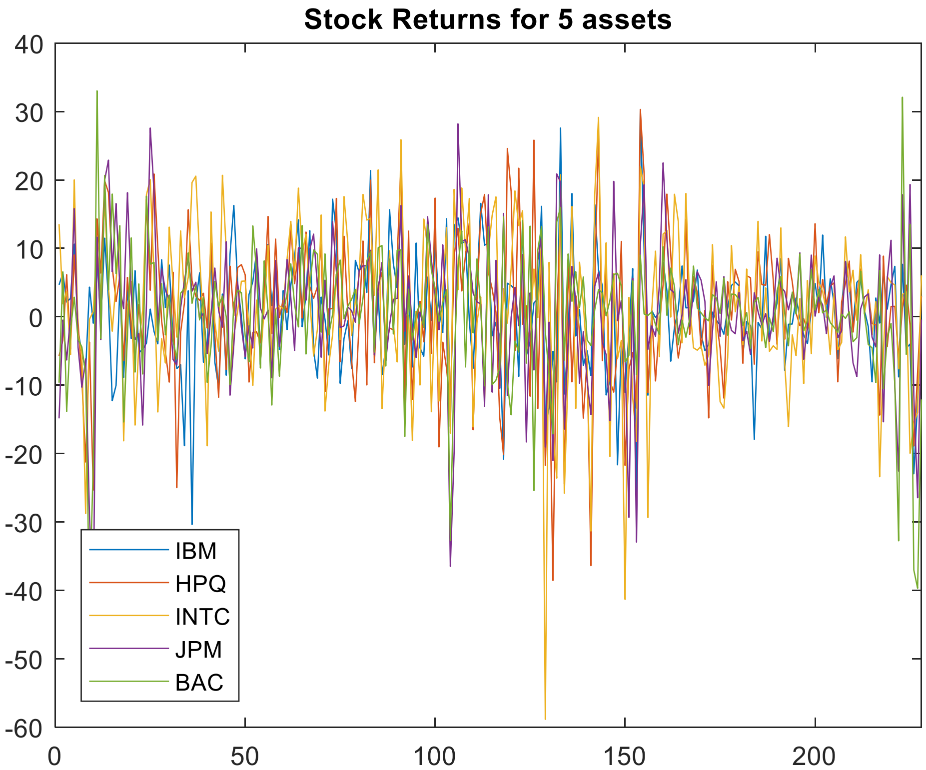
QBUS6830 Financial Time Series and Forecasting
S1, 2019

Solutions to Lab Sheet 6

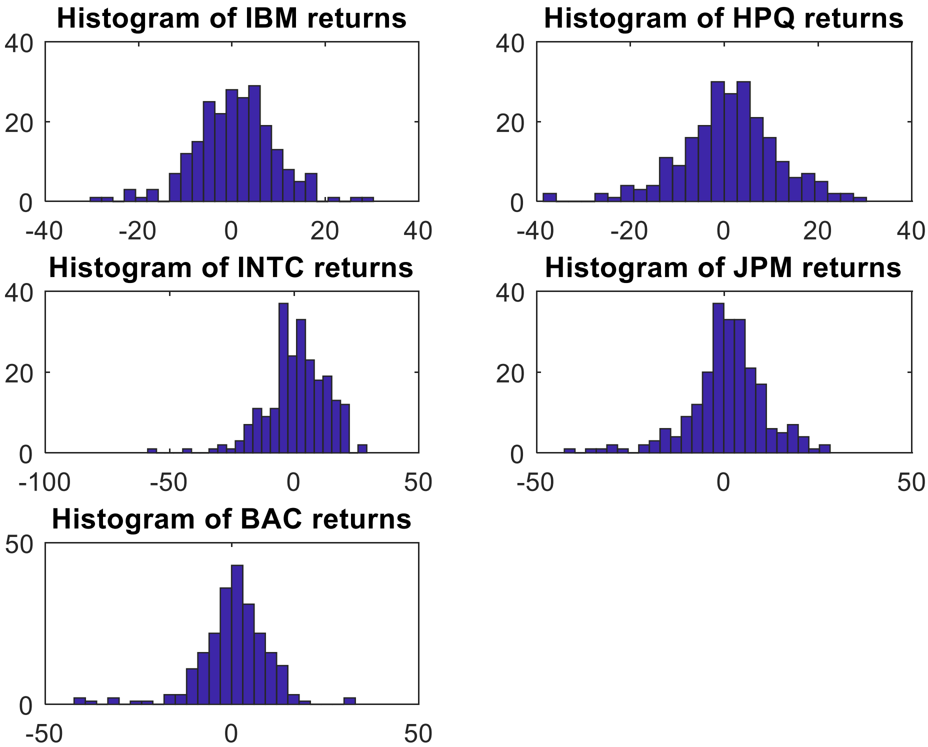
Q1 (Forecasting)
We use the data from the text by Tsay in Chapter 9, being monthly returns on IBM, HPQ, Intel, JP Morgan and Bank of America, from January, 1990 to December, 2008.

