the test data and training data are somewhat differently spread out around the perfect predictor line (when there is no difference between the predicted model and the true model or when the error term is 0). Here, the training data are moderately accurate to estimate residual standard error.

Now, let's test our linear model by using the statistical machine learning –

```
Running OLS using the training data:::
Coefficients of asset 4, LLY:  0.692
Variance score with test data of the asset 4, LLY: 0.27
Variance score with train data of the asset 4, LLY: 0.273
Mean Squared Error of the asset 4, LLY: 0.002
Root Mean Squared Error of the asset 4, LLY: 0.048

Test the training model:::
Mean Squared Error of the asset 4 (Test data), LLY: 0.003
Root Mean Squared Error of the asset 4 (Test data), LLY: 0.054
```

The variance of residuals (0.002) in the training model is slightly smaller than that (0.003) in the test model. The standard deviation of the residuals is 0.048 with the training model and 0.054 with the test model (which is 1.13x of training MSE). Therefore, our training model is highly accurate in predicting standard deviation of the residuals.

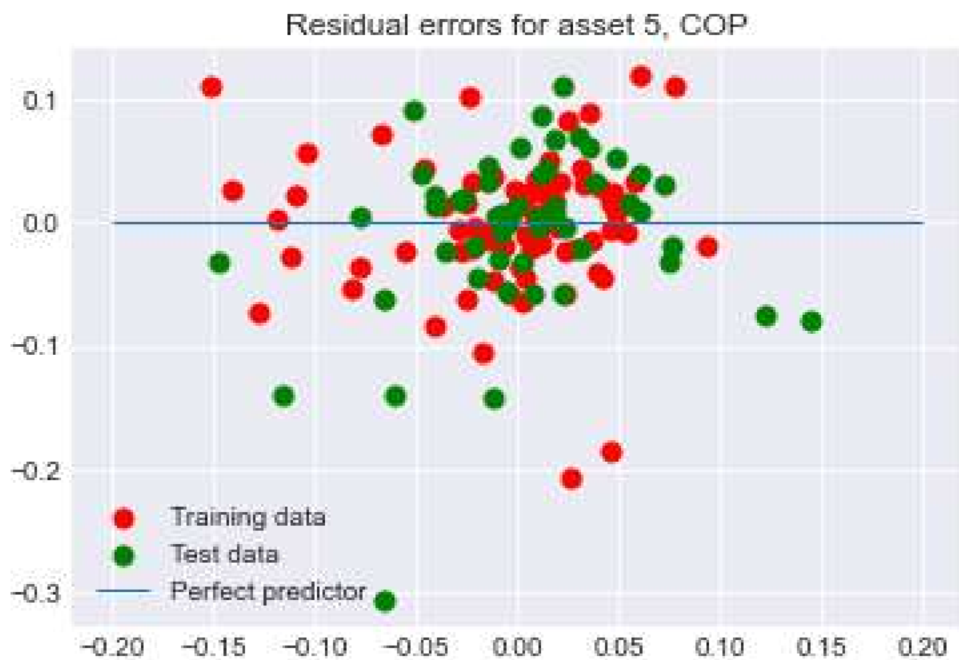


It is observed that the test data and training data are somewhat similarly spread out around the perfect predictor line (when there is no difference between the predicted model and the true model or when the error term is 0). Here, the training data are moderately accurate to estimate residual standard error.

Now, let’s test our linear model by using the statistical machine learning –

```
Running OLS using the training data:::
Coefficients of asset 5, COP:  1.178
Variance score with test data of the asset 5, COP: 0.176
Variance score with train data of the asset 5, COP: 0.459
Mean Squared Error of the asset 5, COP: 0.003
Root Mean Squared Error of the asset 5, COP: 0.054

Test the training model:::
Mean Squared Error of the asset 5 (Test data), COP: 0.005
Root Mean Squared Error of the asset 5 (Test data), COP: 0.068
```

The variance of residuals (0.003) in the training model is slightly smaller than that (0.005) in the test model. The standard deviation of the residuals is 0.054 with the training model and 0.068 with the test model (which is 1.26x of training MSE). Therefore, our training model is highly accurate in predicting standard deviation of the residuals.



It is observed that the test data and training data are quite similarly spread out around the perfect predictor line (when there is no difference between the predicted model and the true model or when the error term is 0). Here, the training data are highly accurate to estimate residual standard error.

The coefficients of all assets using Ordinary Least Squares is shown below -

```
OLS beta:::
                UNH       CAH       XOM       LLY       COP
Intercept  0.000802  0.003741  0.012715  0.011055 -0.000967
mrp        0.790661  1.221021  0.377712  0.749010  1.033872
```

The independent variable (mrp) has more than perfect correlation with the excess return of CAH and COP. Contrarily, the lowest correlation is also observed with the excess return of XOM. In summary, there is a linear positive relationship between market risk premium and excess returns of different assets.

CAPM model parameters under Least absolute deviations method

Least absolute deviations (LAD), also known as least absolute errors (LAE), least absolute residuals (LAR), or least absolute values (LAV), is a statistical optimality criterion and a statistical optimization technique based minimizing the sum of absolute deviations (sum of absolute residuals or sum of absolute errors) or the L1 norm of such values. The following snapshot shows the coefficients of regression for each asset when the residuals are minimized -

Coefficients under least absolute deviations (LAD):::

	Alpha	Beta
UNH	0.000000	0.841859
CAH	-0.001992	0.927014
XOM	0.000000	0.314658
LLY	0.000000	0.764624
COP	-0.005238	0.900959

For each stock, we can develop the CAPM model based on the coefficients under LAD.

$$Excess\ return\ of UNH(\hat{y}) = 0.0 + 0.842 * mrp$$

$$Excess\ return\ of CAH(\hat{y}) = -0.002 + 0.927 * mrp$$

$$Excess\ return\ of XOM(\hat{y}) = 0.0 + 0.315 * mrp$$

$$Excess\ return\ of LLY(\hat{y}) = 0.0 + 0.765 * mrp$$

$$Excess\ return\ of COP(\hat{y}) = -0.005 + 0.901 * mrp$$

Since the residuals are minimized, the factor betas almost perfectly predict the excess return of each stock in our portfolio.

CAPM model parameters under Shrinkage estimator

When sample observations are relatively few but regard multiple assets, shrinkage estimation can improve the determination of single-asset betas. Shrinkage exploits the information in the cross-section of betas, thereby providing estimations for a single asset that are less extreme.

