Black-Litterman Portfolios with Machine Learning derived Views

Preprint · July 2018

DOI: 10.13140/RG.2.2.26727.96160

Pavan Donthireddy
Indian Institute of Technology Madras

Black-Litterman Portfolios with Machine Learning derived Views

Pavan Donthireddy

July 19, 2018

Abstract

In this report, we implement various Black Litterman models on a multi asset ETF portfolio and test their out of sample performance, where 1/N portfolio and market capitalization weights based portfolio are used as benchmarks. Following Meucci(2010) [1], we use qualitative views to set the entries of $\mathbf{Q}_t$ in terms of volatility induced by market. We use ML classifiers to arrive at these qualitative views, and we also compare a number of methods that reduce the estimation error in the covariance matrix. We found out that Black Litterman optimized portfolios significantly outperform the benchmark portfolios in terms of various risk adjusted measures even after controlling different levels of risk aversion.

1 Introduction

The Black-Litterman Model, created by Fischer Black and Robert Litterman, is a sophisticated portfolio construction method that overcomes the problem of unintuitive, highly-concentrated portfolios, input-sensitivity, and estimation error maximization. The Black-Litterman model uses a Bayesian approach to combine the subjective views of an investor regarding the expected returns of one or more assets with the market equilibrium vector of expected returns (the prior distribution) to form a new, mixed estimate of expected returns. The resulting new vector of returns (the posterior distribution), leads to intuitive portfolios with sensible portfolio weights (Idzorek, 2004) [3].

It is an asset allocation model which has its roots in mean-variance optimization model and capital asset pricing model (CAPM). Model builds on mean variance optimization and CAPM by using a Bayesian framework that allows investors to incorporate their views on markets effectively into asset allocation process. Main contribution of Black-Litterman model is that, it enables investors to construct sensible portfolios without using unnecessary constraints which also reflect their views on markets.

It is well known to both academics and practitioners that standard mean variance optimization is very sensitive to expected returns and often generates extreme portfolios (concentration in very few assets, large long and short positions). Black-Litterman model overcomes these issues by choosing a neutral reference point, CAPM equilibrium. It also allows investors to express their views with varying confidence levels and integrate these views into CAPM prior by using a Bayesian framework.

In this report we use predicted views from machine learning classifiers as input to all the dynamic Black-Litterman portfolios. We evaluate the out-of-sample performance of the following dynamic Black-Litterman portfolios for different levels of risk aversion and various estimates of Covariance matrix: the implied BL portfolio with reverse optimization; the SR-BL portfolio with maximal Sharpe ratio; the MVaR-BL portfolio with maximal reward-to-VaR ratio; the MCVaR-BL portfolio with maximal reward-to-CVaR ratio.

The outline of the remainder of the report is as follows. Section 2 describes the data used in the empirical results. Section 3 describes the Black-Litterman allocation framework. Section 4 reports and discusses the empirical results. Section 5 summarizes our findings and offers some concluding remarks.

2 Data

To construct multi-asset portfolios we include highly liquid ETFs which have a wide coverage of the broader asset classes and constitute a globally diversified multi-asset portfolio, representing MSCI-EAFE and MSCI-Emerging markets, U.S. Treasuries (long term), Gold bullion and the U.S. real estate sector. We use the iShares MSCI EAFE ETF (EFA) which tracks MSCI EAFE index, and iShares MSCI Emerging Markets ETF (EEM) which tracks an index composed of large and mid cap emerging markets to cover both developed and emerging markets. To represent US government bond investments we rely on iShares 20+ Year Treasury Bond ETF (TLT). Also we use SPDR Gold Shares (GLD) to represent the gold market, and iShares US Real Estate ETF (IYR) to represent the US Real Estate Market.

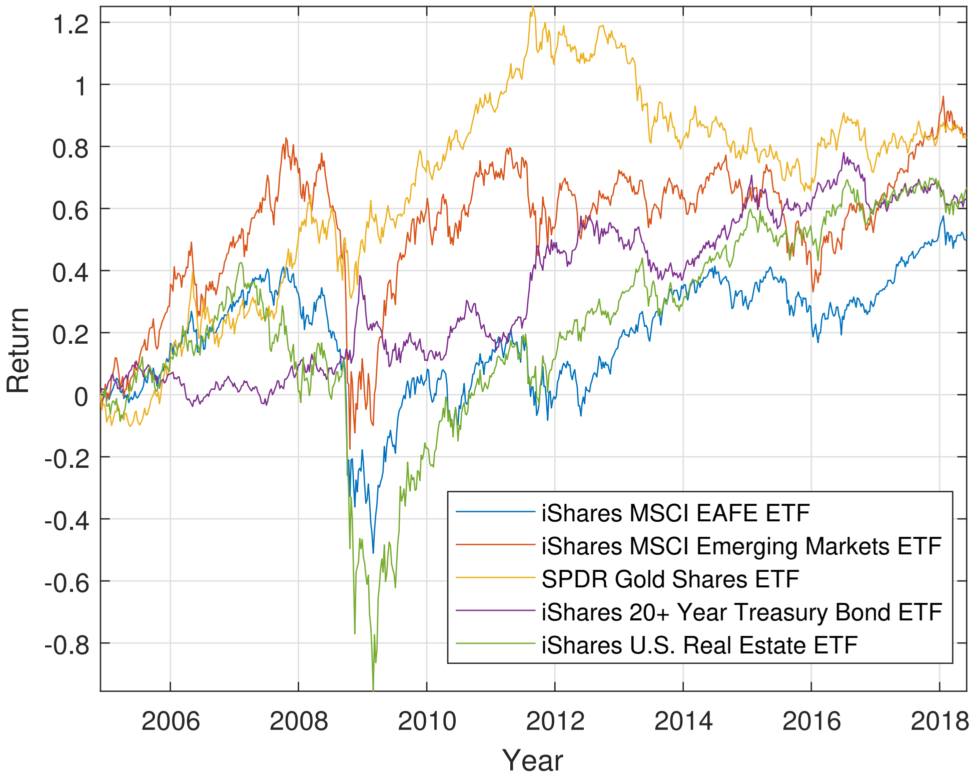


Figure 1: Cumulative excess return of the assets for the period December 2004 - June 2018.

The ETFs selected for this study are relatively recent but have continuous daily pricing history available since December 2004. We obtain weekly data for the period from December 2004 to June 2018 from Yahoo Finance. This period includes one full bear market (2007-2009) and two bull markets (2004-2007 and 2009-May 2013), thus insulating the study from explicit market-phase bias. All data are denominated in US dollar. The Market Capitalization of these ETFs are obtained from ETF Database. We use the 13-week treasury bill (ˆIRX) historical quotes from Yahoo Finance to calculate risk-free rate. Since the T-bills rates are annualized we convert the annualized rate to weekly risk free rate. We use the price data to compute returns and then subtract the calculated weekly risk free rate to get excess returns, which are used throughout the empirical analysis.

Table 1 reports the summary statistics of excess returns for each ETF from December 2004 to June 2018. The null hypothesis of normality is strongly rejected in all cases.

	EFA	EEM	GLD	TLT	IYR
Mean return p.a (%)	3.71	6.20	6.12	4.64	4.92
SD p.a (%)	20.12	25.83	18.11	13.17	25.61
Median	0.0022	0.0022	0.0032	0.0019	0.0042
Min	-0.23	-0.22	-0.10	-0.07	-0.19
Max	0.12	0.25	0.13	0.08	0.20
IQR	0.03	0.04	0.03	0.02	0.03
Skewness	-1.08	-0.36	-0.31	-0.03	-0.28
Kurtosis	10.84	10.26	4.72	4.69	10.05
Hyperflatness	477.91	324.61	52.46	48.74	210.68
Standardized moment of order 8	29934.26	13176.53	921.49	728.68	5238.24
JB-stat	1942.6*	1563.7*	97.8*	83.6*	1469.2*
MDD	0.92	1.00	0.59	0.29	1.38
MDD Start Date	08-10-2007	22-10-2007	29-08-2011	15-12-2008	05-02-2007
MDD End Date	02-03-2009	20-10-2008	23-11-2015	01-06-2009	02-03-2009
Sharpe Ratio	0.18	0.24	0.34	0.35	0.19

Notes: This table provides various descriptive statistics of the assets considered in the empirical analysis. The period covers the months from December 2004 to June 2018. 'Mean return p.a.' denotes annualized time-series mean of monthly returns, while 'SD p.a.' denotes the associated annualized standard deviation. IQR represent the interquartile range of the excess returns for the whole sample. 'Skewness', 'Kurtosis', 'Hyperflatness' represent the third, fourth and sixth standardized moments of the return distribution. 'Sharpe ratio' shows the annualized Sharpe ratios of the respective ETFs. MDD shows the maximum drawdown of the respective ETF during the period from December 2004 to June 2018 and 'JB' is the JarqueBera statistic for testing normality of returns.

*Statistically significant at the 1% level.

Table 1: Descriptive Statistics of Excess Returns (December 2004 - June 2018)

From Table 1, it can be seen that EEM has the highest mean return of 6.20% and EFA has the lowest with 3.71% for the complete sample. The US Treasuries ETF(long term) generates the highest Sharpe ratio of 0.35. Skewness values indicates that all the ETFs have a distribution with an asymmetric tail extending toward more positive values with highest in the case of EFA and lowest in the case of TLT. From Kurtosis and other standardized moments of even order, it is evident that EFA has fatter tail when compared to other ETFs. The maximum drawdowns (MDD) of the ETF is that the maximum loss an investor could have suffered during the evaluation period. By investing in EFA this amounts 92% of the invested capital. This figure was 100% for EEM, 59% for GLD, 29% for TLT, and 138% for IYR. MDD Start Date and MDD End date gives the period which contributed to the MDD of respective ETF's. Clearly the MDD period of IYR (US Real Estate ETF) falls during the Financial crisis (2007-2009). However this is 2011-2015 in the case of GLD. The JarqueBera statistic is significant for all ETF's. Hence, the assumption of normal distributed returns has to be rejected.

Table 2 presents evidence on potential diversification effects in terms of pair-wise correlation coefficients.

	EFA	EEM	GLD	TLT	IYR
MSCI EAFE ETF	1				
MSCI Emerging Markets ETF	0.875*	1			
Gold ETF	0.154*	0.215*	1		
US 20+year Bond ETF	-0.361*	-0.314*	0.093**	1	
US Real Estate ETF	0.664*	0.665*	0.074**	-0.161*	1

Notes: This table provides the correlation matrix of the weekly excess returns considered in the analysis over the period from December 2004 to June 2018.

*Statistically significant at the 1% level.

**Statistically significant at the 5% level.

Table 2: Correlation matrix of excess returns(December 2004-June 2018)

The correlation between the EFA and EEM is highly significant and larger than 0.8, indicating a strong co-movement of developed and emerging stock markets. While the Gold ETF and US Real Estate ETF offer a slightly better diversification effect with correlation coefficients ranging between 0.07 and 0.66, the highest diversification potential during our sample period is provided by investing in the US Long term Bond ETF, which is reflected in negative correlation coefficients and positive Sharpe ratios. Consequently, we expect to find significant portfolio benefits by applying the Black Litterman frameworks on a multi-asset portfolio, including real estate, bonds and gold, rather than on a stock only portfolio.

Table 3 reports the time series properties of excess returns for each ETF from December 2004 to June 2018. In particular, it shows the first five autocorrelation coefficients and the value of the Ljung-Box test for serial correlation of 5, 10 and 15 lags, the ARCH test of Engle. Clearly in many cases excess returns does not display significant autocorrelations. The results of the ARCH test suggest that there is significant volatility clustering for most of the excess return series.

	ACF(1)	ACF(2)	ACF(3)	ACF(4)	ACF(5)	Engle's ARCH	LB-Q(5)	LB-Q(10)	LB-Q(15)
EFA	-0.059	0.022	-0.033	-0.003	0.030	52.894*	4.189	17.495	35.017*
EEM	-0.068	0.047	-0.041	0.009	0.007	88.870*	6.171	17.440	27.104**
GLD	0.017	-0.035	-0.033	0.001	-0.062	11.255*	4.523	9.060	17.480
TLT	-0.103	0.082	0.008	0.068	-0.074	1.903**	19.443*	23.396*	25.998**
IYR	-0.066	-0.001	0.086	-0.074	0.014	198.737*	12.435**	21.434**	38.162*

Notes: This table reports the test statistics for autocorrelation, Engle's ARCH test and Ljung Box Q test for the full sample from December 2004 to June 2018. The Ljung-Box Q test statistic for autocorrelation of orders 5, 10 and 15.
*Statistically significant at the 1% level.
**Statistically significant at the 5% level.