(a) Conduct a brief EDA on the five assets.

The times series are plot here:



The low volatility period leading up to 2008 is readily apparent, as is the increase in volatility marking the GFC period. The returns are clearly stationary in mean.



Clearly these assets are quite fat-tailed and each has many outlying return observations.

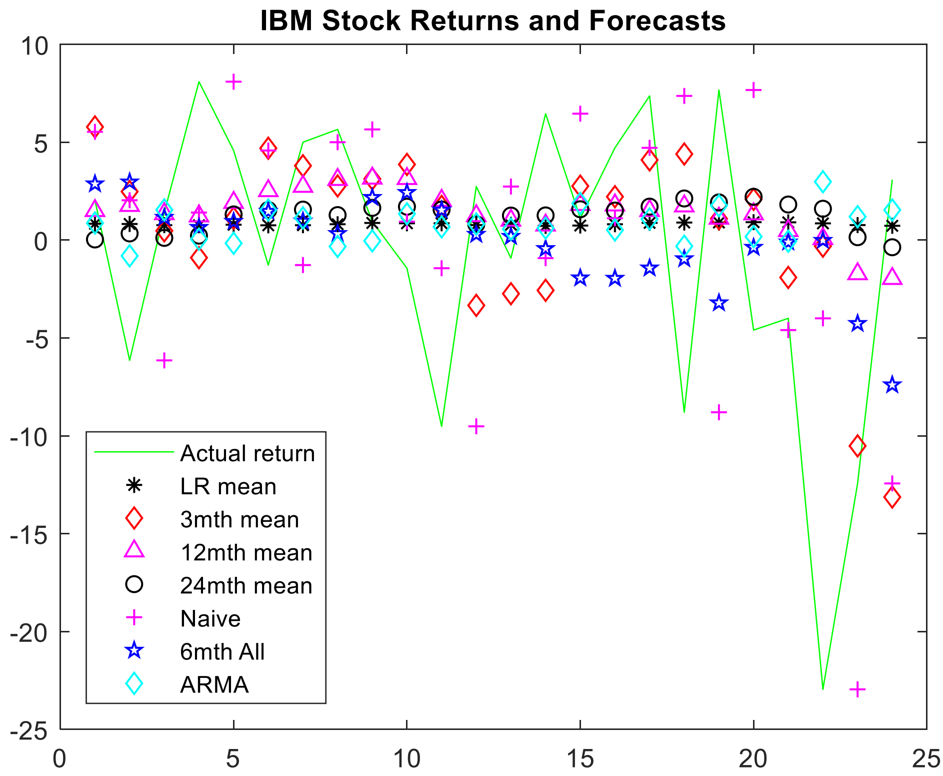
	IBM	HPQ	INTC	JPM	BAC
Mean	0.700	0.986	1.195	0.823	0.410
Median	1.007	1.256	1.715	0.987	0.644
Std	8.639	10.593	12.103	10.297	9.582
Min	-30.368	-38.550	-58.862	-42.583	-42.285
Max	30.291	30.298	29.136	28.197	33.024
Skew	-0.132	-0.356	-0.858	-0.726	-1.024
Kurt	4.274	4.157	5.464	5.436	7.681

All have positive means and medians during this time period. All have negative skew estimates, no doubt influenced by having a few more large negative outlying returns, than positive. From the histograms, most appear quite symmetric, except for INTC which shows a clear negative skew and long left tail. They are all clearly more fat-tailed than a Gaussian (kurtosis > 3).