Shrinkage Estimator parameters:::

	UNH	CAH	XOM	LLY	COP
alpha	0.600	0.600	0.60	0.600	0.600
beta	0.808	1.066	0.56	0.783	0.954

Based on the coefficients under the Shrinkage estimator, we can develop the following CAPM models for each stock –

$$Excess\ return\ of UNH(\hat{y}) = 0.6 + 0.808 * mrp$$

$$Excess\ return\ of CAH(\hat{y}) = 0.6 + 1.066 * mrp$$

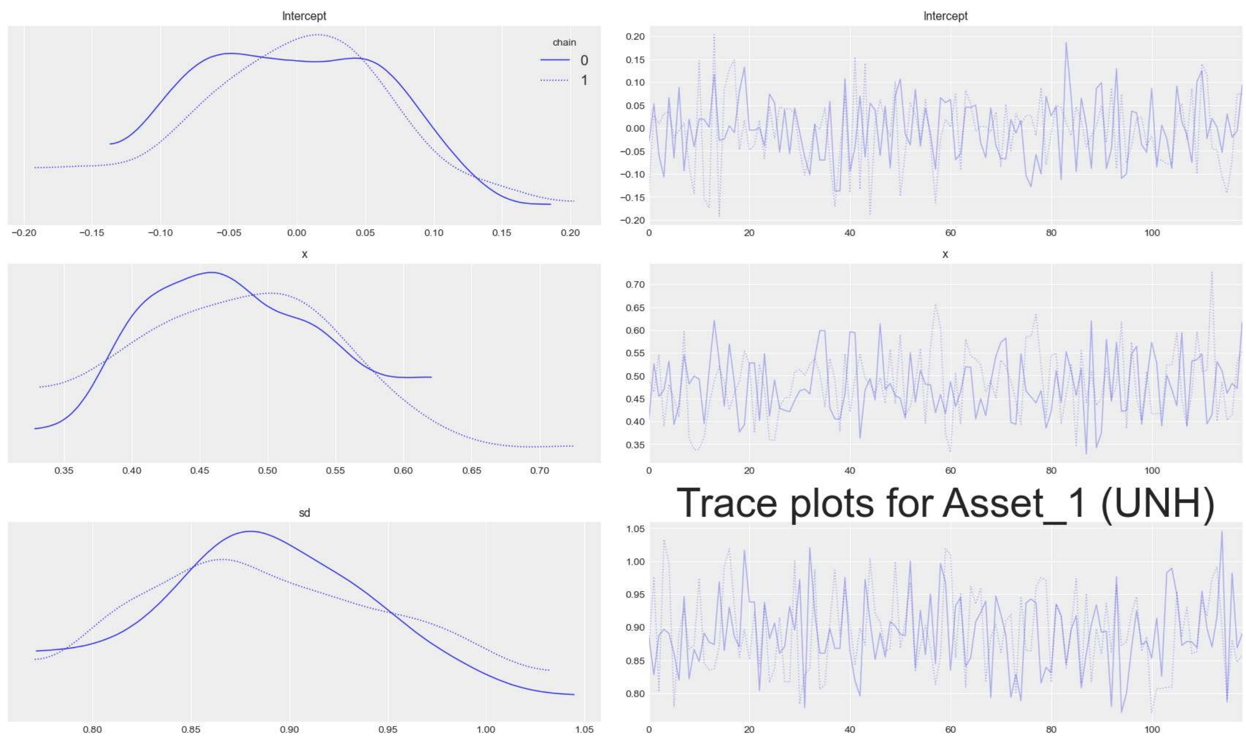
$$Excess\ return\ of XOM(\hat{y}) = 0.6 + 0.56 * mrp$$

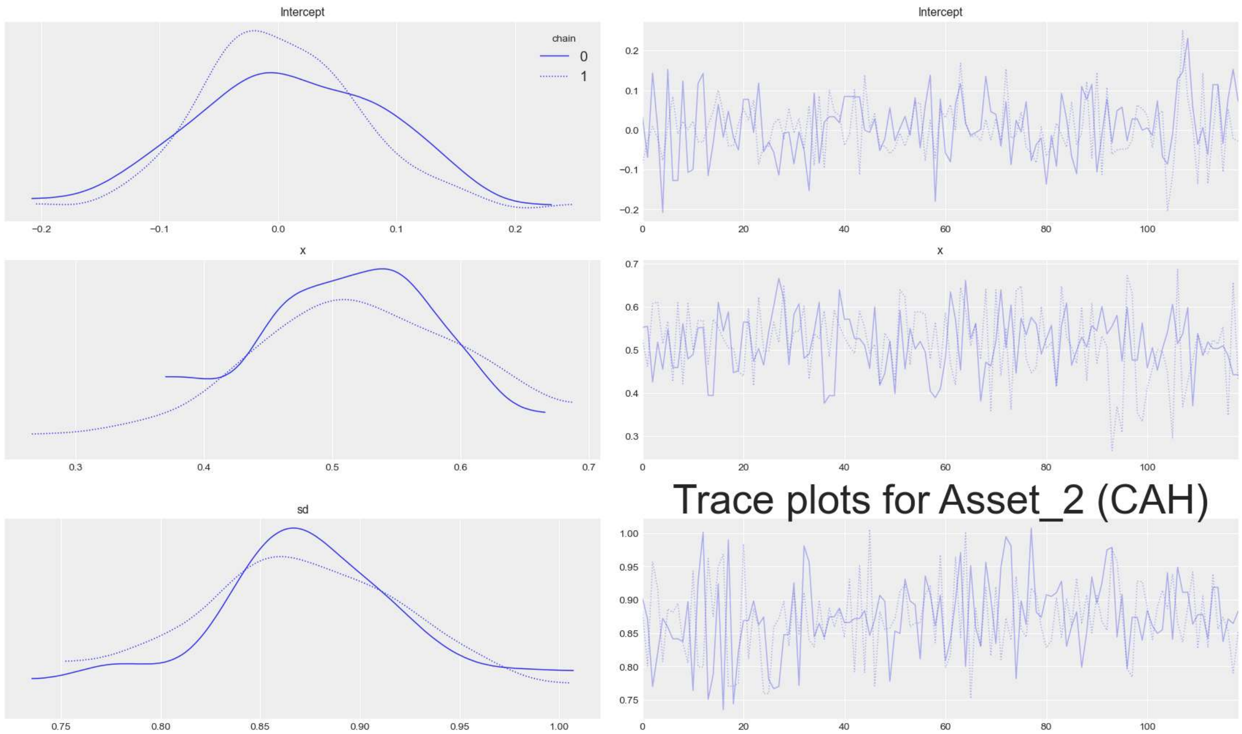
$$Excess\ return\ of LLY(\hat{y}) = 0.6 + 0.783 * mrp$$

