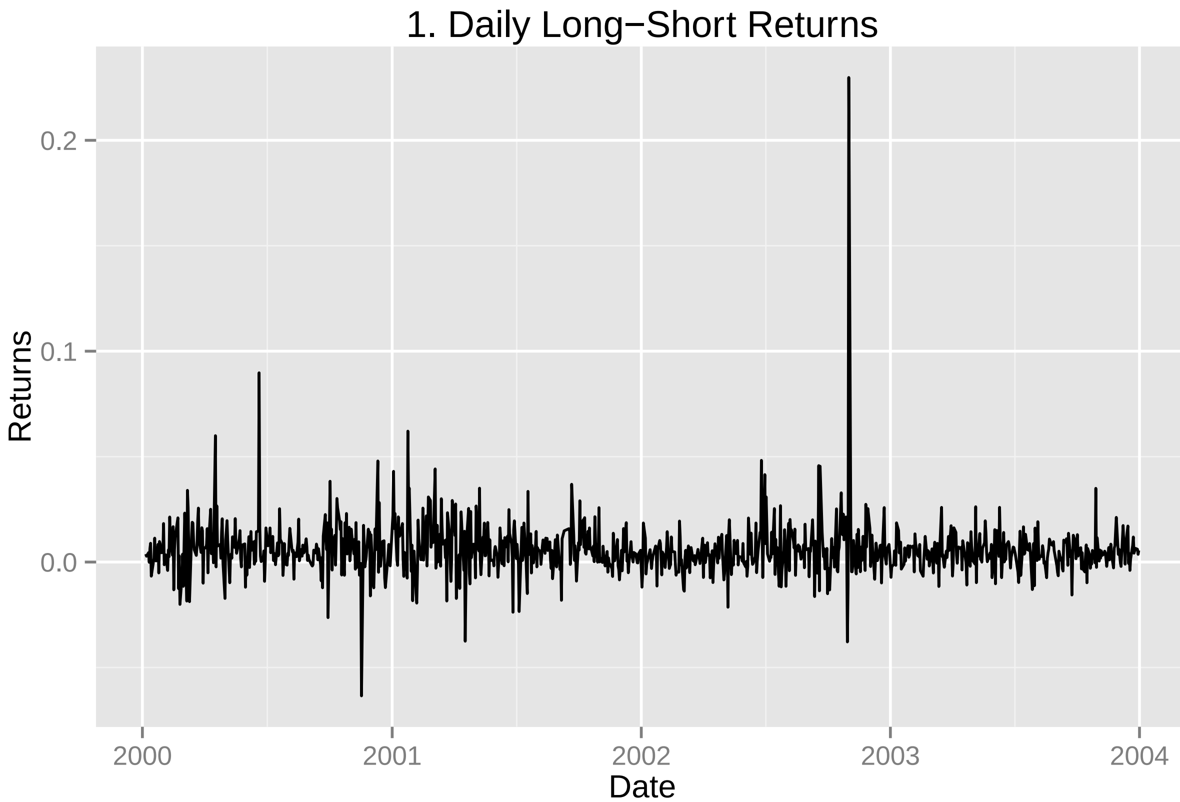
Solution Description - Part (2)