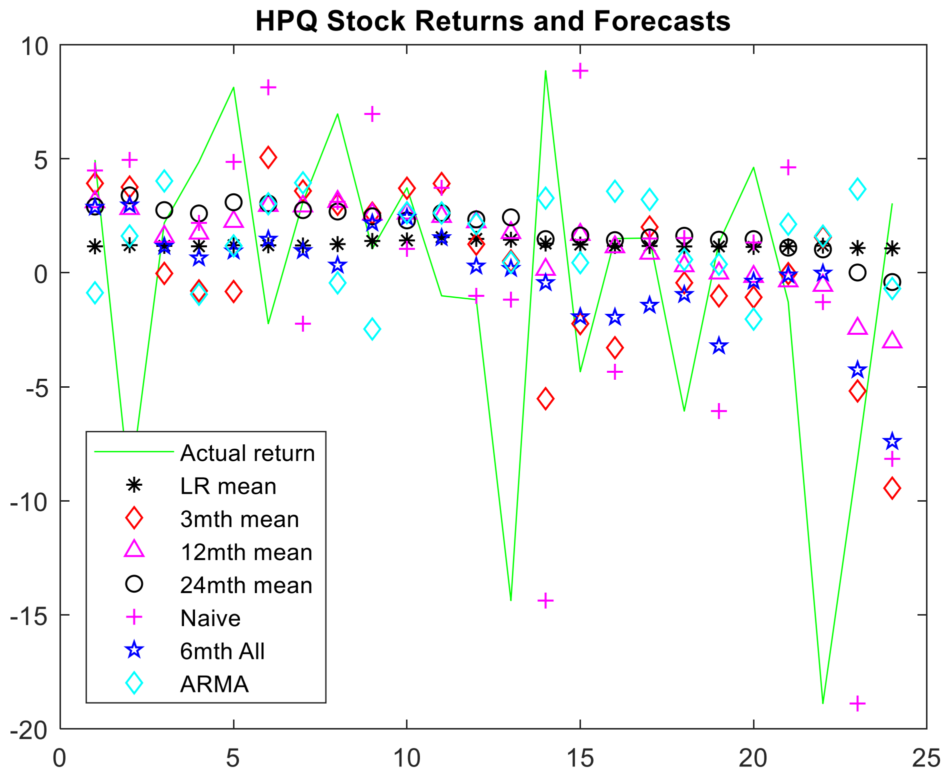
(b) For each asset return series, use the following methods to provide horizon 1 forecasts for the last 24 months of returns in the sample:

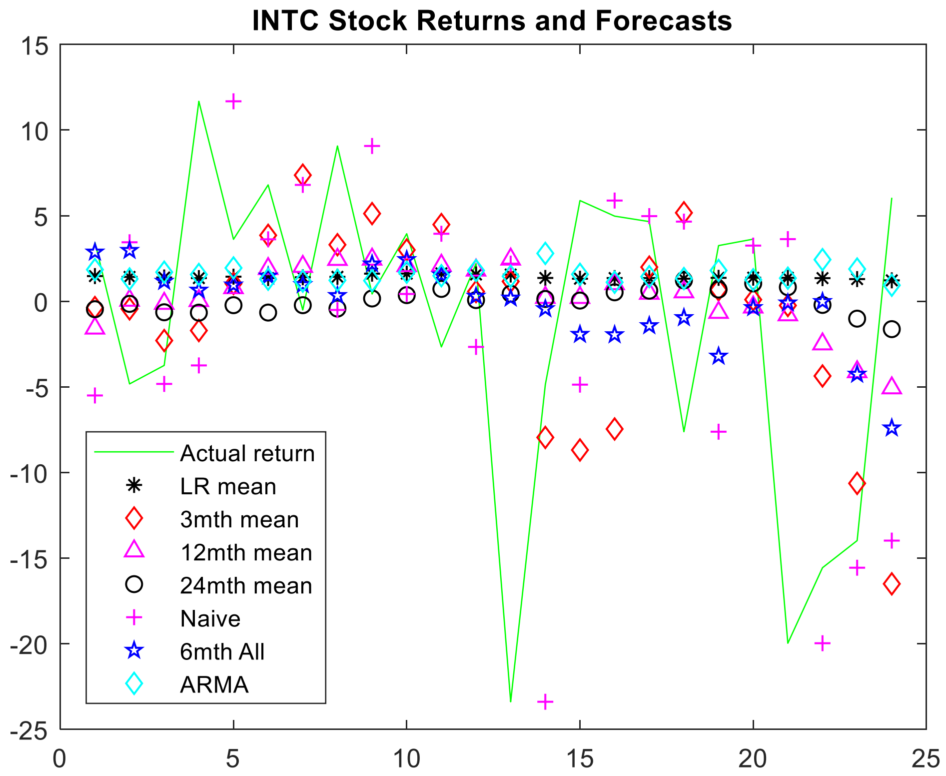
1. Long-run mean for that asset
2. Mean of last 3 months for that asset
3. Mean of last 12 months for that asset
4. Mean of last 2 years for that asset
5. Naive
6. One adhoc method that you make up yourself.
7. One ARMA model chosen for each series