$$Excess\ return\ of COP(\hat{y}) = 0.6 + 0.954 * mrp$$

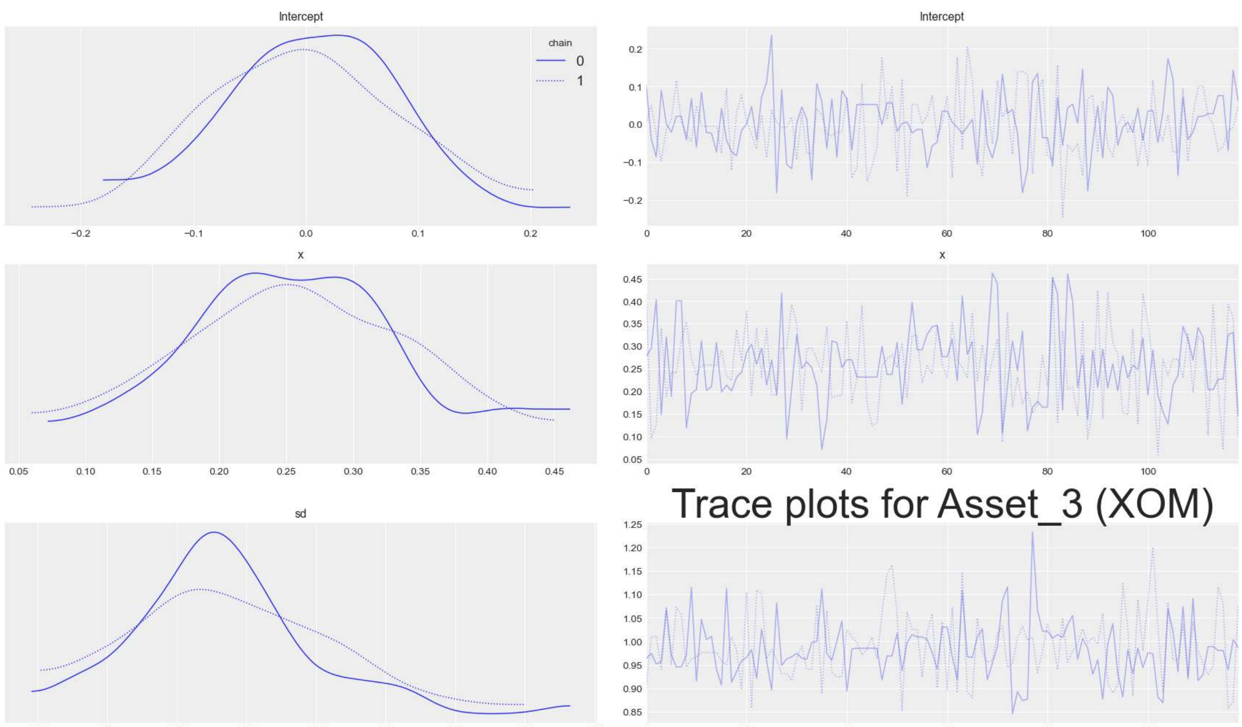
Bayesian regression

Here, we regard the normal distribution from our original sample observations as prior in the Bayesian regression. The following charts shows both priori and posterior distribution of intercept, beta, and standard error for each asset -