Table 3: Time Series properties of excess returns

3 Theoretical Framework

We consider a market of N assets and define the first and second moments of N asset excess returns, conditional on the information set $\mathbf{Y}$, as follows:

$$\mathbf{r}_t = \boldsymbol{\mu}_{\mathrm{BL},t} + \boldsymbol{\epsilon}_t \tag{1}$$

$$\boldsymbol{\epsilon}_t | \mathbf{Y}_{t-1} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_t) \tag{2}$$

$$\boldsymbol{\mu}_{\mathrm{BL},t} \sim \mathcal{N}(\boldsymbol{\pi}_t, \tau \boldsymbol{\Sigma}_{\mathrm{BL},t}) \tag{3}$$

where, $\mathbf{r}_t = (r_{i,j})$ is the $N \times T$ excess return matrix, $\boldsymbol{\mu}_{\mathrm{BL},t}$ is the $N \times 1$ expected excess returns in period t, $\boldsymbol{\epsilon}_t$ is the $N \times 1$ error term vector, $\boldsymbol{\Sigma}_t$ is the $N \times N$ covariance matrix of excess returns, $\boldsymbol{\pi}_t$ is the $N \times 1$ conditional equilibrium return vector of the market portfolio, and the scale parameter τ indicates the uncertainty of the CAPM prior. To calibrate the overall uncertainty level τ in the equation 3 we follow Meucci(2010) [1] and set:

$$\tau \approx \frac{1}{T} \tag{4}$$

3.1 Estimation of time varying Covariance Matrix

There are many methods to forecast Covariance Matrix. In this section we present four methods considered in this study to estimate the time varying Covariance matrix. The methods are sample covariance, Fast minimum covariance determinant estimate, Shrinkage estimate and EWMA Covariance matrix.

3.1.1 Sample Covariance:

The sample covariance method uses all the data available. At time t, with a set of data from 1 to $t-1$, the sample covariance matrix is given by

$$\boldsymbol{\Sigma}_t = \frac{1}{t-2} \sum_{j=1}^{t-1} (\boldsymbol{r}_{t-j} - \bar{\boldsymbol{r}})(\boldsymbol{r}_{t-j} - \bar{\boldsymbol{r}})' \tag{5}$$

where $\bar{\boldsymbol{r}} = \frac{1}{t-1} \sum_{t=1}^{t-1} \boldsymbol{r}_{t-j}$ is the $N \times 1$ vector of the sample mean excess returns.

3.1.2 Fast Minimum Covariance determinant estimate

Minimum covariance determinant (MCD) is the fastest estimator of multivariate location and scatter that is both consistent and robust. However, an exact evaluation of the MCD is impractical because it is computationally expensive to evaluate all possible subsets of the sample data.

The FAST-MCD method selects h observations out of n (where $\frac{n}{2} < h \leq n$) whose classical covariance matrix has the lowest possible determinant.

The MCD covariance is the covariance matrix of the h selected points, multiplied by a consistency factor to obtain consistency at the multivariate normal distribution, and by a correction factor to correct for bias at small sample sizes.

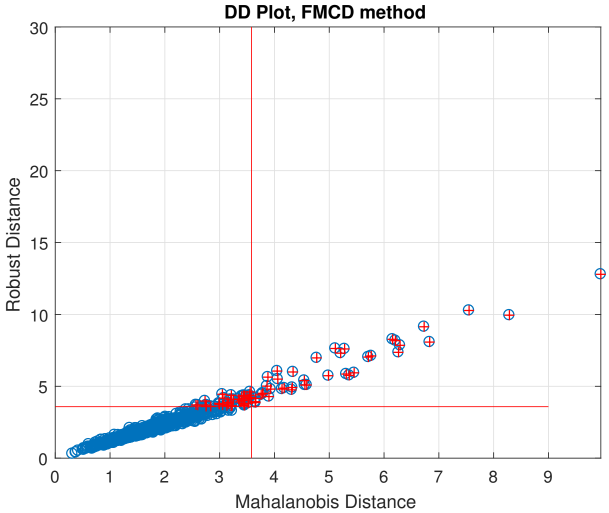
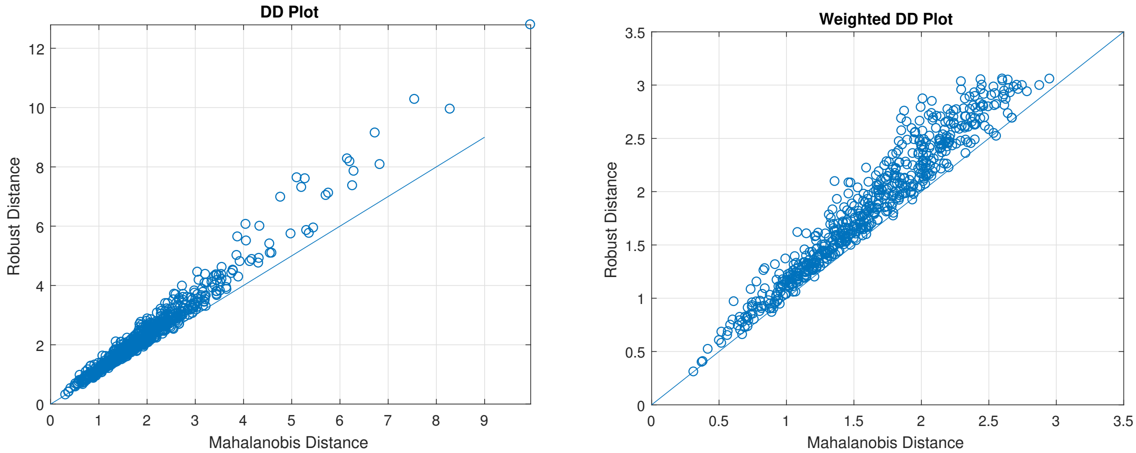


Figure 2: DD plot of full sample indicating outliers.

In a DD plot, the data points tend to cluster in a straight line that passes through the origin. Points that are far removed from this line are generally considered outliers. In figure 2, the red '+'symbol indicates the data points that are considered to be outliers.



(a) DD plot of full sample (b) Weighted DD plot of full sample

Figure 3: Application of Fast MCD on full sample(December 2004-June 2018)

For data with an elliptically-contoured distribution (as shown in the upper left), the plotted points follow a straight line, but are not at a 45-degree angle to the origin. We use a weighted DD plot to magnify the lower corner and reveal features that are obscured when large robust distances exist.

3.1.3 Shrinkage estimate:

In the case of the constant correlation model [8], the shrinkage estimator for the covariance matrix takes the form

$$\mathbf{\Sigma}_{\mathrm{LW},t} = w\mathbf{\Sigma}_{\mathrm{CC},t} + (1-w)\mathbf{\Sigma}_t \tag{6}$$

where $\mathbf{\Sigma}_t$ is the sample covariance matrix as defined in equation 5 , and $\mathbf{\Sigma}_{\mathrm{CC},t}$ is the sample covariance matrix with constant correlation. The Constant Correlation matrix is computed as follows: First we decompose, the sample covariance matrix according to

$$\mathbf{\Sigma}_t = \mathbf{\Lambda}_t \boldsymbol{C}_t \mathbf{\Lambda}_t' \tag{7}$$

where $\boldsymbol{\Lambda}_t$ is a diagonal matrix of the variances of excess returns and $\boldsymbol{C}_t$ is the sample correlation matrix.

Next we replace the sample correlation matrix with constant correlation matrix

$$\boldsymbol{C}_{\mathrm{CC},t} = \begin{bmatrix} 1 & \hat{\rho}_t & \hat{\rho}_t & \dots & \hat{\rho}_t \\ \hat{\rho}_t & 1 & \hat{\rho}_t & \dots & \hat{\rho}_t \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{\rho}_t & \hat{\rho}_t & \hat{\rho}_t & \dots & 1 \end{bmatrix} \tag{8}$$

where $\hat{\rho}_t$ is the average of all sample correlations, in other words

$$\hat{\rho}_t = \frac{2}{N(N-1)} \sum_{i=1}^{N} \sum_{j=i+1}^{N} \rho_{i,j,t} \tag{9}$$

The optimal shrinkage intensity can be shown to be proportional to a constant divided by the length of the history, T.

3.1.4 EWMA Covariance matrix

This approach to forecasting the covariance-matrix is popularized by the RiskMetrics group [15]. In this approach, the exponentially weighted covariance-matrix is estimated using the following recursive form:

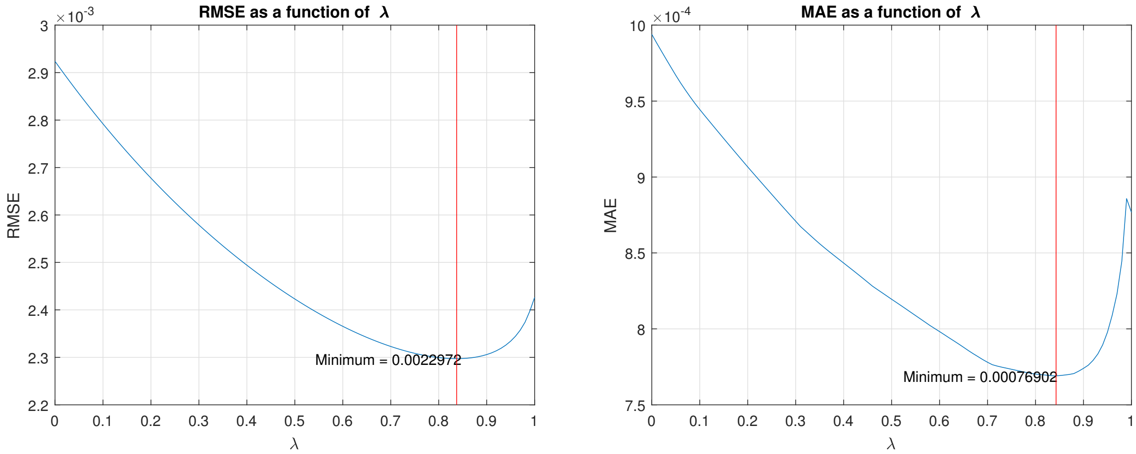
$$\boldsymbol{\Sigma}_t = \lambda \boldsymbol{\Sigma}_{t-1} + (1-\lambda)(\boldsymbol{r}_{t-1} - \bar{\boldsymbol{r}}_{t-1})(\boldsymbol{r}_{t-1} - \bar{\boldsymbol{r}}_{t-1})' \tag{10}$$

where λ is the decay factor which determines how rapidly the weights on past observations decline and typically estimated to be between 0.92 and 0.97. In RiskMetrics (J.P. Morgan, 1994), the decay factor is set to 0.94. The EWMA model is a special case of the IGARCH(1,1) model where volatility innovations have infinite persistence. The assumption that volatility innovations infinitely persist through time may appear theoretically tenuous, however it appears to be a reasonable assumption for short term volatility forecasting.

Since there is no generally accepted statistical procedure for choosing λ, we use the average optimal value of λ for which RMSE and MAE of individual variances are minimized.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (\sigma_t^2 - \hat{\sigma}_t^2)^2} \tag{11}$$

$$\mathrm{MAE} = \frac{1}{n} \sum_{t=1}^{n} \left| (\sigma_t - \hat{\sigma}_t^2) \right| \tag{12}$$



(a) label 1 (b) label 2

Figure 4: RMSE and MAE as a function of λ on full sample(Decmber 2004-June 2018) for EFA

	RMSE		MAE	
	Optimal λ	Minimum	Optimal λ	Minimum
EFA	0.8374	0.0023	0.8429	0.0008
EEM	0.7341	0.0036	0.8804	0.0013
GLD	0.8954	0.0012	0.8594	0.0006
TLT	0.9193	0.0006	0.9251	0.0003
IYR	0.6479	0.0032	0.6309	0.0012
Average	0.80682		0.82774	

Table 4: Optimal λ for individual ETFs on the whole sample (December 2004 - June 2018)

where, σ^2 is the realized variance and $\hat{\sigma}_t^2$ is the forecast from the EWMA model. Clearly, a better alternative would be to minimize the norm($||.||_2$) of the error matrix(difference between estimated EWMA covariance matrix and realized covariance matrix) for the whole sample.