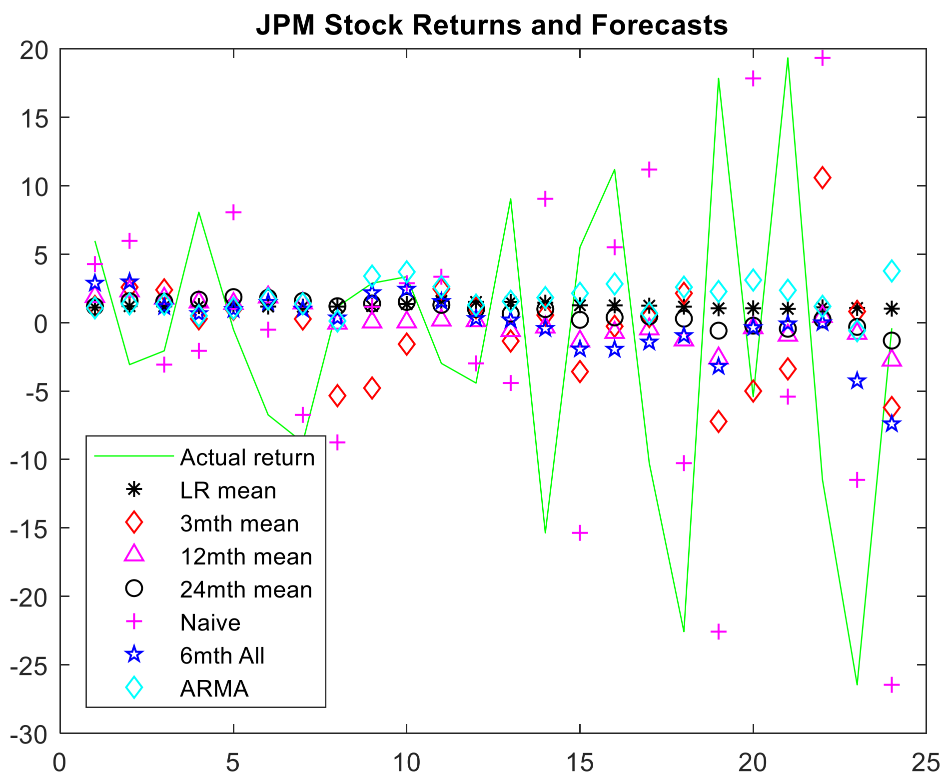
The code I wrote evaluates and calculates forecasts for these methods. For my 6^{th} method I chose the average return across all assets over the previous 6 months. There are literally millions of such strategies we could employ here. I chose quite a few, partly to communicate this to you clearly.