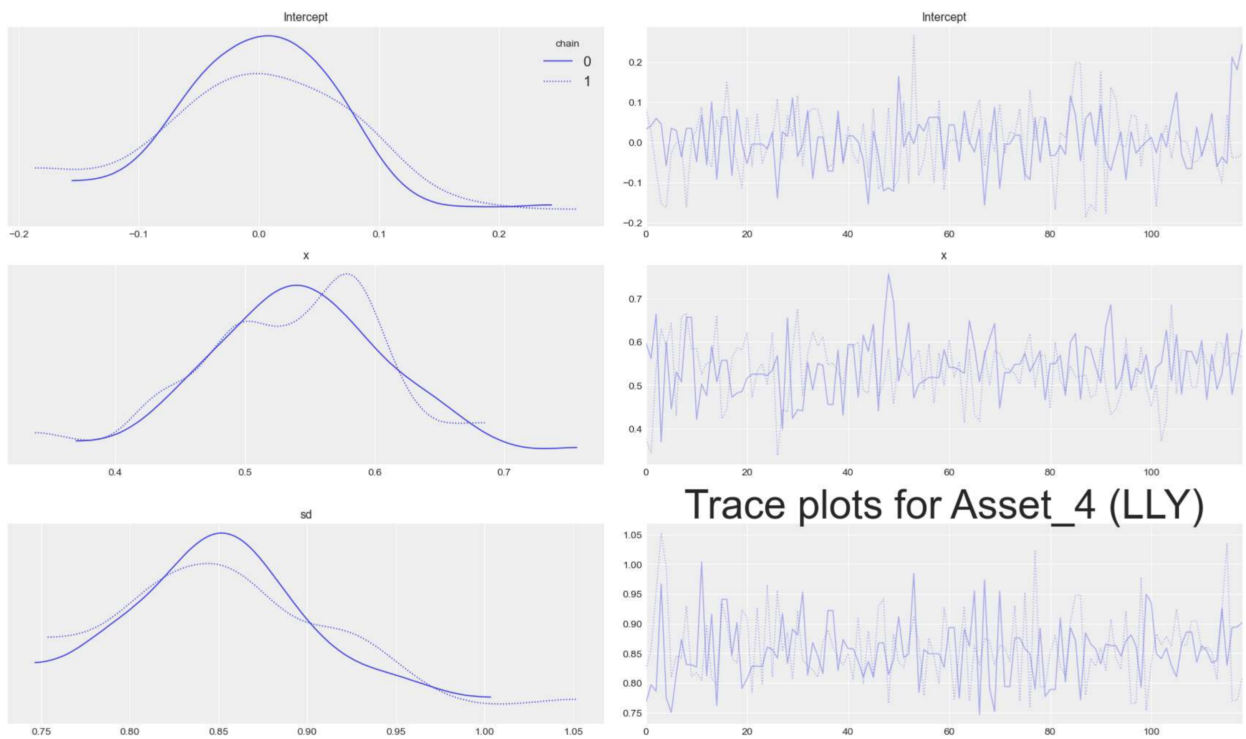


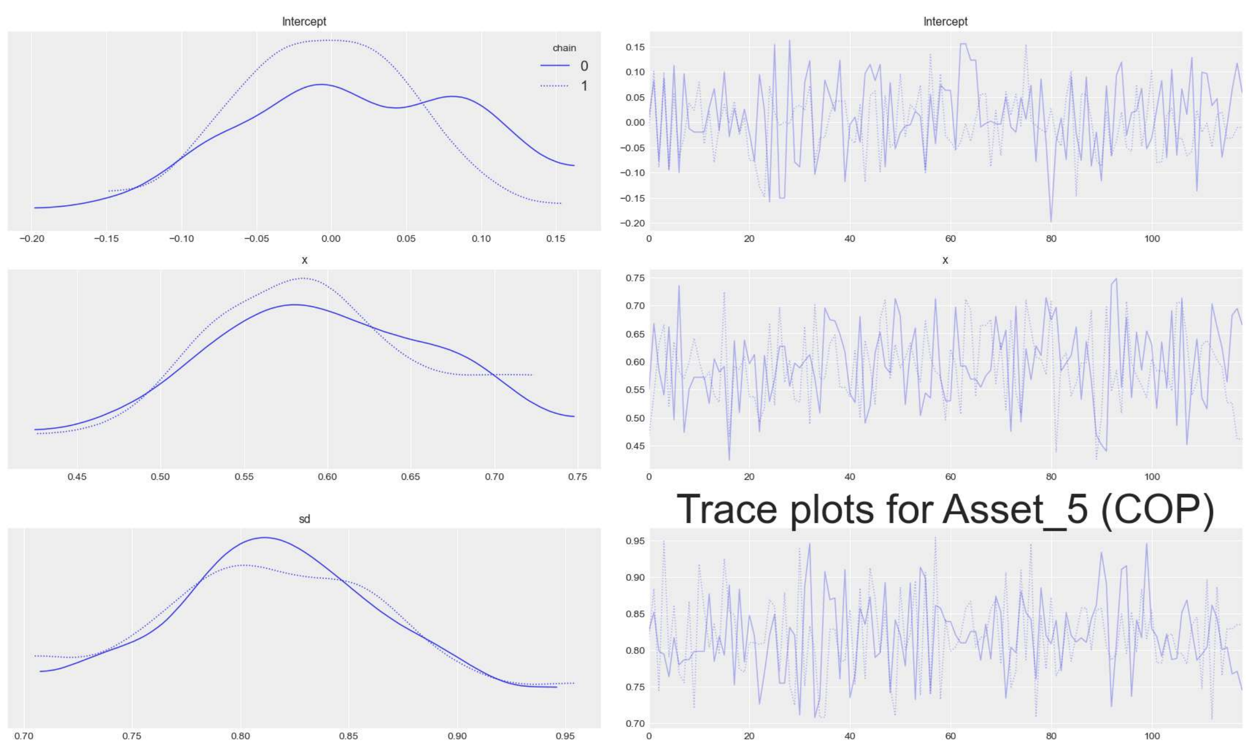


Trace plots for Asset_2 (CAH)



Trace plots for Asset_3 (XOM)



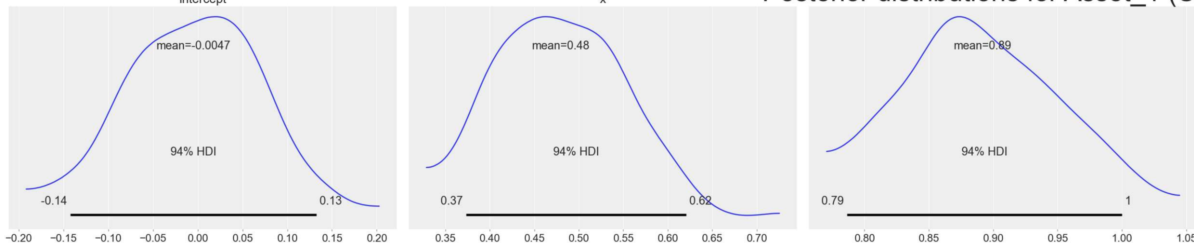


Now, we put the coefficient results under three methods – OLS, Shrinkage estimator, and Least absolute deviations to observe the evolution of model parameters under these methods.

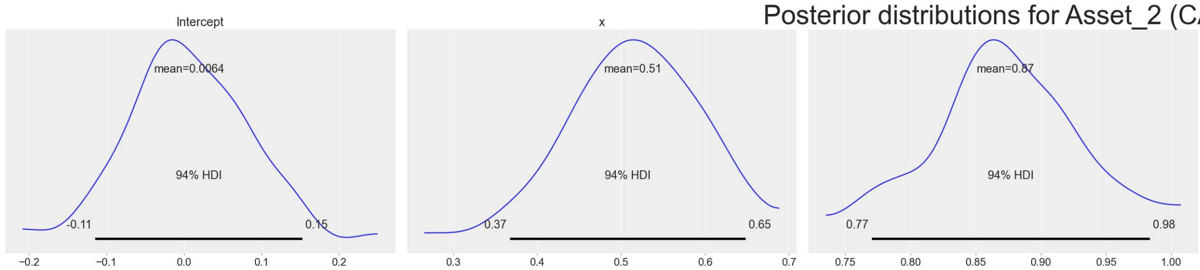
	OLS beta	Shrinkage estimator beta	LAD beta
UNH	0.790661	0.808178	0.841853
CAH	1.221021	1.066395	0.927014
XOM	0.377712	0.560409	0.314666
LLY	0.749010	0.783188	0.764627
COP	1.033872	0.954105	0.902974

Now, we can plot posterior distributions for each asset separately -

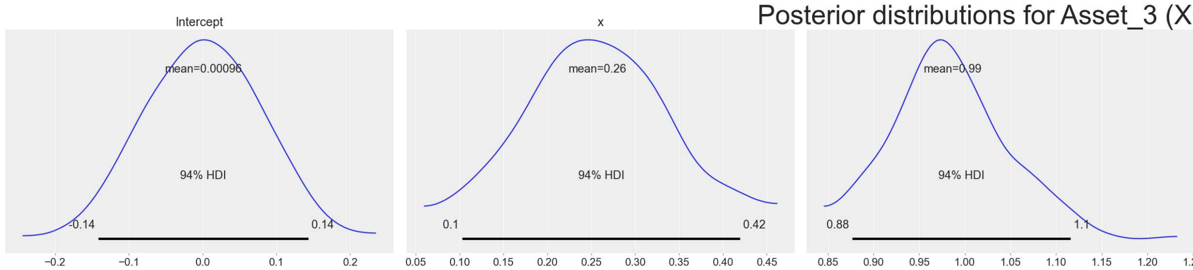
Posterior distributions for Asset_1 (UNH)



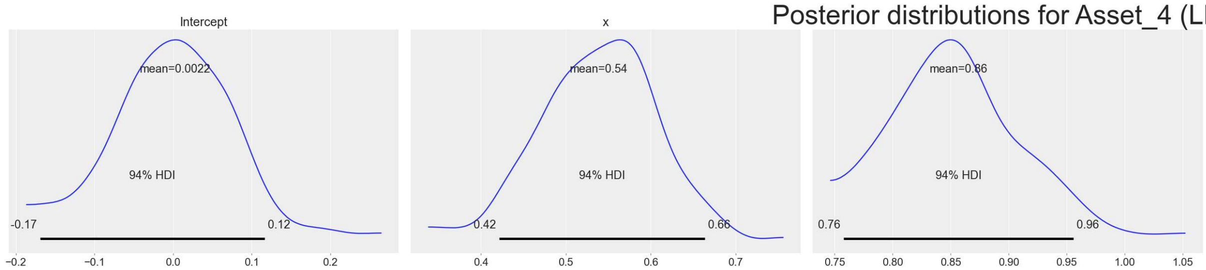
Posterior distributions for Asset_2 (CAH)

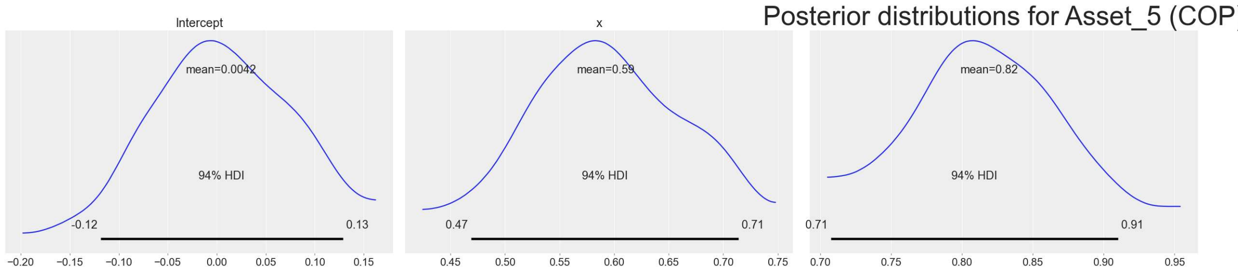


Posterior distributions for Asset_3 (XOM)



Posterior distributions for Asset_4 (LLY)





Now, we can consider the estimates from OLS and use the t-statistic to check whether any of the stocks have alphas that are significantly different from 0.