3.2 Conditional Equilibrium returns

Let $\boldsymbol{\pi}_t$ be the $N \times 1$ conditional mean vector and let $\boldsymbol{\Sigma}_t$ be the $N \times N$ covariance matrix estimate of these returns given information available at time $t-1$. In addition, define $\boldsymbol{w}_{t-1}$ to be the vector of market capitalization weights at time $t-1$. When the CAPM holds, the conditional mean vector satisfies the following equation:

$$\boldsymbol{\pi}_t = \boldsymbol{\delta} \boldsymbol{\Sigma}_t \boldsymbol{w}_{t-1} \tag{13}$$

where δ is the risk aversion coefficient.

3.3 The Investor's Views

An investor might possess views about some or all of the returns of the assets in a portfolio, which may differ from the above implied equilibrium returns. The uncertainty of the views is given by the error vector e_t with a mean of zero and covariance matrix $\boldsymbol{\Omega}_t$. The error terms are unknown and independent. The investors views at time t, can thus be expressed as:

$$\boldsymbol{Q}_t = \boldsymbol{P}_t \boldsymbol{\mu}_{\mathrm{BL},t} + \boldsymbol{e}_t \tag{14}$$

At time t, let $K \leq N$ be the total number of the views, $\boldsymbol{P}_t$ be the $K \times N$ matrix of view portfolios and $\boldsymbol{Q}_t$ be the vector of expected returns on the view portfolios.

We use various classification algorithms to impose views. The methodology for imposing qualitative views is as follows:

We set entries of $\boldsymbol{Q}_t$ in terms of the volatility imposed by the market.

$$\boldsymbol{Q}_t(k) = (\boldsymbol{P}_t\boldsymbol{\pi}_t)(k) + \eta_k\sqrt{(\boldsymbol{P}_t\boldsymbol{\Sigma}_t\boldsymbol{P}_t')(k,k)} \qquad k = 1, 2, \cdots, K \tag{15}$$

where, $\eta_k \in \{-2, -1, 1, 2\}$ defines "Very Bearish","Bearish","Bullish" and "Very Bullish" views respectively.

To impose views we use two classifiers, one to predict the sign of η_k and the other to predict its absolute value. i.e, we use the following two binary level responses to train various classifiers and predict the qualitative view of next period.

$$Y_1 = \begin{cases} -1, & \text{if sign of excess return of the next week's is negative} \\ +1, & \text{if sign of excess return of the next week's is positive} \end{cases}$$

$$Y_2 = \begin{cases} 1, & \text{if } z_t = \dfrac{r_t - \bar{r}_{t,3}}{\sigma_{t,3}} \leq 1 \\ 2, & \text{if } z_t = \dfrac{r_t - \bar{r}_{t,3}}{\sigma_{t,3}} > 1 \end{cases}$$

We use the above two responses to identify the view of next week's excess return.i.e.,

$$Y = \left\{ \begin{array}{lr} -2, & \text{Very Bearish} \\ -1, & \text{Bearish} \\ 1, & \text{Bullish} \\ 2, & \text{Very Bullish} \end{array} \right\} = Y_1 Y_2$$

3.3.1 Features

For each period and each ETF, we examine the combination of following information sets as features for both classification problems. Note that for any given ETF's change in return we allow for it's own information as well as the other assets information to influence the sign and magnitude of the return over the given horizon. We define our information sets as follows:

- Information set 1: Previous n week returns and their j lags, where $n \in \{1, 2, 3, 4\}$ and $j \in \{0, 1, 2, 3, 4, 5\}$ for respective ETFs along with S&P 500, VIX and 10 Year treasury note yield index.

- Information set 2: Previous n week excess returns and their j lags, where $n \in \{1, 2, 3, 4\}$ and $j \in \{0, 1, 2, 3, 4, 5\}$ for respective ETFs along with S&P 500, VIX and 10 Year treasury note yield index.

- Information set 3: Rate of change of volume in past n weeks and their j lags, where $n \in \{1, 2, 3, 4\}$ and $j \in \{0, 1, 2, 3, 4, 5\}$ for respective ETFs.

- Information set 4: Momentum related technical indicators like CCI, RSI, ROC, Slow/Fast Stochastic oscillator,William Indicator, Aroon indicator and true strength index up to 6 periods and their five lags values on all ETFs and S&P 500, 3 month T-bill rate, VIX and 10 Year treasury note yield index.

- Information set 5: Trend related technical indicators like SMA, EMA, MACD, ADX and T3 up to 6 periods and their five lags values on all ETFs and S&P 500, 3 month T-bill rate, VIX and 10 Year treasury note yield index.

- Information set 6: Volume based technical indicators like OBV, Money Flow Index, CMF and Force indicator up to 6 periods and their five lags values on all ETFs and S&P 500, 3 month T-bill rate, VIX and 10 Year treasury note yield index.
- Information set 7: Volatility based indicators like Bollinger Bands, Average True Range, VR and HHLL up to 6 periods and their five lags values on all ETFs and S&P 500, 3 month T-bill rate, VIX and 10 Year treasury note yield index.

First we remove all the columns which have more than 10 NaN and instead of deleting all the rows with missing values, we impute them based on kNN. The k nearest neighbors algorithm can be used for imputing missing data by finding the k closest neighbors to the observation with missing data and then imputing them based on the the non-missing values in the neighbors.

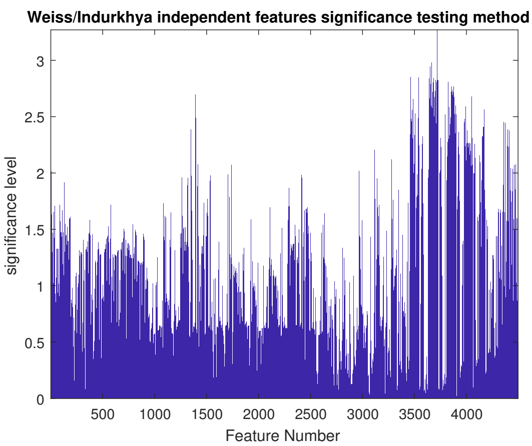
3.3.2 Feature Selection

After imputing the missing values, we use the Weiss/Indurkhya 'independent features'significance testing method to select independent features. The below two equations summarize the test for a classification problem, where A and B are the same feature measured for class 1 and class 2 respectively, and n_1 and n_2 are the respective numbers of cases.

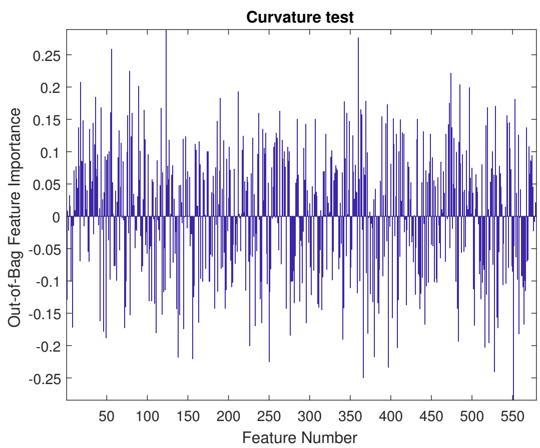
$$S.E(A-B) = \sqrt{\frac{var(A)}{n_1} + \frac{var(B)}{n_2}} \tag{16}$$

$$\frac{|\mu_A - \mu_B|}{S.E(A-B)} > c \tag{17}$$

We choose $c = 1.5$ and eliminate all features which fail this test. Once all the unwanted features are removed, we use bagging to rank variable importance. Although the collection of bagged trees is much more difficult to interpret than a single tree, one can obtain an obtain an overall summary of the importance of each predictor using the Gini index from bagging classification trees. In the context of classification, we can add up the total amount that the Gini index is decreased by splits over a given predictor, averaged over all B trees. In order to avoid over-fitting we choose top 10 variables with the largest mean decrease in Gini Index.

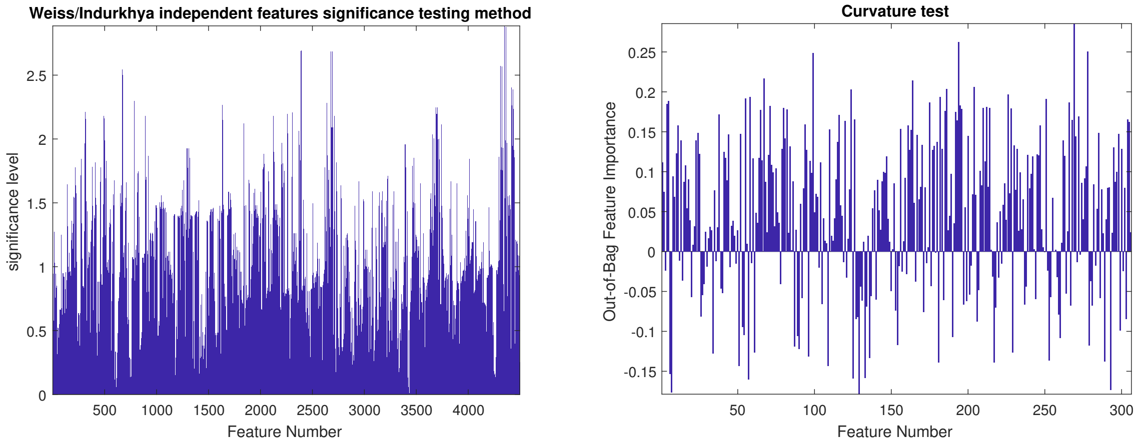


(a) Weiss/Indurkhta independent features significance testing method



(b) Curvature test based on Bagging

Figure 5: Features selection for Classifier1 on EFA



(a) Weiss/Indurkhta independent features significance testing method

(b) Curvature test based on Bagging

Figure 6: Features selection for Classifier2 on EFA

3.3.3 Classifiers

We explain the details of the classifiers used in this subsection.

- Logistic Regression : The first classier we used was Logistic regression. In our study we used LR to evaluate the relationship between binary class labels (-1/+1 and 1/2), and the final 10 features derived for each classification problem. The Logistic regression model gave us the probability of the following week being decided as +1 or -1, and 1 or 2 in respective problems. In our case, we determined the threshold value as 0.5 and we assigned the class label if the probability exceeded the threshold. Logistic regression estimates the probability of the output as follows:

$$p(y|x_1, x_2, \cdots x_{10}) = \frac{e^{w_0+w_1x_1+w_2x_2+\cdots+w_{10}x_{10}}}{1+e^{w_0+w_1x_1+w_2x_2+\cdots+w_{10}x_{10}}} \tag{18}$$

  In this equation, y is the class label and $p(y|x_1, x_2, \cdots x_{10})$ is the prediction probability class label based on 10 final features $(x_1, x_2, \cdots, x_{10})$. The maximum likelihood method is used to optimize the model parameters $(w_0, w_1, \cdots, w_{10})$

- Support Vector Machine: An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall. More formally, a support vector machine constructs a hyperplane or set of hyperplanes in a high- or infinite-dimensional space, which can be used for classification. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training-data point of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier.

- Näive Bayes Classifier: Abstractly, Näive Bayes is a conditional probability model: given a problem instance to be classified, represented by a vector $\mathbf{x} = (x_1, \ldots, x_{10})$ representing 10 features (independent variables), it assigns to this instance probabilities

$$p(C_k \mid x_1, \ldots, x_{10})$$

for each of 2 possible outcomes. since Näive Bayes classifiers assume that the value of a particular feature is independent of the value of any other feature, given the class variable this becomes

$$p(C_k \mid x_1, \ldots, x_{10}) \propto p(C_k) \prod_{i=1}^{10} p(x_i|C_k)$$

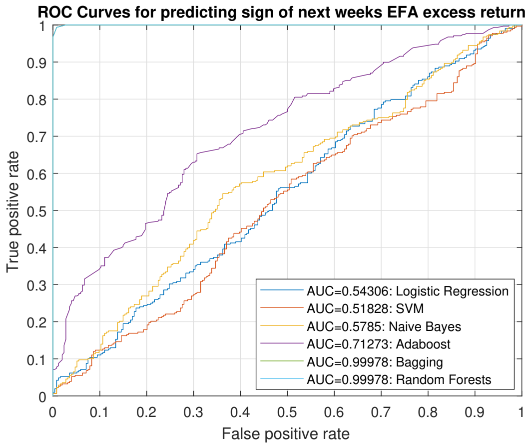
Now the Bayes classifier, assigns a class label $\hat{Y} = C_k$ for some k as follows:

$$\hat{Y} = \underset{k\in\{1,\ldots,K\}}{\operatorname{argmax}} \; p(C_k) \prod_{i=1}^{10} p(x_i \mid C_k).$$

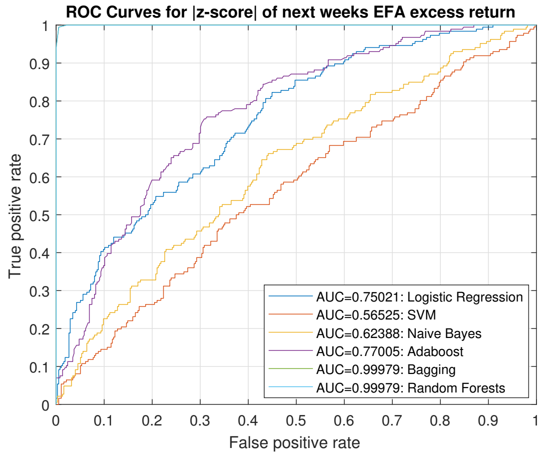
- Bagging, Random Forests and Boosting: Bagging, Random Forests and Boosting uses trees as building blocks to construct more powerful predictive models. The decision trees usually suffer from high variance. This means that if we split the training data into two parts at random, and fit a decision tree to both halves, the results we get could be quite different.