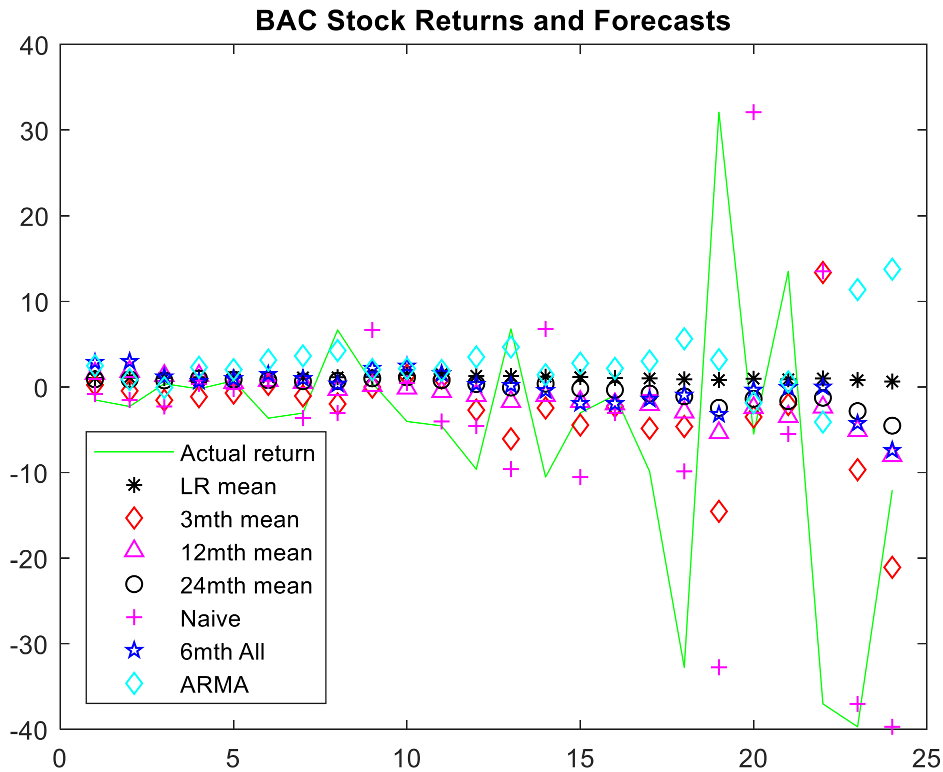


The symbols used for each method are the same in the plots below. None of the forecasts seem to "follow" the directions or magnitudes of the actual IBM returns they are trying to forecast. This pattern is repeated for all the other assets too.




INTC Stock Returns and Forecasts
15
10
5
0
-5
-10
-15
-20
-25
0
5
10
15
20
25
Actual return
LR mean
3mth mean
12mth mean
24mth mean
Naive
6mth All
ARMA


JPM Stock Returns and Forecasts
20
15
10
5
0
-5
-10
-15
-20
-25
-30
0
5
10
15
20
25
Actual return
LR mean
3mth mean
12mth mean
24mth mean
Naive
6mth All
ARMA



(c) Assess the accuracy of these forecasting methods using RMSE and MAD.

For IBM

	Mean	3 mth	12 mth	24 mth	naive	mean over assets	AR(4)
	1	2	3	4	5	6	7
RMSE	7.4037	8.0989	**7.2561**	7.5729	<u>9.1781</u>	7.7593	7.6178
MAD	5.4269	6.1751	**5.4123**	5.6249	<u>7.6163</u>	6.1224	5.4494

The units for these measures are the same units as percentage returns. The typical errors made are between 5% and 11% in terms of percentage returns. These seem very large! The best method, most accurate under both accuracy measures, is the 12 month average for IBM returns, followed by the 24 month average (RMSE). The AR(4) models ranks 2nd best under MAD and 3rd best under RMSE. The naive method ranks last.

For HPQ

	Mean	3 mth	12 mth	24 mth	naive	mean over assets	AR(7)
	1	2	3	4	5	6	7
RMSE	6.9568	7.9224	**6.8128**	7.0721	<u>9.5984</u>	7.0207	7.4201
MAD	4.9745	5.8034	4.9160	**4.9121**	<u>7.7031</u>	5.3174	5.8081

For INTC

	Mean	3 mth	12 mth	24 mth	naive	mean over assets	ARMA(1,1)
	1	2	3	4	5	6	7
RMSE	9.3333	10.4181	**8.9583**	9.2150	<u>11.4107</u>	9.3153	9.4945
MAD	**6.8075**	7.9009	6.8789	7.0164	<u>8.9456</u>	7.3503	6.9182

For JPM

	Mean	3 mth	12 mth	24 mth	naive	mean over assets	AR(4)
	1	2	3	4	5	6	7
RMSE	11.2844	13.2072	11.2778	11.2487	17.0078	**11.0187**	11.1615
MAD	**8.7453**	10.7233	8.9909	8.7634	13.2058	8.8261	8.7462

For BAC

	Mean	3 mth	12 mth	24 mth	naive	mean over assets	AR(9)
	1	2	3	4	5	6	7
RMSE	15.9958	17.3274	15.0590	15.3205	21.1524	15.3365	**13.0031**
MAD	10.6433	10.4629	9.5948	9.9289	13.0770	10.0568	**9.5882**

Since these are measures that are in percentage return units, these seven methods are all highly inaccurate at forecasting monthly returns for these assets. The lowest typical error is about 5% a month! The highest errors are of the order of 20% per month!

(d) As an experiment, find the returns that would have been made from investing under the following strategies:

1. *Equally weighted across all five assets, long 24 month holding.*
2. *Invest in the single asset, long and for 1 month holding, based on the forecast that gave the highest predicted return in each month, across the five assets. I.e. invest in the asset that had the maximum predicted return in each month. (Repeat this for all 6 forecast methods).*
3. *Weight the assets proportionally to their predicted returns, in each month.*
4. *Weight the assets inversely proportionally to their absolute predicted returns.*

Conduct these strategies for the last 24 months in the sample and examine risk and return performance.

These are all simulated investing strategies. They all use only data available prior to investment and hence generate exactly the returns that could have been achieved in real investing using these approaches.