The result from OLS is shown below -

OLS beta:::

	UNH	CAH	XOM	LLY	COP
alpha	0.0008	0.0037	0.0127	0.0111	-0.0010
beta	0.7907	1.2210	0.3777	0.7490	1.0339
t (alpha)	0.1457	0.4924	2.4109	2.5056	-0.1829
t (beta)	6.1748	6.9071	3.0779	7.2959	8.3997
p (alpha)	0.8844	0.6233	0.0173	0.0135	0.8552
p (beta)	0.0000	0.0000	0.0025	0.0000	0.0000

To test for significance, we set our hypothesis first –

H_0: alpha = 0

H_a: alpha ≠ 0

Considering a 5% level of significance and 130 degrees of freedom (i.e., df = n – 1 = 131 – 1), we get the critical value of t-statistic as 1.98.

We observe that two stocks – Exxon Mobil (‘XOM’) and Elly & Lilly (‘LLY’) – have statistical t values 2.41 and 2.51 respectively, that are greater than our critical value. Therefore, these stocks have alphas that are significantly greater than 0.

Out-of-sample tests:

Now, we can consider a 60-month initial window to get the 1-month out of sample root mean squared error (RMSE) for each stock when fitting returns using OLS with Rolling fixed 60-month estimation window and Cumulative rolling estimation window.

```
1-month out of sample RMSE for each stock with
      a rolling fixed 60-month estimation window:::

                                 UNH       CAH       XOM       LLY  \
60 observations start at month:
1                           0.042766  0.062587  0.043156  0.044698
2                           0.042739  0.062863  0.043155  0.044653
3                           0.045583  0.062859  0.043475  0.044383
4                           0.045478  0.063095  0.043812  0.044429
5                           0.045803  0.062967  0.043772  0.044728
...                              ...       ...       ...       ...
56                          0.071126  0.092588  0.071175  0.050549
57                          0.071224  0.092479  0.070782  0.052470
58                          0.071214  0.092301  0.069358  0.051513
59                          0.072244  0.092415  0.069537  0.052690
60                          0.072575  0.100206  0.070716  0.052731

                                 COP
60 observations start at month:
1                           0.033759
2                           0.033715
3                           0.033714
4                           0.033615
5                           0.032690
...                              ...
56                          0.060213
57                          0.060238
58                          0.060332
59                          0.060325
60                          0.072198

[60 rows x 5 columns]
```

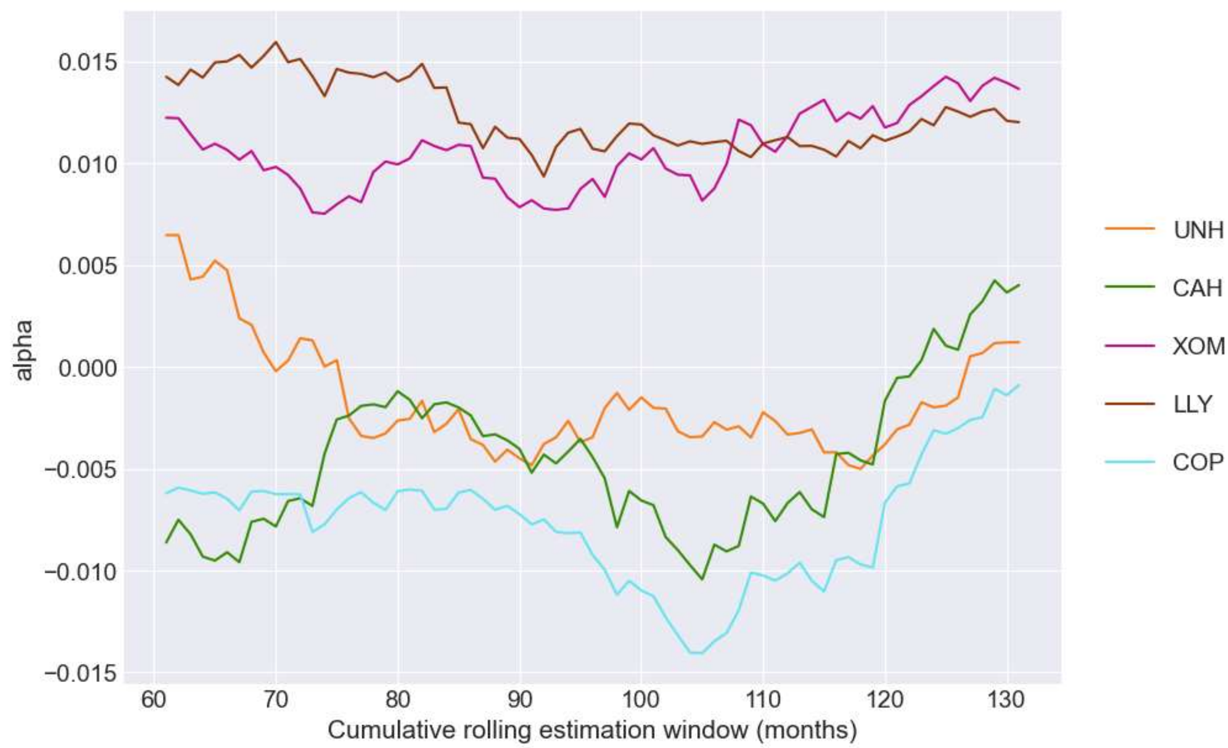
```
1-month out of sample RMSE for each stock with
      a cumulative rolling estimation window:::

                   UNH       CAH       XOM       LLY       COP
Observations
61            0.042414  0.062071  0.042801  0.044330  0.033481
62            0.042070  0.062008  0.042455  0.044051  0.033257
63            0.044574  0.061726  0.042496  0.044037  0.033006
64            0.044236  0.061783  0.042533  0.043788  0.032770
65            0.044242  0.061321  0.042253  0.043763  0.032520
...                ...       ...       ...       ...       ...
127           0.062924  0.086044  0.059015  0.048999  0.058252
128           0.062695  0.085918  0.059193  0.048868  0.058036
129           0.062726  0.086449  0.059147  0.048702  0.060161
130           0.062485  0.086384  0.058985  0.048973  0.060034
131           0.062246  0.086128  0.058835  0.048790  0.060002
```