  - Bootstrap aggregation or Bagging is a general purpose procedure for reducing the variance of the statistical model. In this approach we generate B different bootstrapped training data sets. We then train the method on the b^{th} bootstrapped training set in order to get $\hat{Y}_b$, and finally take a majority vote: the overall prediction is the most commonly occurring class among B predictions.
  - Random Forests provide an improvement over bagged trees by way of a small tweak that decorrelates trees. As in bagging, we build a number of trees on bootstrapped training samples. But when building these decision trees, each time a split in a tree is considered, a random sample of m predictors is chosen as split candidates from the full set of p predictors. The split is allowed to use only one of those m predictors. A fresh sample of m predictors is taken at each split, and typically we choose $m \approx \sqrt{p}$.
  - Boosting works in similar way to that of bagging, except that the trees are grown sequentially: each tree is grown using information from previously grown trees. Boosting does not involve bootstrap sampling; instead each tree is fit on a modified version of the original dataset.

    Figures 7-11 shows the training dataset (80% of full sample) ROC curves for all the classifiers considered for both the problems. From Table 5 It can be seen that Random Forecast classifier significantly outperforms the other classifiers considered.

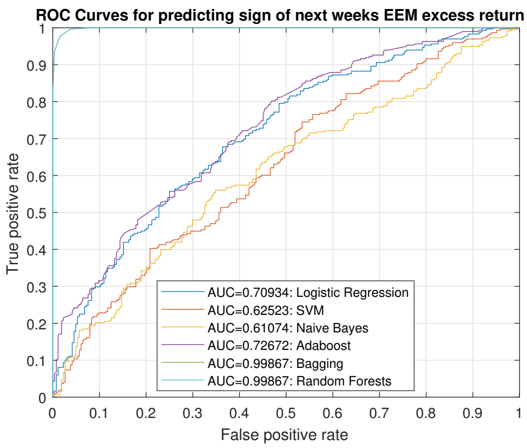


(a) ROC curves for Classification problem 1

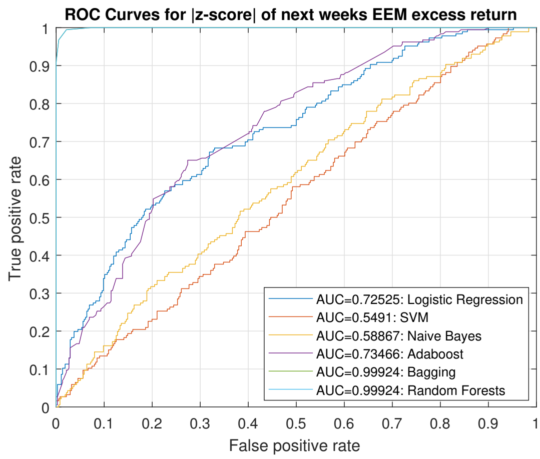


(b) ROC curves for Classification problem 2

Figure 7: ROC curves related to EFA

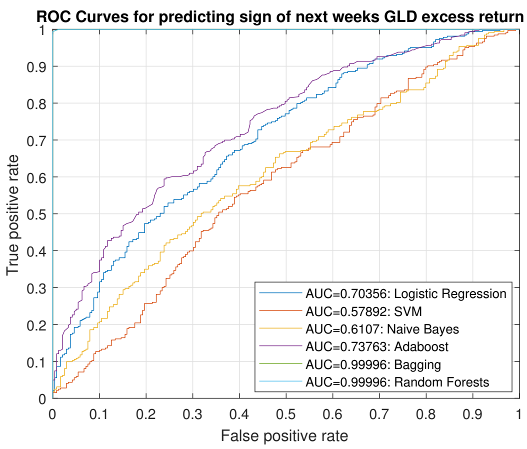


(a) ROC curves for Classification problem 1

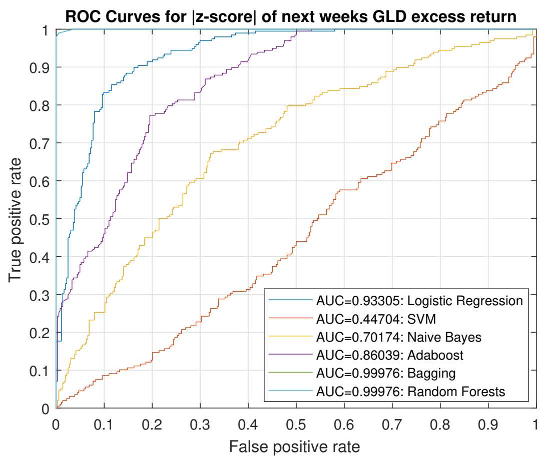


(b) ROC curves for Classification problem 2

Figure 8: ROC curves related to EEM

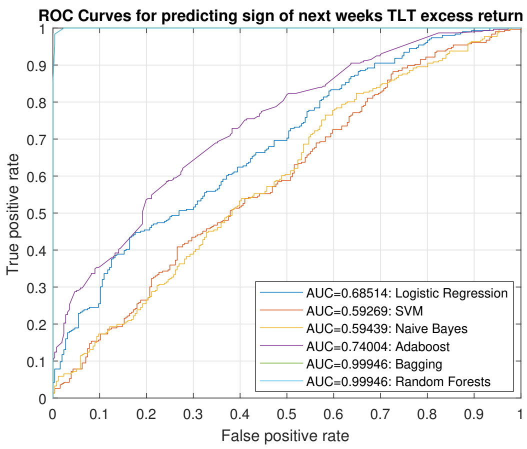


(a) ROC curves for Classification problem 1

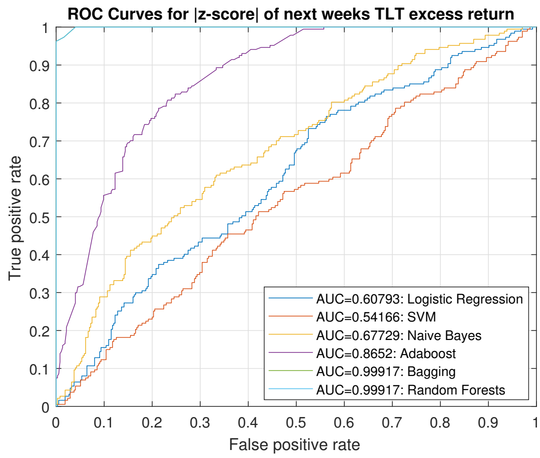


(b) ROC curves for Classification problem 2

Figure 9: ROC curves related to GLD

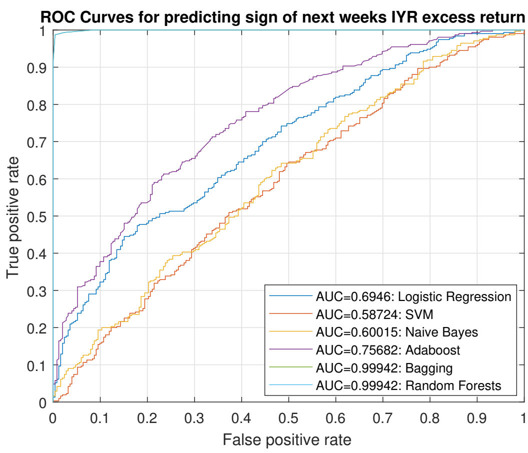


(a) ROC curves for Classification problem 1

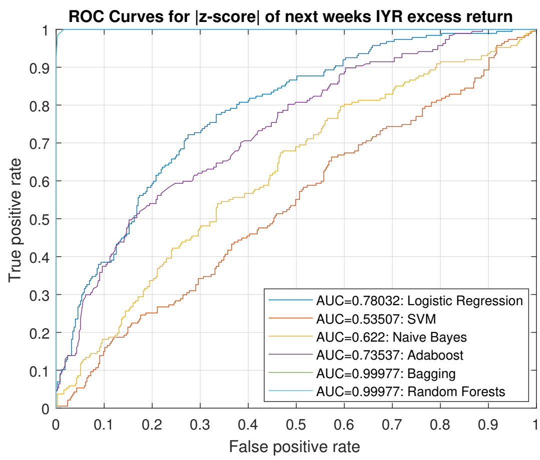


(b) ROC curves for Classification problem 2

Figure 10: ROC curves related to TLT



(a) ROC curves for Classification problem 1



(b) ROC curves for Classification problem 2

Figure 11: ROC curves related to IYR

ETF	Algorithm	Classification problem 1					Classification problem 2				
		TP	FP	FN	TN	Accuracy	TP	FP	FN	TN	Accuracy
EFA	Logistic Regression	35	29	35	42	54.61%	78	21	27	15	65.96%
	SVM	0	64	0	77	54.61%	99	0	42	0	70.21%
	Naive Bayes	47	17	41	36	58.87%	74	25	26	16	63.83%
	Adaboost	30	34	39	38	48.23%	94	5	40	2	68.09%
	Bagging	44	20	44	33	54.61%	90	9	37	5	67.38%
	Random Forests	37	27	35	42	56.03%	91	8	39	3	66.67%
EEM	Logistic Regression	32	30	43	36	48.23%	75	21	39	6	57.45%
	SVM	19	43	19	60	56.03%	96	0	45	0	68.09%
	Naive Bayes	26	36	41	38	45.39%	66	30	25	20	60.99%
	Adaboost	14	48	17	62	53.90%	81	15	40	5	60.99%
	Bagging	26	36	23	56	58.16%	84	12	38	7	64.54%
	Random Forests	26	36	36	43	48.94%	80	16	41	4	59.57%
GLD	Logistic Regression	27	46	30	38	46.10%	73	19	17	32	74.47%
	SVM	1	72	0	68	48.94%	92	0	49	0	65.25%
	Naive Bayes	36	37	36	32	48.23%	52	40	20	29	57.45%
	Adaboost	42	31	36	32	52.48%	72	20	20	29	71.63%
	Bagging	39	34	40	28	47.52%	77	15	19	30	75.89%
	Random Forests	39	34	28	40	56.03%	75	17	19	30	74.47%
TLT	Logistic Regression	28	40	34	39	47.52%	81	10	38	12	65.96%
	SVM	29	39	28	45	52.48%	91	0	50	0	64.54%
	Naive Bayes	43	25	43	30	51.77%	46	45	17	33	56.03%
	Adaboost	24	44	20	53	54.61%	66	25	20	30	68.09%
	Bagging	29	39	27	46	53.19%	81	10	22	28	77.30%
	Random Forests	30	38	29	44	52.48%	75	16	17	33	76.60%
IYR	Logistic Regression	9	52	8	72	57.45%	80	14	40	7	61.70%
	SVM	2	59	4	76	55.32%	94	0	47	0	66.67%
	Naive Bayes	24	37	19	61	60.28%	90	4	44	3	65.96%
	Adaboost	20	41	17	63	58.87%	85	9	45	2	61.70%
	Bagging	26	35	43	37	44.68%	81	13	42	5	60.99%
	Random Forests	26	35	28	52	55.32%	78	16	41	6	59.57%