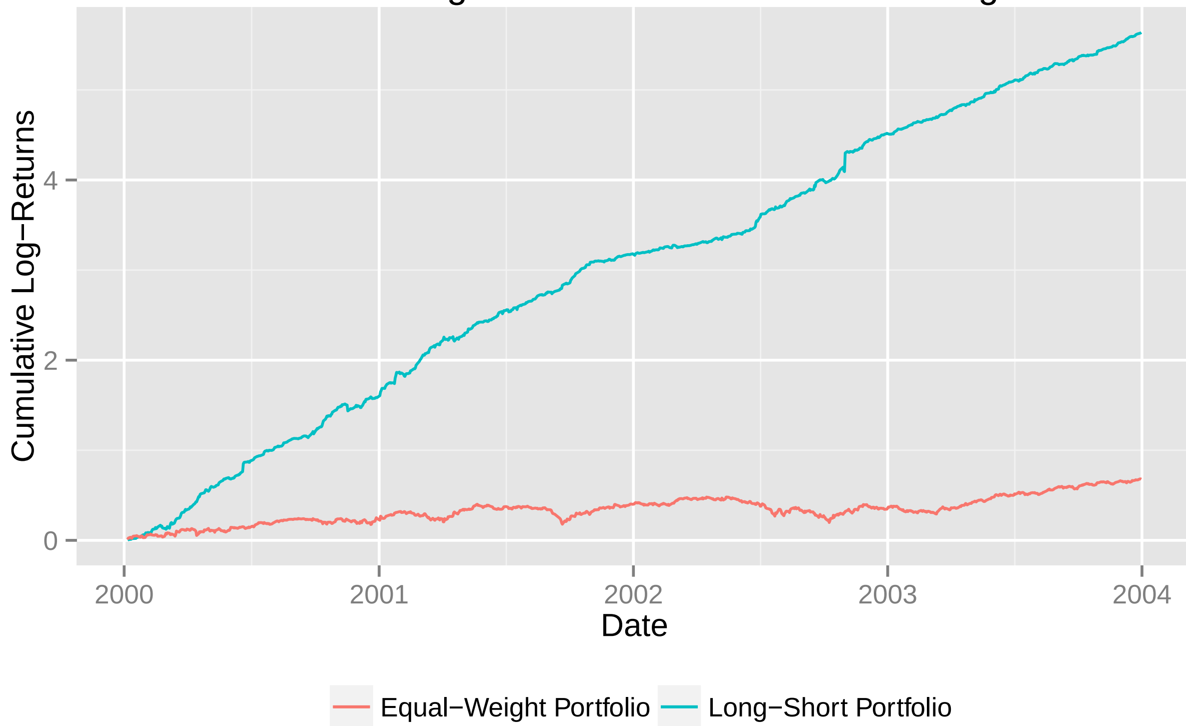
Team A - Mufan Li, Mengye Ren, Tian Xia

March 19, 2016

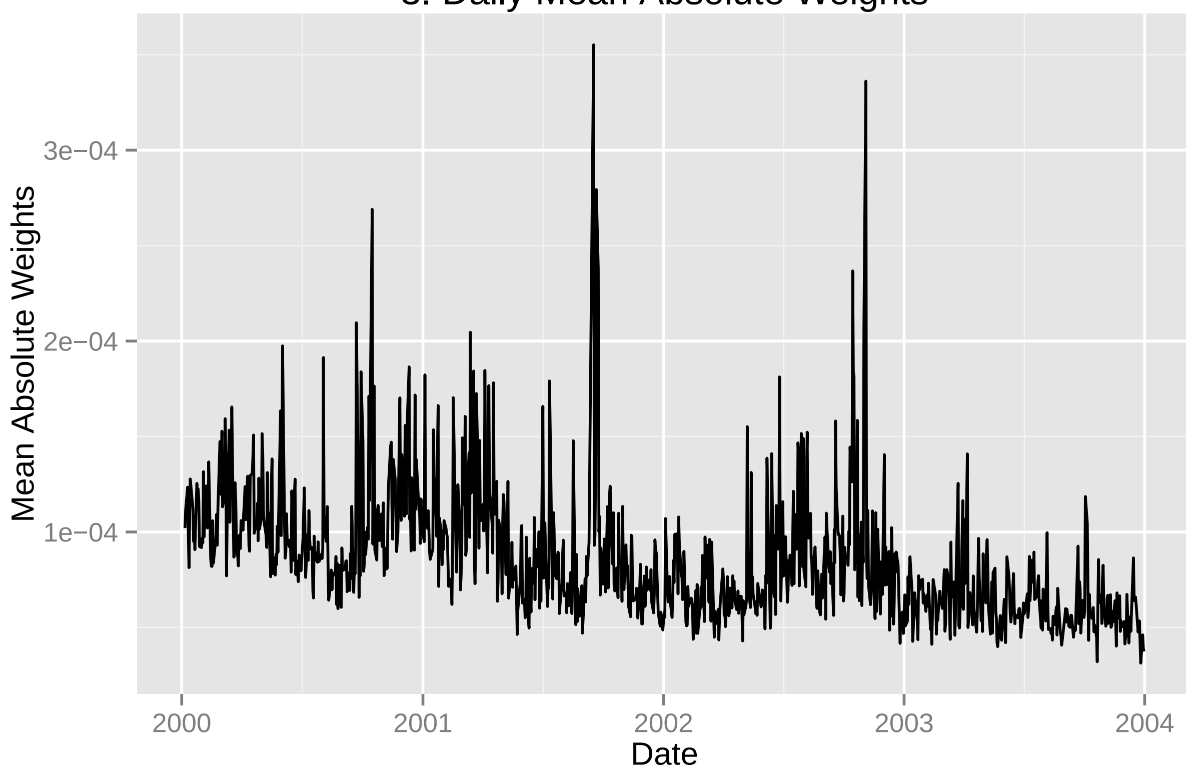
We first display the four time series plots.