Portfolios need to assign weights to the individual assets within them. There are many millions of ways to assign such weights. It is usually hard to beat the simplest weighting that applies equal weights (e.g. 1/5 when there are 5 assets) to each asset. The weights tell us how much each asset comprises of the portfolio, in terms of number of shares, and is also the respective contributions of each asset's return to the portfolio return.

$$r_{p,t} = \sum_{i=1}^{5} w_{i,t} r_{i,t}$$

Equal weighting sets each $w_{i,t} = 0.2$ = 1/5 when there are 5 assets.

One method of weighting is to assign a weight of 1 to one asset and 0 to the others. But these weights can change over time if we re-balance the portfolio during the period. The combination of these is what I have employed in strategy 2; where the asset with the highest forecast return next month has a weight of 1 and the others get a weight of 0. Since there are 6 methods to predict what the next month's return will be, I used them all and produced six variants of this strategy. This strategy is labelled "Max return pred" in the table below.

Strategy 3 sets weights at each time to be $w_{i,t} = \dfrac{\hat{r}_{i;t+1|t}}{\sum_{k=1}^{5} \hat{r}_{k;t+1|t}}$. This strategy allows for negative weights, which is effectively selling short an asset. Here we invest in each asset by an amount related to how much money we are predicted to make for that asset. Again, there are 6 methods for forecasting returns and I employed them all. This strategy is labelled “Wt prop pred” in the table below.

Strategy 4 sets weights at each time to be $w_{i,t} = \dfrac{1/\left|\hat{r}_{i;t+1|t}\right|}{\sum_{k=1}^{5} 1/\left|\hat{r}_{k;t+1|t}\right|}$. This could be seen as a risk-conservative strategy, since it weights assets higher of their predicted returns are closer to 0. This strategy is labelled “Wt prop inv pred” in the table below.

Strategy	Mean	Std
Equal weights	-1.7170	7.7655
Max return pred 1	-1.5834	9.5860
Max return pred 2	-3.7795	11.4434
Max return pred 3	-1.5329	9.0961
Max return pred 4	-0.7817	8.1310
Max return pred 5	-2.6841	7.4398
Max return pred 6	-4.9962	15.1487
Max return pred 7	-1.8774	14.9205
Wt prop pred ret 1	-1.7315	7.4450
Wt prop pred ret 2	**2.1422**	12.3123
Wt prop pred ret 3	-1.4799	39.8587
Wt prop pred ret 4	-5.8518	16.8359
Wt prop pred ret 5	-0.9064	20.1797
Wt prop pred ret 6	-1.7170	7.7655
Wt prop pred ret 7	-2.2022	10.0131
Wt prop inv pred 1	-1.6782	8.1511
Wt prop inv pred 2	-2.7591	8.1099
Wt prop inv pred 3	-1.5739	7.1981
Wt prop inv pred 4	-0.6396	6.9849
Wt prop inv pred 5	-1.8012	7.1149
Wt prop inv pred 6	-1.7170	7.7655
Wt prop inv pred 7	-1.8924	**6.8327**

(e) Comment on the results obtained regarding forecast accuracy and investment performance over all methods.

In three out of five assets (the tech stocks), by RMSE, the most accurate forecaster was the 12 month mean method. For the financial stocks the ARMA models had the lowest RMSE and MAD. However, these RMSE figures were well above the standard deviations obtained in the in-sample period (see (a) above). The 24 month mean also did well across the assets regarding accuracy. In fact, there is usually very little separating the long-run, 12-month and 24-month means and the ARMA models in terms of accuracy.

Clearly smoothness is important. This is likely because return directions are very difficult to forecast and so methods that are smoother sit in the middle of the data, while methods that jump around tend to make larger errors (since about 50% of the time they jump the wrong way!). The worst method regarding accuracy is always the naive. This would do well if the returns were behaving like a random walk over time. Since they instead are mean stationary, with a mean close to 0, this method does very poorly in accuracy.

Regarding profit and risk, since the investing period is in 2008 and includes some GFC period, none of the strategies has done very well. In fact the results are highly variable across the 6 forecast methods and the four different weighting methods.

Some possible points to note:

1. It's hard to beat equal weighting. This is very, very often true in real data.
2. The range of performance is quite scary! These are only very simple investment methods and forecast models, but they could be disastrous!
3. The methods that did the best in terms of accuracy don't necessarily do the best in terms of risk or profit. In fact, the worst forecaster re accuracy, naive method, is one of the best in terms of risk and return over this period (across the weighting schemes).