Table 5: Out of sample performance of all the classifiers

3.4 Combining Conditional Equilibrium returns and Views

The next step in the estimation is to combine the conditional equilibrium returns with the views using a Bayesian approach. In the dynamic case, the $N \times 1$ vector of conditional expected returns $\boldsymbol{\mu}_{\mathrm{BL},t}$ at time t is given by:

$$\boldsymbol{\mu}_{\mathrm{BL},t} = \boldsymbol{\pi}_t + \tau \boldsymbol{\Sigma}_t \boldsymbol{P}_t' (\boldsymbol{P}_t \boldsymbol{\Sigma}_t \boldsymbol{P}_t' \tau + \boldsymbol{\Omega}_t)^{-1} (\boldsymbol{Q}_t - \boldsymbol{P}_t \boldsymbol{\pi}_t) \tag{19}$$

and the estimated covariance matrix is given by

$$\begin{aligned} \boldsymbol{\Sigma}_{\mathrm{BL},t} &= \boldsymbol{\Sigma}_t + ((\tau \boldsymbol{\Sigma}_t)^{-1} + \boldsymbol{P}_t' \boldsymbol{\Omega}_t^{-1} \boldsymbol{P}_t)^{-1} && (20) \\ &= (1+\tau)\boldsymbol{\Sigma}_t - \tau^2 \boldsymbol{\Sigma}_t \boldsymbol{P}_t' (\tau \boldsymbol{P}_t \boldsymbol{\Sigma}_t \boldsymbol{P}_t')^{-1} \boldsymbol{P}_t \boldsymbol{\Sigma}_t && (21) \end{aligned}$$

3.5 Risk Adjusted Black-Litterman Portfolios

We estimate the time-varying expected returns and covariance matrix from Equations 19 and 20. We use reverse optimization to compute the implied weights $\boldsymbol{w}^*_{\mathrm{BL},t}$ at time t, given by:

$$\boldsymbol{w}^*_{\mathrm{BL},t} = \frac{1}{\delta} \boldsymbol{\Sigma}^{-1}_{\mathrm{BL},t} \boldsymbol{\mu}_{\mathrm{BL},t} \tag{22}$$

3.5.1 Sharpe Ratio Maximization

In this method we use mean variance optimization and maximize the Sharpe ratio.

$$\begin{aligned} \max \quad & \frac{\boldsymbol{w}'_{\mathrm{BL},t}\boldsymbol{\mu}_{\mathrm{BL},t}}{\sqrt{\boldsymbol{w}'_{\mathrm{BL},t}\boldsymbol{\Sigma}_{\mathrm{BL},t}\boldsymbol{w}_{\mathrm{BL},t}}} \\ \text{s.t.} \quad & \boldsymbol{w}'_{\mathrm{BL},t}\mathbf{1} = 1 \end{aligned} \tag{23}$$

where $\boldsymbol{\mu}_{\mathrm{BL},t}$ is the expected excess return of the BL portfolio in equation 19, $\sqrt{\boldsymbol{w}'_{\mathrm{BL},t}\boldsymbol{\Sigma}_{\mathrm{BL},t}\boldsymbol{w}_{\mathrm{BL},t}}$ is the conditional portfolio standard deviation, $\boldsymbol{w}_{\mathrm{BL},t}$ is the $N \times 1$ vector of portfolio weights and 1 is an $N \times 1$ vector of ones.

Thus the vector of optimal weights for the SR-BL portfolio is given by

$$\boldsymbol{w}^*_{\mathrm{BL},t} = \frac{\boldsymbol{\Sigma}^{-1}_{\mathrm{BL},t}\boldsymbol{\mu}_{\mathrm{BL},t}}{\mathbf{1}'\boldsymbol{\Sigma}^{-1}_{\mathrm{BL},t}\boldsymbol{\mu}_{\mathrm{BL},t}} \tag{24}$$

3.5.2 Reward to VaR Maximization

In both methods mentioned above, the standard deviation is used to measure the portfolio risk. As this is only appropriate when returns are normally distributed, we consider a more general formulation of the maximal reward-to-risk portfolio, where risk is measured by the VaR or CVaR of the BL portfolio. In particular, the optimization problem with maximal reward-to-VaR (MVaR-BL) is:

$$\begin{aligned} \max \quad & \frac{\boldsymbol{w}'_{\mathrm{BL},t}\boldsymbol{\mu}_{\mathrm{BL},t}}{\mathrm{VaR}_{\alpha,t}} \\ \text{s.t.} \quad & \boldsymbol{w}'_{\mathrm{BL},t}\mathbf{1} = 1 \end{aligned} \tag{25}$$

where $\mathrm{VaR}_{\alpha,t}$ is the expected maximum loss on the BL portfolio at time t with a confidence level of α . The VaR of the BL portfolio at time t can be expressed as:

$$\mathrm{VaR}_{\alpha,t} = \xi_\alpha \sqrt{\delta \boldsymbol{w}'_{\mathrm{BL},t}\boldsymbol{\Sigma}_{\mathrm{BL},t}\boldsymbol{w}_{\mathrm{BL},t}} - \boldsymbol{w}'_{\mathrm{BL},t}\boldsymbol{\mu}_{\mathrm{BL},t} \tag{26}$$

where, $\xi_\alpha = \Phi^{-1}(\alpha)$ is the $1-\alpha\%$ quantile of the cumulative normal distribution and α is equal to $1\%, 5\%$ and 10%.

3.5.3 Reward to CVaR Maximization

Similarly, the optimization problem with maximal reward-to-CVaR (MCVaR-BL) is:

$$\begin{aligned} \max \quad & \frac{\boldsymbol{w}'_{\mathrm{BL},t}\boldsymbol{\mu}_{\mathrm{BL},t}}{\mathrm{CVaR}_{\alpha,t}} \\ \text{s.t.} \quad & \boldsymbol{w}'_{\mathrm{BL},t}\mathbf{1} = 1 \end{aligned} \tag{27}$$

where $\mathrm{CVaR}_{\alpha,t}$ is the average loss exceeding the expected maximum loss $\mathrm{VaR}_{\alpha,t}$ at time t on the BL portfolio with a certain confidence level of $1-\alpha$. The CVaR of the BL portfolio at time t is:

$$\mathrm{CVaR}_{\alpha,t} = \zeta_\alpha \sqrt{\delta \boldsymbol{w}'_{\mathrm{BL},t}\boldsymbol{\Sigma}_{\mathrm{BL},t}\boldsymbol{w}_{\mathrm{BL},t}} - \boldsymbol{w}'_{\mathrm{BL},t}\boldsymbol{\mu}_{\mathrm{BL},t} \tag{28}$$

where, $\zeta_\alpha = \frac{\phi(\Phi^{-1}(\alpha))}{1-\alpha}$

4 Empirical Results

In this section, we discuss the out of sample performance of the portfolios detailed above. Using the weights obtained from the optimization of the above BL models, we calculate buy-and-hold returns on the portfolio for a holding period of one week and repeat the calculation until the end of the sample, and thus obtain the time series of realized portfolio returns. We report the average, standard deviation, skewness and kurtosis of the portfolio returns in the out-of-sample period. We also use the Average Herfindahl index (AHI) to measure diversification. The higher the value of the AHI, the less diversified the portfolio.

$$\mathrm{HHI}_t = \sum_{i=1}^{5} w_{i,t}^2 \tag{29}$$

To evaluate the performance of the portfolio, we use the Sharpe ratio and the ratio of reward-to down side risk, measured by VaR or CVaR computed from the empirical distribution. As an alternative to Sharpe Ratio we use Omega ratio also referred to as gain-loss ratio. It measures the average gains to average losses in which we define gains as returns above the risk-free rate and losses as returns below the risk-free rate. Hence, investments with a larger Omega measure are preferable. Formally, the omega measure is

$$\Omega = \frac{\frac{1}{T}\sum_{t=1}^{T} \max(0, r_{exs,t})}{\frac{1}{T}\sum_{t=1}^{T} \max(0, -r_{exs,t})} \tag{30}$$

The advantage of the Omega measure is that it does not require any assumption on the distribution of returns. As alternative risk measures besides volatility, we compute the maximum drawdown (MDD). The MDD reflects the maximum accumulated loss that an investor may suffer during the entire investment period if she buys the portfolio at a high price and subsequently sells at the lowest price. We compute MDD of returns as follows:

$$\mathrm{MDD}_T = \max_{\tau \in (0,T)} \mathrm{DD}_t = \max_{\tau \in (0,T)} \Big(0, \max_{t \in (0,\tau)} (0, r_{\{t,T\}})\Big) \tag{31}$$

Furthermore, we compute the portfolio turnover, which quantifies the extent of trading required to implement a certain strategy. The portfolio turnover is the average absolute change of the portfolio weights w over the T rebalancing points in time and across the 5 assets.

$$\mathrm{PT} = \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{5}(w_{j,t+1} - w_{j,t}) \tag{32}$$

in which $w_{j,t}$ is the weight of asset j before re-balancing at time $t+1$ and $w_{j,t+1}$ is the weight of asset j after re-balancing.

With these evaluation criteria, we compare the BL portfolios to the market weights portfolio, the naive 1/N portfolio. Finally, we investigate the impact of the choice of covariance estimates and the choice of δ on the performance of the BL portfolios.

We initially train both the random forest classifiers using first 561 observations and we predict the outcome for 562^{th} observation which the qualitative view for 563^{rd} week. Similarly we use the data available till 562^{th} observation to estimate the covariance matrix and using these both estimates we generate the portfolio weights for 563^{rd} week (This is the first week in the out of sample). Fixing the initial observation, the estimation sample is then rolled forward by one week, the models are re-estimated and used to generate out-of-sample forecasts for month 564, and so on until the end of the sample. This procedure results in a total of 141 out-of-sample weekly forecasts (20% of the full sample). We report the out of sample portfolio performance results in the below subsections.

4.1 Benchmark Portfolios

We compute two naive diversified portfolios that serve as the benchmark portfolios for the optimization models. First we calculate 1/N strategy which invests equally in the 5 assets. As a second näive diversification strategy we use a weighted portfolio in which each asset obtains the market capitalization weights that is constant over time, which is the neutral point in the Black-Litterman model. Table 6 shows that while the 1/N portfolio is more diversified, the market weight portfolio performs better. In particular the market weight portfolio generates a 6% higher Sharpe ratio with almost same reward to downside risk ratio.

4.2 Unconstrained BL Portfolios

In this section we evaluate the performance of the unconstrained Bl portfolios. Table 6 suggests that all the unconstrained BL portfolios outperform the 1/N and market portfolio with a huge margin. However, these both portfolios have a better Sharpe ratio, Omega, Reward per unit risk and little drawdown when compared to any other BL portfolios. Average HHI suggests that all the unconstrained BL portfolios are heavily concentrated and not as diversified as both benchmarks.

		Mean p.a (%)	SD p.a. (%)	Skewness	Kurtosis	μ/VaR	μ/CVaR	Average HHI	Sharpe Ratio	Omega	MDD	Turnover
	Market Weights	7.94	11.65	-0.02	4.02	-0.0604	-0.0453	0.34	0.6810	1.280	0.14	0.00
	1/N	6.45	10.01	0.02	3.58	-0.0605	-0.0451	0.20	0.6440	1.257	0.11	0.00
	Sample	14674.09	74572.83	0.03	3.04	-0.0159	-0.0133	41648823.51	0.1968	1.073	1116.28	-11.36
$\delta = 0.01$	Fast MCD	20241.60	87330.28	-0.03	3.21	-0.0176	-0.0151	56586765.79	0.2318	1.087	1211.48	-16.85
	EWMA	44756.42	137498.59	0.22	3.76	-0.0283	-0.0226	142056076.00	0.3255	1.127	2069.02	-13.21
	Shrinkage	14798.66	74332.10	0.02	3.05	-0.0161	-0.0135	41146753.75	0.1991	1.074	1112.13	-11.35
	Sample	73.41	334.88	0.01	3.08	-0.0175	-0.0147	831.18	0.2192	1.082	4.98	-0.05
$\delta = 2.24$	Fast MCD	100.35	391.00	-0.04	3.25	-0.0199	-0.0167	1128.84	0.2567	1.098	5.40	-0.08
	EWMA	207.71	615.45	0.21	3.75	-0.0299	-0.0234	2831.88	0.3375	1.132	9.23	-0.06
	Shrinkage	73.97	333.79	0.01	3.09	-0.0177	-0.0148	821.19	0.2216	1.083	4.96	-0.05
	Sample	32.38	126.58	-0.01	3.17	-0.0202	-0.0170	116.32	0.2558	1.096	1.85	-0.02
$\delta = 6$	Fast MCD	41.62	147.43	-0.07	3.31	-0.0225	-0.0183	158.02	0.2823	1.108	2.01	-0.03
	EWMA	82.52	230.97	0.19	3.73	-0.0328	-0.0247	395.08	0.3573	1.140	3.44	-0.02
	Shrinkage	32.59	126.17	-0.01	3.17	-0.0204	-0.0171	114.93	0.2583	1.097	1.85	-0.02