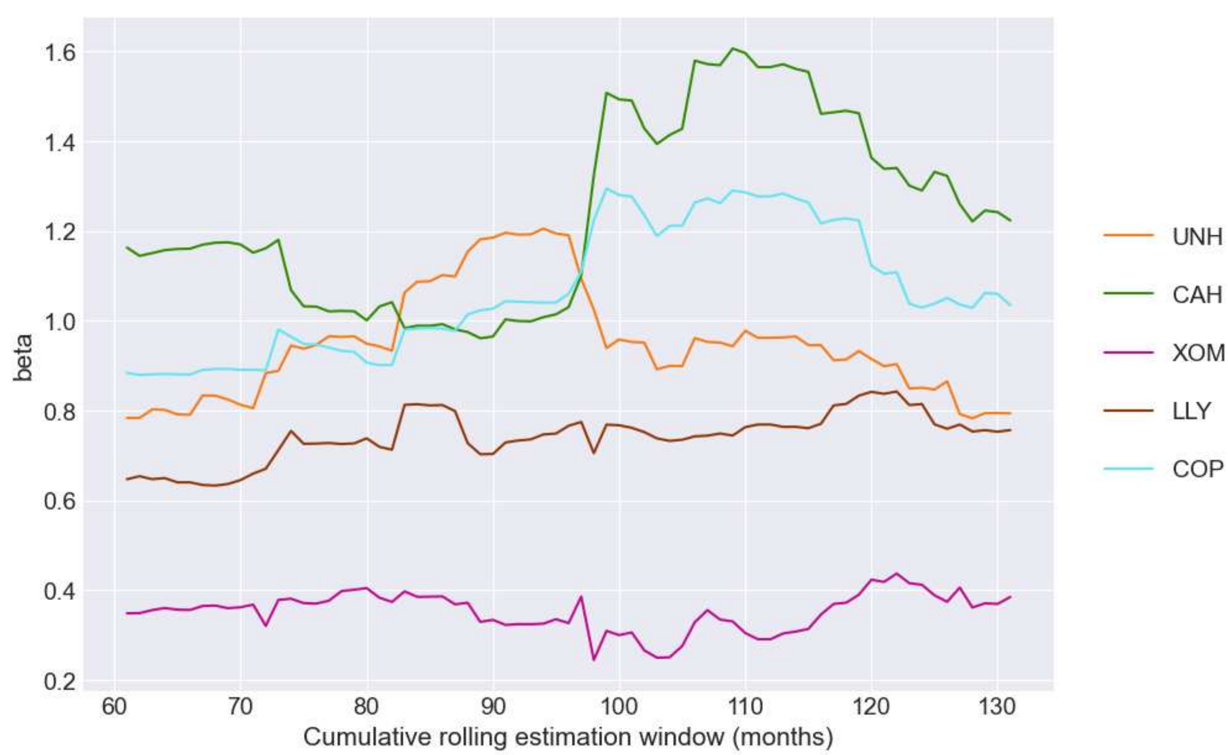
Now, let’s measure the RMSE from assuming an expected return of 0 for each stock -

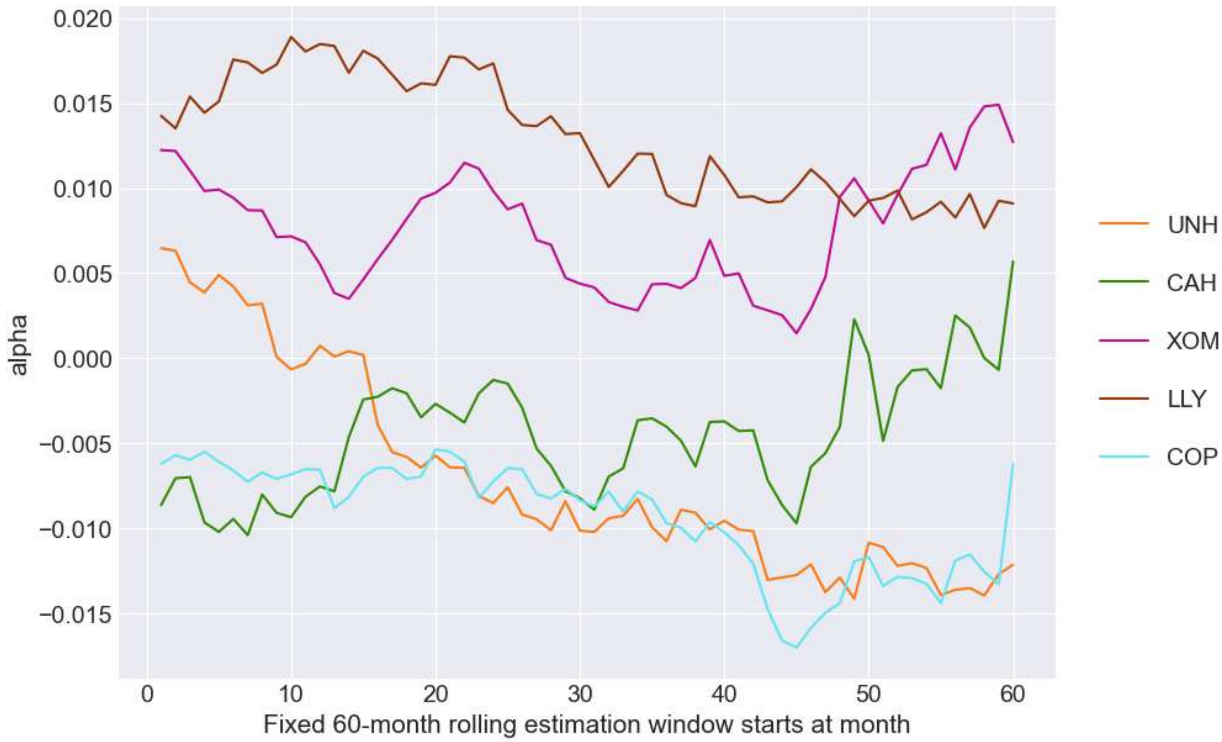
```
1-month out of sample RMSE for each stock (assuming an expected return of 0)
      with a cumulative rolling estimation window:::

                   UNH       CAH       XOM       LLY       COP
Observations
61            0.026775  0.033802  0.019078  0.027943  0.025770
62            0.026576  0.033102  0.019035  0.027672  0.025465
63            0.025903  0.032958  0.018561  0.027993  0.025286
64            0.025811  0.032828  0.018085  0.027718  0.025109
65            0.025735  0.032645  0.018234  0.027943  0.024897
...                ...       ...       ...       ...       ...
127           0.032842  0.051893  0.026250  0.036978  0.041866
128           0.032916  0.051377  0.025654  0.036724  0.042344
129           0.033983  0.053072  0.026747  0.037462  0.044342
130           0.034177  0.052842  0.026926  0.037283  0.044288
131           0.034174  0.052429  0.027009  0.037275  0.043533
```