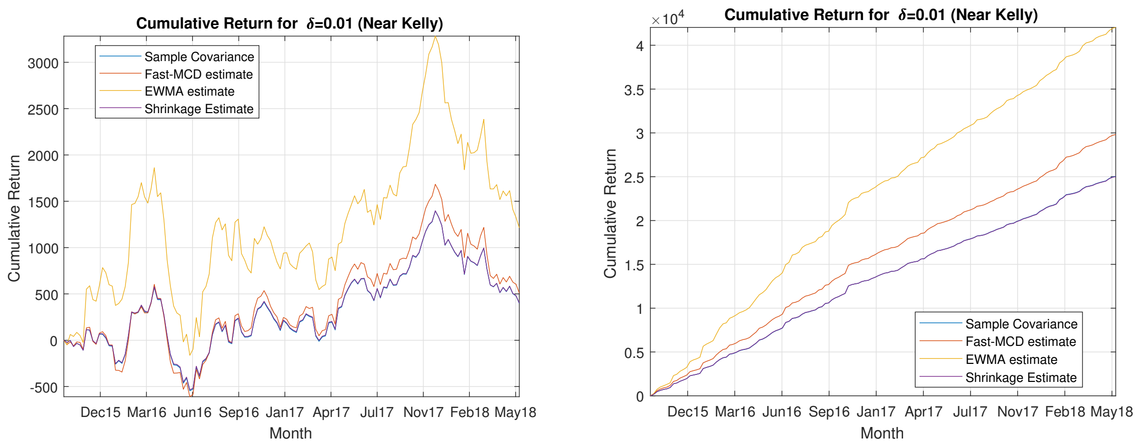
Table 6: Out of sample performance of Unconstrained Black-Litterman portfolios with Random Forest based views

By examining the unconstrained Black-Litterman portfolios, we also find that the portfolios with $\delta = 0.01$ (Near Kelly Investors) have more concentrated weights and thus higher return and high risk when compared to other values of δ. However their reward per unit risk is very high. Also, one can observe that the portfolio turnover values from table 6 suggest that these strategies are difficult to implement when compared to other unconstrained dynamic portfolios. And within the group of portfolios with same δ the portfolios which uses EWMA estimate outperforms the other by a factor close to 2(when compared with Fast MCD estimates). The portfolios which uses sample covariance and shrinkage have almost similar measures. An interesting observation is that portfolios with Fast MCD covariance estimates are difficult to execute when compared to other BL portfolios which have different covariance estimates and although, BL portfolios with EWMA Covariance estimates have highest SR and Ω, they are prone to higher drawdown.

δ	Method	Mean p.a (%)	SD p.a. (%)	Skewness	Kurtosis	μ/VaR	μ/CVaR	Average HHI	Sharpe Ratio (p.a)	Omega	MDD	Turnover
	Sample	923447.1	84076.4	2.36	11.81	2.74	3.44	38390741.14	10.98	∞	0.00	44.52
$\delta = 0.01$	Fast MCD	1099457.9	97141.1	2.15	9.86	2.74	3.39	51179373.51	11.26	∞	0.00	58.11
	EWMA	1550374.5	152796.0	2.21	9.76	3.14	3.87	129656485.74	10.15	∞	0.00	66.52
	Shrinkage	922231.8	83901.6	2.35	11.68	2.73	3.43	38009690.54	10.99	∞	0.00	44.21
	Sample	4130.4	375.0	2.34	11.65	2.72	3.46	766.83	11.01	∞	0.00	0.20
$\delta = 2.24$	Fast MCD	4916.7	434.0	2.15	9.89	2.73	3.41	1022.81	11.33	∞	0.00	0.26
	EWMA	6929.2	681.0	2.19	9.63	3.11	3.86	2586.35	10.18	∞	0.00	0.30
	Shrinkage	4125.0	374.2	2.32	11.52	2.74	3.46	759.23	11.02	∞	0.00	0.20
	Sample	1547.0	140.1	2.29	11.36	2.93	3.55	107.49	11.04	∞	0.00	0.07
$\delta = 6$	Fast MCD	1842.8	162.7	2.12	9.73	2.69	3.43	143.39	11.33	∞	0.00	0.10
	EWMA	2591.9	253.7	2.15	9.40	3.09	3.87	361.24	10.22	∞	0.00	0.11
	Shrinkage	1545.0	139.8	2.28	11.23	2.89	3.53	11.05	57.00	∞	0.00	0.07

Table 7: Out of sample performance of unconstrained Black-Litterman with Correct views

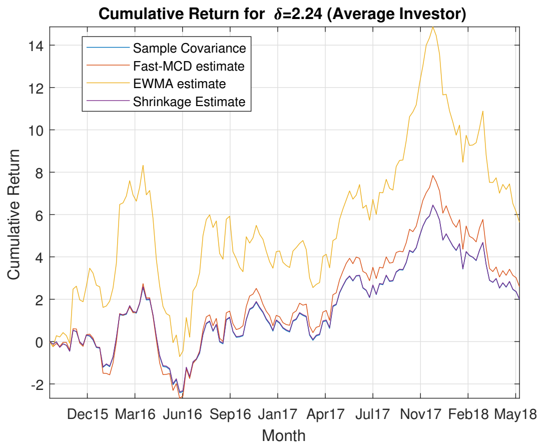
Figures shows the cumulative return of all the Unconstrained BL portfolios with random forest based views as a comparison with the same portfolio with all the correct views. Table 7 summarizes the out of sample performance of these Unconstrained dynamic portfolios with true views. The huge difference in these portfolios is due to the accuracy of classifier 1. Though the classifier 2 has a decent accuracy (since most of the weeks return fall under 1 standard deviation) it's the accuracy of classifier 1(sign of next week's return) which resulted in a Q_t which is far from the true Q matrix representing the qualitative views. It is observed that both classifiers are able to correctly predict only 38% of the total Y values (141×5).



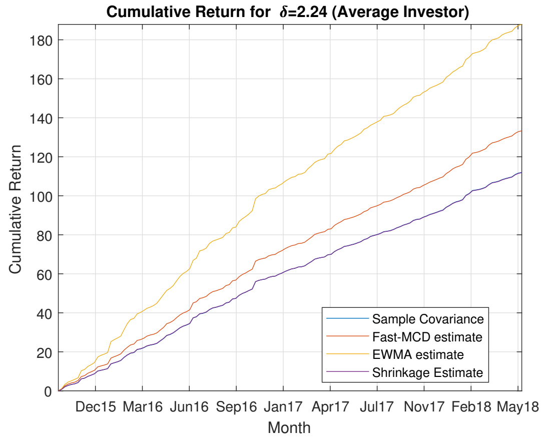
(a) Cumulative returns with Random Forest Views

(b) Cumulative returns with Correct Views

Figure 12: Cumulative returns of unconstrained Black-Litterman portfolios of Near Kelly Investor

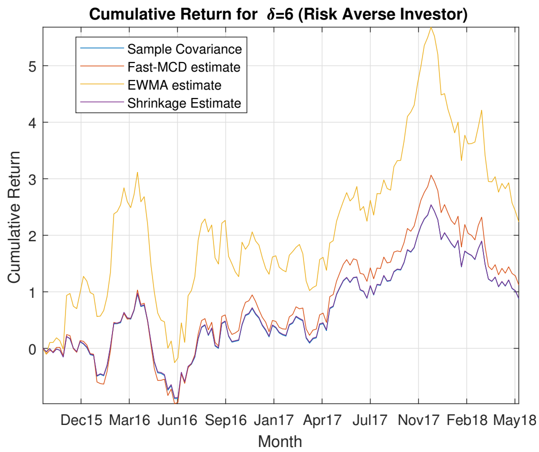


(a) Cumulative returns with Random Forest Views

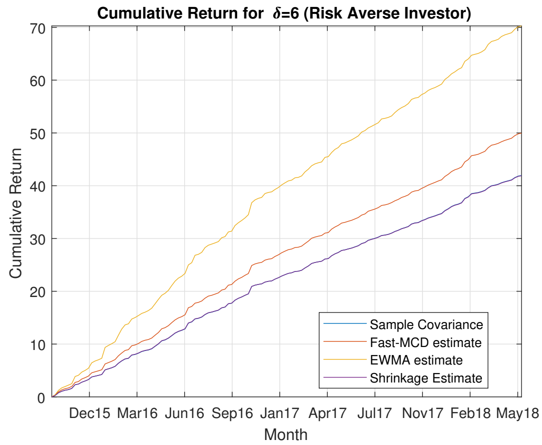


(b) Cumulative returns with Correct Views

Figure 13: Cumulative returns of unconstrained Black-Litterman portfolios of Average Investor



(a) Cumulative returns with Random Forest Views



(b) Cumulative returns with correct Views

Figure 14: Cumulative returns of unconstrained Black-Litterman portfolios of Risk Averse Investor

4.3 Constrained BL Portfolios

Table 8 reports the out of sample performance results of the BL portfolios with no short selling and budget constraints. It can be seen that the constrained BL portfolios slightly outperform the benchmark portfolios. The findings and impact of the value of δ is same as discussed in the previous section. In particular, we find that EWMA based BL portfolios are superior in terms of performance and inferior in terms of diversification. Also the Skewness and Kurtosis values reveal that extreme negative skewness and high kurtosis leads to higher tail risk in the case of EWMA based Constrained Portfolios.

		Mean p.a (%)	SD p.a. (%)	Skewness	Kurtosis	μ/VaR	μ/CVaR	Average HHI	Sharpe Ratio	Omega	MDD	Turnover
	Market Weights	7.94	11.65	-0.02	4.02	-0.0604	-0.0453	0.34	0.6810	1.280	0.14	0.00
	1/N Portfolio	6.45	10.01	0.02	3.58	-0.0605	-0.0451	0.20	0.6440	1.257	0.11	0.00E+00
$\delta = 0.01$	Sample	8.91	15.32	-0.11	5.49	-0.0550	-0.0361	1.00	0.5820	1.242	0.14	-2.49E-18
	Fast MCD	8.24	15.36	-0.11	5.45	-0.0508	-0.0333	1.00	0.5362	1.222	0.15	2.22E-18
	EWMA	11.15	16.18	-0.47	6.18	-0.0659	-0.0400	1.00	0.6888	1.303	0.14	3.56E-18
	Shrinkage	9.54	15.27	-0.13	5.55	-0.0589	-0.0386	1.00	0.6251	1.263	0.14	8.06E-19
$\delta = 2.24$	Sample	10.49	15.26	-0.14	5.56	-0.0648	-0.0425	0.96	0.6879	1.293	0.14	-1.45E-18
	Fast MCD	9.83	15.32	-0.14	5.49	-0.0607	-0.0397	0.96	0.6414	1.271	0.13	-2.30E-18
	EWMA	10.77	16.08	-0.48	6.24	-0.0637	-0.0389	0.99	0.6702	1.294	0.15	1.84E-19
	Shrinkage	10.61	15.21	-0.15	5.56	-0.0655	-0.0432	0.95	0.6972	1.297	0.14	2.85E-18
$\delta = 6$	Sample	12.10	14.98	-0.14	5.53	-0.0816	-0.0520	0.89	0.8080	1.350	0.12	4.82E-18
	Fast MCD	10.34	15.33	-0.12	5.43	-0.0638	-0.0422	0.91	0.6747	1.284	0.12	1.69E-18
	EWMA	10.26	15.96	-0.46	6.34	-0.0606	-0.0374	0.96	0.6428	1.281	0.15	7.92E-18
	Shrinkage	11.91	14.89	-0.16	5.52	-0.0803	-0.0512	0.88	0.7995	1.345	0.12	2.84E-19

Table 8: My caption

Figure 15 presents the optimized portfolio weights for constrained BL portfolios for the three different investor types during the out of sample period. In line with the turnover and diversification measures, the figures reveals less extreme portfolios in the case of Risk Averse investor.

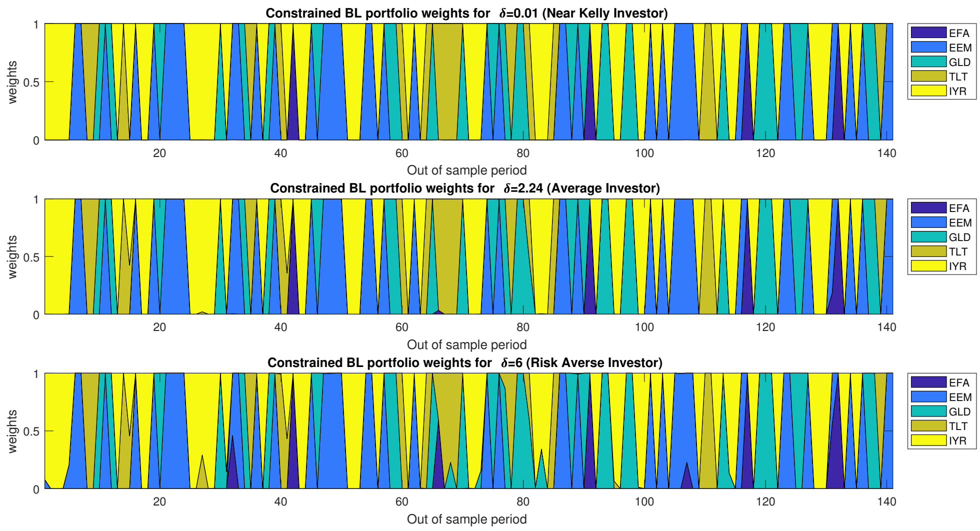


Figure 15: Optimized out of sample weights for Constrained Black-Litterman portfolios.

4.4 Maximum Sharpe Portfolios

In this subsection we evaluate the performance of the Maximum Sharpe Ratio - BL portfolios with constraints and compare its performance with that of BL portfolios with no short selling and budget constraints. Table 9 indicates that SR-BL portfolios with constraints has a better risk adjusted performance as compared with that of BL portfolios with no short selling and budget constraints. Also the value of HHI indicates that the weights are less concentrated when compared to the constrained BL portfolios. As in the case of section 4.3 the EWMA based portfolios have higher negative skewness and kurtosis leading to fat tail risks.

		Mean p.a (%)	SD p.a. (%)	Skewness	Kurtosis	μ/VaR	μ/CVaR	Average HHI	Sharpe Ratio	Omega	MDD	Turnover
	Market Weights	7.94	11.65	-0.02	4.02	-0.0604	-0.0453	0.34	0.6810	1.280	0.14	0.00
	1/N Portfolio	6.45	10.01	0.02	3.58	-0.0605	-0.0451	0.20	0.6440	1.257	0.11	0.00E+00
$\delta = 0.01$	Sample	9.46	11.16	-0.65	6.74	-0.0941	-0.0557	0.59	0.8477	1.373	0.09	3.77E-18
	Fast MCD	11.17	11.85	-0.11	7.10	-0.1083	-0.0657	0.54	0.9425	1.433	0.09	4.59E-18
	EWMA	9.21	11.12	-0.59	6.77	-0.0897	-0.0556	0.71	0.8281	1.369	0.11	2.34E-18
	Shrinkage	9.59	11.16	-0.55	6.47	-0.0959	-0.0569	0.58	0.8596	1.378	0.09	7.05E-18
$\delta = 2.24$	Sample	10.25	11.17	-0.38	5.78	-0.1028	-0.0619	0.57	0.9179	1.407	0.09	-7.57E-19
	Fast MCD	11.36	11.69	0.08	6.47	-0.1112	-0.0690	0.53	0.9716	1.444	0.09	-2.82E-18
	EWMA	9.10	11.11	-0.58	6.78	-0.0895	-0.0550	0.70	0.8184	1.366	0.11	-1.27E-18
	Shrinkage	10.33	11.17	-0.32	5.66	-0.1039	-0.0628	0.56	0.9251	1.411	0.09	-5.21E-18
$\delta = 6$	Sample	10.39	10.97	-0.17	4.99	-0.1029	-0.0652	0.54	0.9476	1.417	0.08	-7.29E-18
	Fast MCD	11.48	11.51	0.22	5.94	-0.1135	-0.0723	0.51	0.9976	1.454	0.08	-1.73E-18
	EWMA	8.81	11.03	-0.61	6.88	-0.0882	-0.0535	0.68	0.7985	1.357	0.11	-4.22E-19
	Shrinkage	10.47	10.98	-0.12	4.98	-0.1037	-0.0660	0.54	0.9535	1.420	0.08	-6.80E-18

Table 9: Out of sample performance of SR-BL portfolios

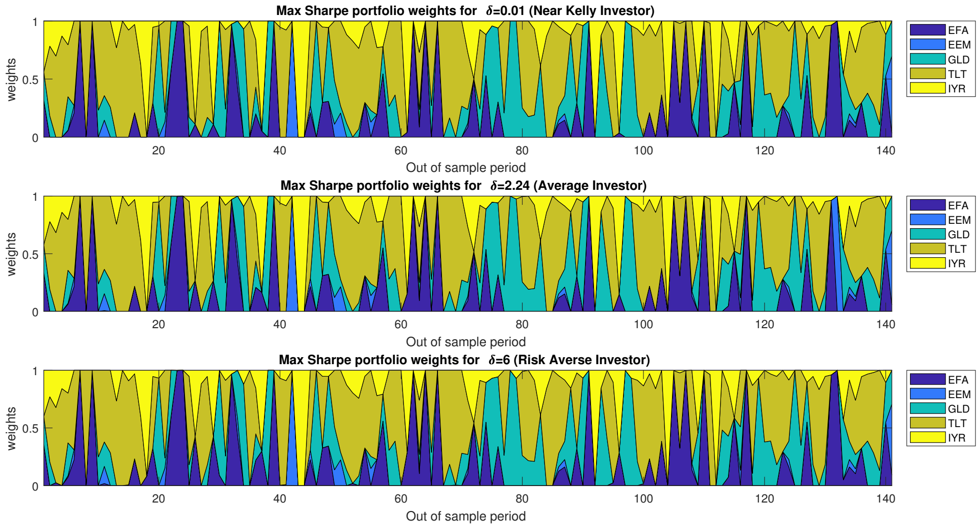


Figure 16: Optimized out of sample weights for SR BL portfolios.

In line with average HHI measures, Figure 16 reveals that Near Kelly investors have concentrated bets when compared to other types and the diversification is more in the case of risk averse investors. Additionally comparing Figure 15 and 16 we find that average HHI is higher for BL portfolios with constraints as compared to SR-BL portfolios with constraints. Consequently SR-BL portfolios offer more diversification across asset classes and have less extreme allocations.

4.5 Maximum Reward to VaR Portfolios

In this subsection, we discuss the out of sample performance of the Maximum Reward to VaR portfolio introduced in section 3.x. Contrast to the results discussed in the above subsections, Table 10 suggests that near kelly investors in this case outperform other types and also has less standard deviation. The regular trend of EWMA based estimates having higher negative skewness and kurtosis is not observed in this case. Also the average HHI values and the Figure 17 reveals that the near kelly investors have less concentrated bets and thus more diversification across asset classes. However, when compared with other portfolios it is observed that Max reward to VaR do not outperform both SR-BL and BL portfolios with constraints.

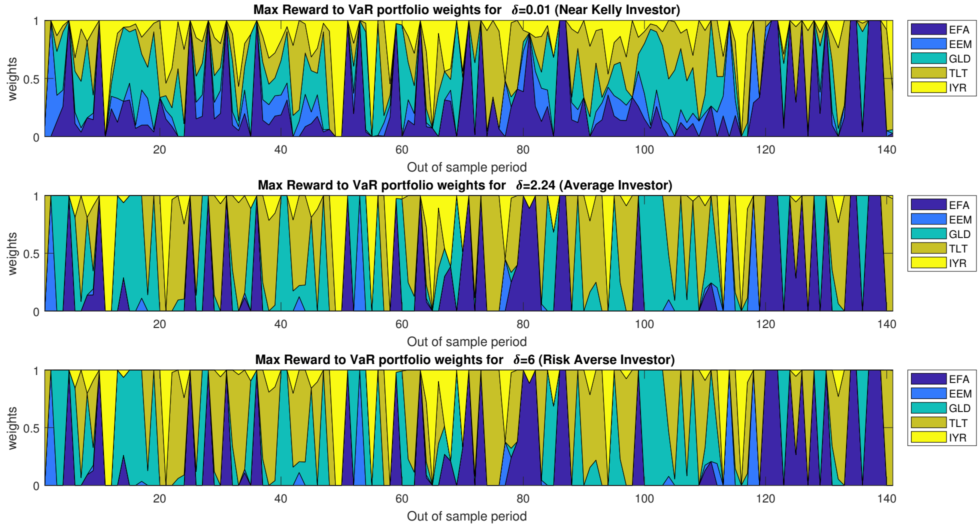


Figure 17: Optimized out of sample weights for MVaR-BL portfolios.

		Mean p.a (%)	SD p.a. (%)	Skewness	Kurtosis	μ/VaR	μ/CVaR	Average HHI	Sharpe Ratio	Omega	MDD	Turnover
	Market Weights	7.94	11.65	-0.02	4.02	-0.0604	-0.0453	0.34	0.6810	1.280	0.14	0.00
	1/N Portfolio	6.45	10.01	0.02	3.58	-0.0605	-0.0451	0.20	0.6440	1.257	0.11	0.00E+00
	Sample	8.22	11.40	-0.28	3.61	-0.0737	-0.0483	0.50	0.7212	1.289	0.17	4.54E-19
$\delta = 0.01$	Fast MCD	6.72	11.80	-0.49	4.31	-0.0555	-0.0356	0.48	0.5696	1.226	0.16	6.50E-19
	EWMA	8.35	11.87	-0.38	3.82	-0.0642	-0.0459	0.57	0.7038	1.281	0.17	2.05E-18
	Shrinkage	7.64	11.46	-0.28	3.54	-0.0657	-0.0446	0.50	0.6663	1.263	0.18	-4.48E-19
	Sample	4.36	12.42	-0.34	3.17	-0.0285	-0.0223	0.80	0.3512	1.130	0.20	-8.62E-19
$\delta = 2.24$	Fast MCD	3.36	12.85	-0.43	3.63	-0.0220	-0.0159	0.74	0.2616	1.098	0.20	-2.59E-18
	EWMA	3.15	12.54	-0.40	3.47	-0.0219	-0.0159	0.86	0.2510	1.092	0.21	-5.90E-18
	Shrinkage	4.47	12.44	-0.33	3.16	-0.0292	-0.0228	0.79	0.3591	1.133	0.20	-7.78E-19
	Sample	3.92	12.46	-0.35	3.19	-0.0256	-0.0200	0.82	0.3150	1.116	0.20	-2.28E-18
$\delta = 6$	Fast MCD	2.90	12.87	-0.43	3.60	-0.0190	-0.0137	0.77	0.2257	1.083	0.20	-5.91E-18
	EWMA	3.03	12.57	-0.41	3.51	-0.0206	-0.0152	0.87	0.2410	1.088	0.21	-1.08E-18
	Shrinkage	3.94	12.45	-0.35	3.18	-0.0258	-0.0202	0.82	0.3168	1.117	0.20	2.85E-18

Table 10: Out of sample performance of MVaR-BL portfolios

4.6 Maximum Reward to CVaR Portfolios

Table 11 indicates that there is no regular trend in terms of performance or diversifcation as in the case of the previously discussed portfolios. It can be observed that in terms of mean return some of these portfolios doesn't even outperform the Market weights based portfolio . Also it can be seen that all of these portfolios have a positive skewness and less kurtosis as compared to other risk adjusted portfolios discussed above.

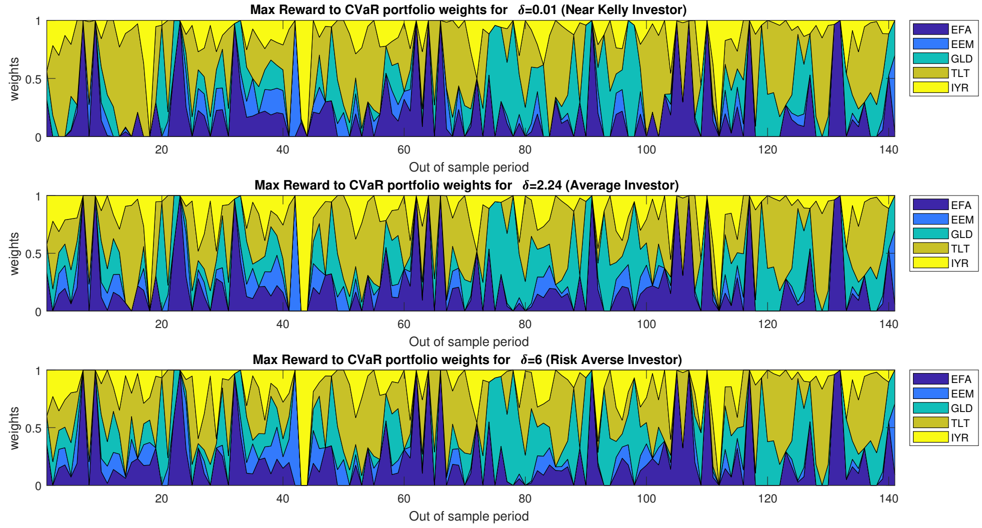


Figure 18: Optimized out of sample weights for MCVaR-BL portfolios.