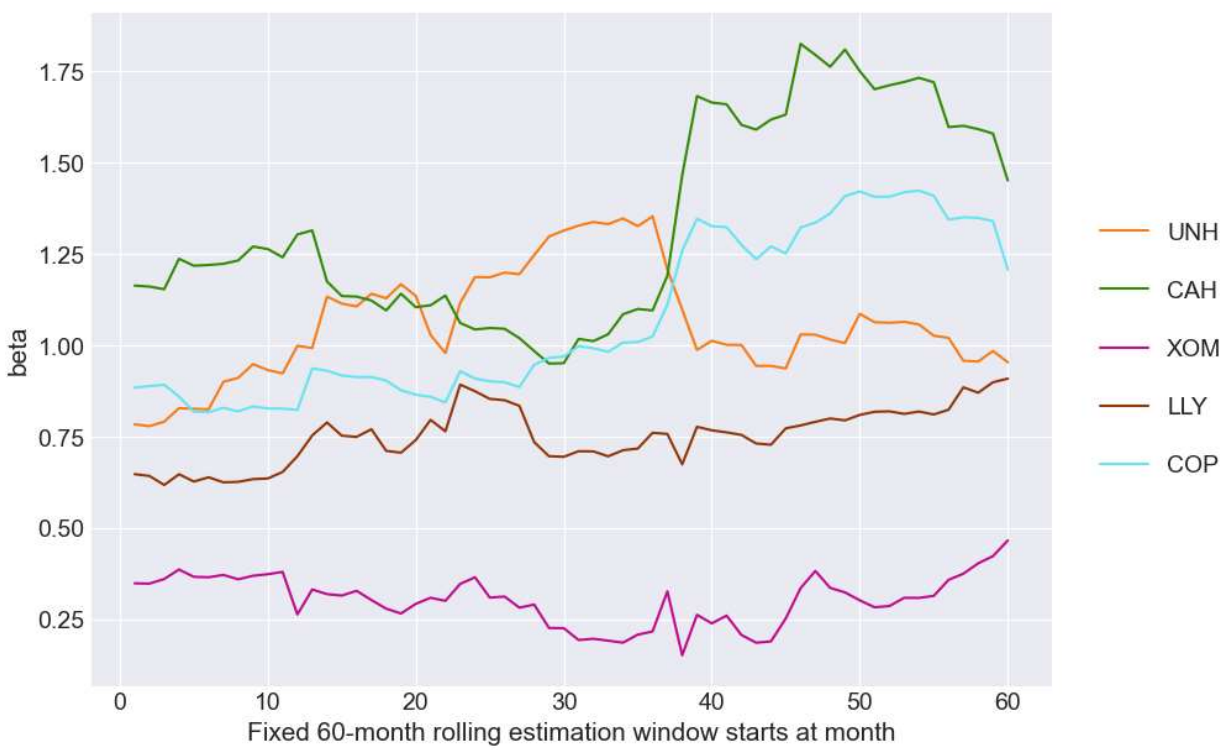
```
1-month out of sample RMSE for each stock (assuming expected return
      of each stock is 0) with a rolling fixed 60-month estimation window:::

                                    UNH       CAH       XOM       LLY  \
60 observations start at month:
1                              0.001202  0.001202  0.001202  0.001202
2                              0.001455  0.001455  0.001455  0.001455
3                              0.001741  0.001741  0.001741  0.001741
4                              0.001972  0.001972  0.001972  0.001972
5                              0.002163  0.002163  0.002163  0.002163
...                                 ...       ...       ...       ...
56                             0.008387  0.008387  0.008387  0.008387
57                             0.008323  0.008323  0.008323  0.008323
58                             0.008397  0.008397  0.008397  0.008397
59                             0.008429  0.008429  0.008429  0.008429
60                             0.008574  0.008574  0.008574  0.008574

                                    COP
60 observations start at month:
1                              0.001202
2                              0.001455
3                              0.001741
4                              0.001972
5                              0.002163
...                                 ...
56                             0.008387
57                             0.008323
58                             0.008397
59                             0.008429
60                             0.008574
```

For both cumulative rolling and fixed rolling estimation window, we can plot the time series of regression coefficients for each stock -


0.015
0.010
0.005
0.000
−0.005
−0.010
−0.015
alpha
60
70
80
90
100
110
120
130
Cumulative rolling estimation window (months)
UNH
CAH
XOM
LLY
COP


1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
beta
60
70
80
90
100
110
120
130
Cumulative rolling estimation window (months)
UNH
CAH
XOM
LLY
COP





Now, for the rolling fixed 60-month window, we can use the fixed rolling factor model results to allocate stocks into the maximum Sharpe ratio portfolio and get the out-of-sample utility for a mean variance investor with a risk aversion coefficient of 4.

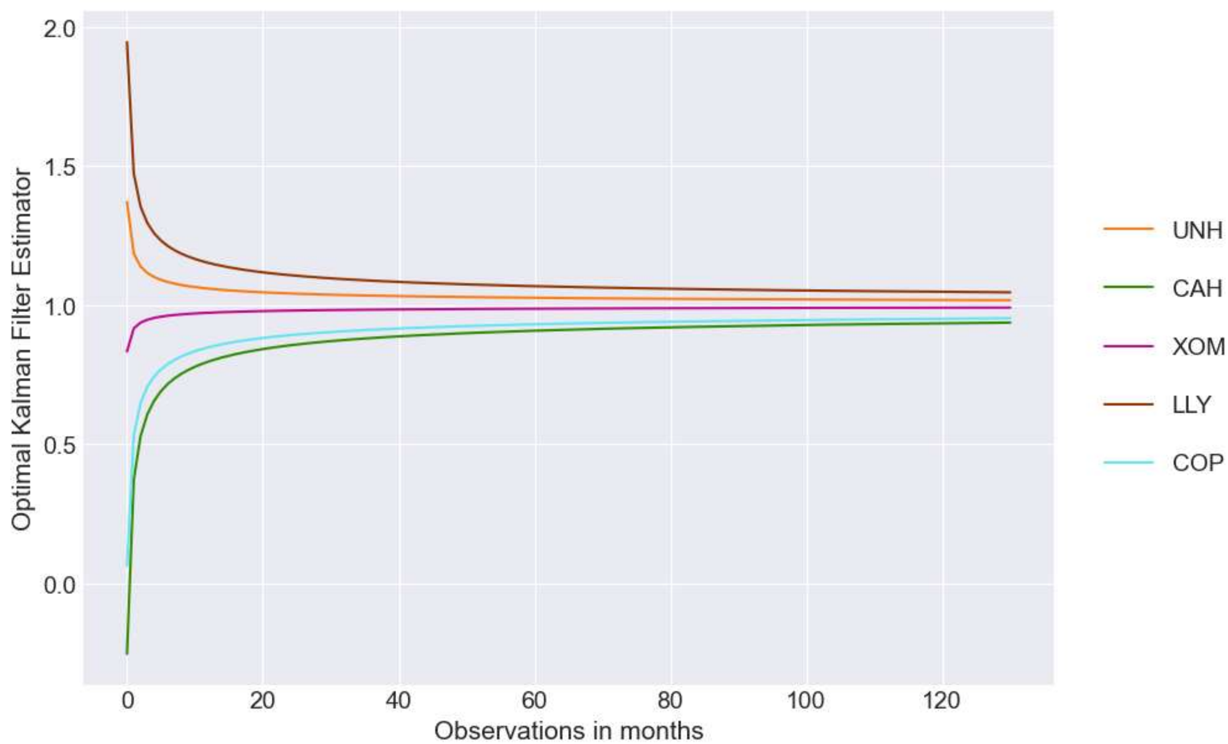
```
MSR portfolio weights for 60-month initial window
             UNH       CAH       XOM       LLY       COP
Weight  0.227005 -0.006267  0.386796  0.510264 -0.117798
```

```
MSR portfolio allocation in the rolling factor model results:::


                         Weight of UNH  Weight of CAH  Weight of XOM  \
Rolling fixed 60-months
1                             0.221758       0.012610       0.388806
2                             0.178743       0.018974       0.353774
3                             0.188880      -0.018763       0.315833
4                             0.205065      -0.018582       0.314507
5                             0.173534       0.011194       0.248290
...                                ...            ...            ...
56                           -0.051941       0.613277       0.675651
57                           -0.054886       0.646430       0.774325
58                           -0.058453       0.668107       0.761105
59                           -0.149924       0.543694       0.602418
60                           -0.172792       0.573848       0.586015

                         Weight of LLY  Weight of COP
Rolling fixed 60-months
1                             0.461889      -0.085063
2                             0.556102      -0.107594
3                             0.523132      -0.009082
4                             0.545232      -0.046222
5                             0.719000      -0.152018
...                                ...            ...
56                            0.767963      -1.004950
57                            0.730409      -1.096278
58                            0.765051      -1.135810
59                            0.683820      -0.680008
60                            0.695300      -0.682370

[60 rows x 5 columns]


Out-of-sample utility for the for a mean variance investor
      in the rolling factor model results is : 0.0062
```

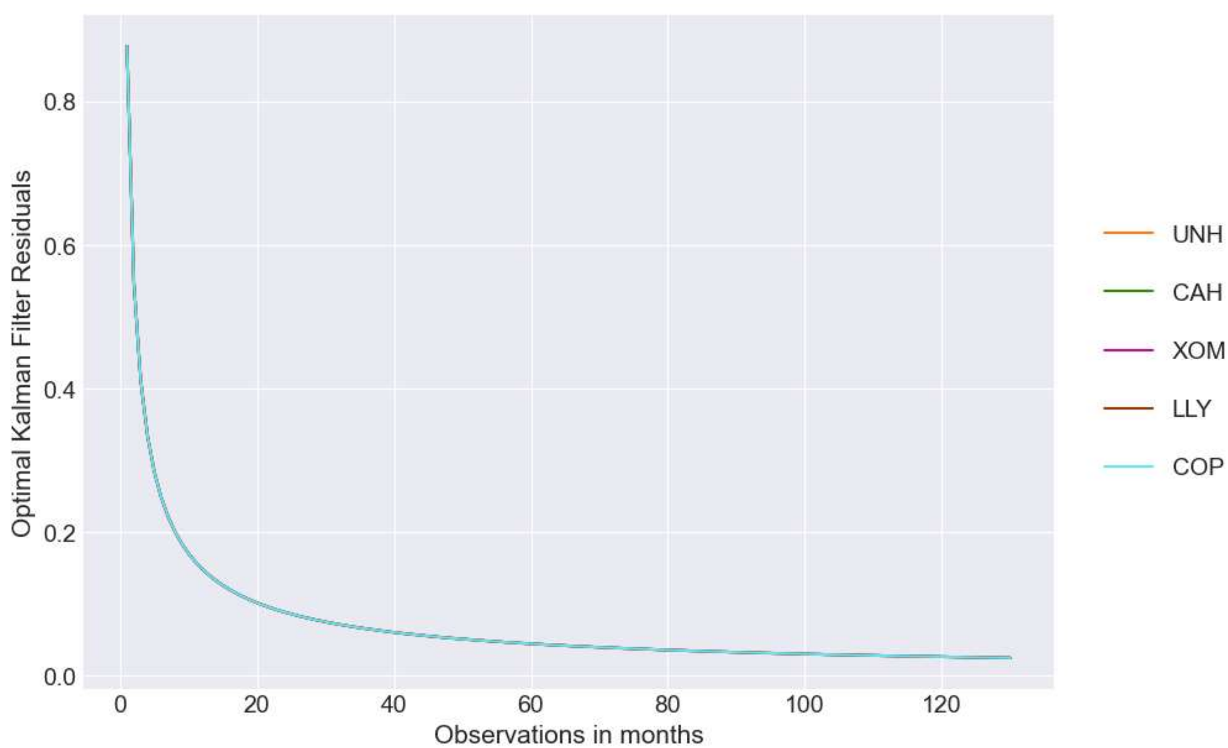
Again, for the cumulative fixed 60-month window, we can use the cumulative rolling factor model results to allocate stocks into the maximum Sharpe ratio portfolio and get the out-of-sample utility for a mean variance investor with a risk aversion coefficient of 4.

```
MSR portfolio allocation in the cumulative factor model results:::


                             Weight of UNH  Weight of CAH  Weight of XOM  \
Cumulative fixed 60-months
61                                0.224793       0.010866       0.384345
62                                0.164961      -0.008174       0.390353
63                                0.183413      -0.029787       0.372363
64                                0.187250      -0.031487       0.369467
65                                0.182783      -0.018976       0.367800
...                                    ...            ...            ...
127                               0.085431       0.154645       0.476416
128                               0.086258       0.147348       0.472103
129                               0.096742       0.138900       0.474995
130                               0.090614       0.131316       0.479268
131                               0.093238       0.125227       0.470205

                             Weight of LLY  Weight of COP
Cumulative fixed 60-months
61                                0.466114      -0.086117
62                                0.530079      -0.077219
63                                0.540330      -0.066319
64                                0.541292      -0.066522
65                                0.554781      -0.086389
...                                    ...            ...
127                               0.521977      -0.238470
128                               0.513504      -0.219212
129                               0.499475      -0.210113
130                               0.498937      -0.200135
131                               0.484019      -0.172690

[71 rows x 5 columns]


Out-of-sample utility for the for a mean variance investor
      in the cumulative factor model results is : 0.0061
```

Now, we can use the Kalman filter to plot the time series of the factor loading for each of the stocks over the full sample period.



Optimal Kalman Filter Estimator
Observations in months
UNH
CAH
XOM
LLY
COP



Optimal Kalman Filter Residuals
Observations in months
UNH
CAH
XOM
LLY
COP