When compared to Maximal reward to VaR portfolios it can be observed that Maximal reward to CVaR have higher sharpe ratios and Omega ratio with little Maximum Drawdown. However, in terms of SR and Omega many of the these portfolio clearly cannot outperform SR-BL portfolios even with better diversification among assets. It can also be seen that Maximum reward to CVaR portfolios have less drawdown and risk when compared with all other BL portfolios.

		Mean p.a (%)	SD p.a. (%)	Skewness	Kurtosis	μ/VaR	μ/CVaR	Average HHI	Sharpe Ratio	Omega	MDD	Turnover
	Market Weights	7.94	11.65	-0.02	4.02	-0.0604	-0.0453	0.34	0.6810	1.280	0.14	0.00
	1/N Portfolio	6.45	10.01	0.02	3.58	-0.0605	-0.0451	0.20	0.6440	1.257	0.11	0.00E+00
	Sample	6.76	10.54	0.08	3.62	-0.0590	-0.0447	0.47	0.6415	1.262	0.08	1.53E-18
$\delta = 0.01$	Fast MCD	9.69	10.88	0.18	3.78	-0.0857	-0.0647	0.42	0.8912	1.382	0.09	-1.81E-18
	EWMA	7.65	10.52	0.13	3.87	-0.0715	-0.0519	0.55	0.7273	1.306	0.10	-3.61E-18
	Shrinkage	5.84	10.49	0.10	3.64	-0.0510	-0.0387	0.46	0.5571	1.223	0.09	-1.04E-18
	Sample	8.02	10.67	0.14	3.80	-0.0753	-0.0540	0.44	0.7522	1.313	0.09	3.71E-18
$\delta = 2.24$	Fast MCD	7.69	11.14	0.24	3.92	-0.0722	-0.0514	0.40	0.6906	1.285	0.10	-3.32E-18
	EWMA	7.36	10.70	0.21	4.05	-0.0694	-0.0503	0.53	0.6879	1.287	0.09	-9.40E-18
	Shrinkage	8.03	10.65	0.15	3.78	-0.0756	-0.0543	0.44	0.7540	1.313	0.09	-1.17E-18
	Sample	7.80	10.95	0.13	3.95	-0.0754	-0.0516	0.43	0.7123	1.294	0.09	4.51E-18
$\delta = 6$	Fast MCD	8.03	10.85	0.17	3.84	-0.0770	-0.0547	0.40	0.7400	1.306	0.09	-4.62E-20
	EWMA	7.88	10.91	0.25	4.52	-0.0746	-0.0530	0.51	0.7217	1.308	0.09	-8.20E-19
	Shrinkage	7.97	10.91	0.16	3.89	-0.0770	-0.0532	0.43	0.7303	1.302	0.09	-2.05E-18

Table 11: Out of sample performance of MVaR-BL portfolios

4.7 Summary

In this section we discuss the performance ranking of the BL based portfolios discussed in the previous sections. We select the EWMA based portfolios in each section and report their rankings based on different measures. Based on Average HHI 1/N and Benchmark portfolio have the best performance. However, these two portfolios are outperformed by most constrained BL portfolios in terms of risk adjusted performance. Table 12 shows at an overall level the best performing portfolios are Max SR-Bl and Max Reward to CVaR portfolios. SR-BL portfolios have higher Sharpe Ratio ($\approx 14\%$) and Omega ratio ($\approx 5\%$). However, the average HHI for SR-BL portfolios are almost 40% more than that of Max reward to CVaR portfolios. It can be seen that except for the case of Max Reward to VaR BL-1, all the Max Reward CVaR outperforms Max reward to VaR portfolios.

	Mean p.a (%)	Rank	SD p.a. (%)	Rank	Average HHI	Rank	Sharpe Ratio	Rank	Omega	Rank	MDD	Rank
Benchmark	7.94	8	11.65	8	0.34	2	0.6818	9	1.280	11	0.14	8
1/N Portfolio	6.45	12	10.01	1	0.20	1	0.6440	11	1.257	12	0.11	7
BL-1	11.15	1	16.18	14	1.00	14	0.6888	7	1.303	6	0.14	9
BL-2	10.77	2	16.08	13	0.99	13	0.6702	10	1.294	7	0.15	10
BL-3	10.26	3	15.96	12	0.96	12	0.6428	12	1.281	9	0.15	11
SR BL -1	9.21	4	11.12	7	0.71	9	0.8281	1	1.369	1	0.11	6
SR BL -2	9.10	5	11.11	6	0.70	8	0.8184	2	1.366	2	0.11	5
SR BL -3	8.81	6	11.03	5	0.68	7	0.7985	3	1.357	3	0.11	4
MVaR BL-1	8.35	7	11.87	9	0.57	6	0.7038	6	1.281	10	0.17	12
MVaR BL-2	3.15	13	12.54	10	0.86	10	0.2510	13	1.092	13	0.21	13
MVaR BL-3	3.03	14	12.57	11	0.87	11	0.2410	14	1.088	14	0.21	14
MCVaR BL-1	7.65	10	10.52	2	0.55	5	0.7273	4	1.306	5	0.10	3
MCVaR BL-2	7.36	11	10.70	3	0.53	4	0.6879	8	1.287	8	0.09	1
MCVaR BL-3	7.88	9	10.91	4	0.51	3	0.7217	5	1.308	4	0.09	2

Table 12: Out of sample performance ranking for various Black-Litterman Portfolios

5 Conclusion

In this paper we've implemented the Black Litterman framework to a dynamic asset allocation framework and analyzed its out of sample performance with naive diversified portfolios. We've derived the qualitative views using machine learning classifiers which are converted into quantitative views and used as a input the BL models. We analyze different constrained BL portfolios which maximizes Sharpe, Reward to VaR and Reward to CVaR respectively. We also use different performance measures, which lead to different rankings of these BL portfolios. Further, we examine the effect of the choice of covariance estimates and the choice of risk aversion coefficient on portfolio diversification and portfolio performance. The out of sample analysis highlights the following important findings.

- When compared to classifier 2, the classifier 1 highly influences the out of sample portfolio performance.
- Almost all BL portfolios outperform 1/N and market weight portfolios. However they have more concentrated weights and consequently less diversification among the assets.
- Except for Maximum Reward to CVaR portfolios, EWMA based portfolios outperform all the other portfolio with different covariance estimates.However, they have higher negative values of skewness and kurtosis resulting in fat tail risks.
- Both SR-BL and Maximum Reward to CVaR portfolios outperform Maximum reward to VaR portfolios.
- Maximum Reward to CVaR and SR-BL portfolios have better risk adjusted performance.
- Maximum Reward to CVaR portfolios have lowest MDD when compared to other constrained BL portfolios.

References

[1] Attilio Meucci (October 2010), *The Black litterman Approach: Originla Model and Extensions.*

[2] Attilio Meucci (2005). *Risk and Asset Allocation.* Berlin, Heidelberg. Springer-Verlag Berlin Heidelberg.

[3] Thomas M. Idzorek, *A Step-By-Step guide to the Black Litterman Model: Incorporating user specified confidence intervals.*, (July 2004).

[4] Black, F., and R. Litterman. 1992. *Global Portfolio Optimization.* Financial Analysts Journal 48 (5): 28-43.

[5] DeMiguel, V., L. Garlappi, and R. Uppal. 2009. *Optimal Versus Naive Diversification: How Inefficient Is the 1/N Portfolio Strategy?* Review of Financial Studies 22 (5): 1915-1953.

[6] He, G., and R. Litterman. 1999. *The Intuition Behind the Black Litterman Model Portfolios.* Investment Management Research, Goldman Sachs Quantitative Resources Group, 115.

[7] Walters, Jay. 2014. *The Black-Litterman Model in Detail.* Working paper available on SSRN.

[8] Olivier Ledoit, MichaelWolf. *Honey, I shrunk the sample covariance matrix.* UPF Economics and Business Working Paper, 2003.

[9] Rousseuw, P.J. and Van Driessen, K. *A fast algorithm for the minimum covariance determinant estimator.* Technometrics, Vol. 41, 1999.

[10] Richard D F Harris, Evarist Stoja and Linzhi Tan. *Staff Working Paper No. 596 The dynamic Black-Litterman approach to asset allocation.* Bank of England.

[11] Wolfgang Bessler, Heiko Opfer & Dominik Wolff. *Multi-asset portfolio optimization and outof- sample performance: an evaluation of BlackLitterman, mean-variance, and nave diversification approaches.*. The European Journal of Finance 2014.

[12] Olivier Ledoit and Michael Wolf. *Improved estimation of the covariance matrix of stock returns with an application to portfolio selection.* Journal of empirical finance, 10(5) 603 621, 2003.

[13] Breiman, L. *Random Forests. Machine Learning* 45, pp. 5-32, 2001.

[14] Jim Kyung-Soo Liew and Boris Mayster. *Forecasting ETFs with Machine Learning Algorithms.*

[15] RiskMetrics. *Technical Document.*4th edition, 1996.

Appendices

Description of Codes used in MatLab

Sl.No	Function Name	Notes
1	bl_cov	Helper function which gives us choice of the covarariance estimate to be used in obtaining weights.
2	covCor	Computes the shrinkage estimate of covariance matrix. Written by Ledoitt and Wolf. Released under BSD-2 licensce
3	ewma_cov	Computes the ewma covariance matrix based on the algorithm described in Risk Metrics Technical document.
4	getDescriptiveStatistics	Helper function obtianed from MatLab file exchange to compute various descriptive statistics at once.
5	hist_stock_data	Used to extract the ticker related data from Yahoo Finance
6	indep_features	Computes the significance level of each features based on Weiss/Indurkhya test
7	indicators	Computes various technical indicators in MatLab. Obtained from MatLab File Exchange

Table 13: Description of MatLab functions used

Sl.No	Code Name	Notes
1	p1_data_extract1	Obtain OHLC data of ETFs from Yahoo Finance
2	p1_data_extract2	Obtain OHLC data of IRX from Yahoo finance
3	p1_data_extract3	Obtain data fo SPY, VIX and 10 year treasury bond index from Yahoo Finance
4	p2_returns	Compute the excess returns
5	p2_returns_descriptive_stats	Computes various descriptive statistics and performs statistical tests on Excess returns
6	p3_ewma_1	Computes ewma estimate for the sample(illustration purpose for report)
7	p3_ewma_2	Computes optimal lamba for ewma estimates based on RMSE and MAE
8	p3_sample_robust_cov	Computes fast MCD estimate for the sample(illustration purpose for report)
9	p3_shrinkage_cov	Computes fast shrinkage estimate for the sample(illustration purpose for report)
10	p4_features_prep_1	Calculates Information sets 1-3
11	p4_features_prep_2	Calculates various technical indicators to create other information sets
12	p4_features_prep_3	Combines features from p4_features_prep_1 and p4_features_prep_2
13	p5_data_prep1	Remove NaN and impute missing values
14	p5_data_prep2	Feature selection for classification problem1
15	p5_data_prep3	Feature selection for classification problem2
16	p6_sample_classifier1	Performance of classifier 1 on full sample
17	p6_sample_classifier2	Performance of classifier 2 on full sample
18	p7_views_helper_1	Prepares classifier 1 views
19	p7_views_helper_2	Prepares classifier 2 views
20	p7_views_helper_3	Computes the final entries which goes into Q matrix for all portfolios
21	p8_1_marketcap_portfolio	Computes out of sample performance of market cap portfolio
22	p8_2_1byn_portfolio	Computes out of sample performance of 1/N portfolio
23	p8_3_main_bl_rfviews	Computes out of sample performance of unconstrained BL portfolio based on Random Forest views
24	p8_3_main_bl_trueviews	Computes out of sample performance of unconstrained BL portfolio based on Correct views
25	p8_4_main_bl_constraints	Computes out of sample performance of Constrained BL portfolio based on Random Forest views
26	p8_5_sr_bl_portfolios	Computes out of sample performance of SR-BL portfolio based on Random Forest views
27	p8_6_VaR_bl_portfolios	Computes out of sample performance of MVaR-BL portfolio based on Random Forest views
28	p8_7_CVaR_bl_portfolios	Computes out of sample performance of MCVaR-BL portfolio based on Random Forest views

Table 14: Description of MatLab Codes used for Out of Sample Performance