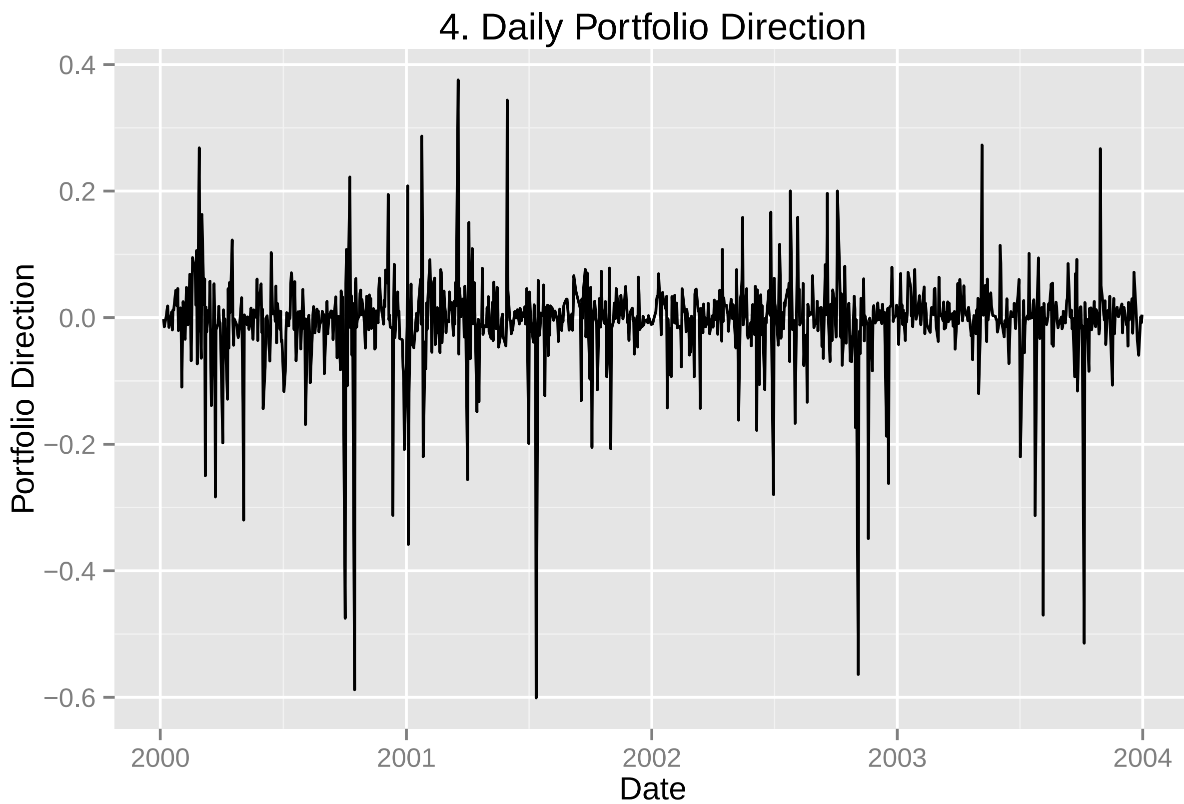


2. Cumulative Long−Short Returns − In Natural Logarithms



3. Daily Mean Absolute Weights





We can also look at the relevant statistics below.

	Names	Values
1	Average Daily Log Returns	0.0056
2	Standard Deviation of Daily Log Returns	0.0125
3	Annualized Sharpe Ratio	7.1674
4	Skewness	4.5924
5	Kurtosis	70.3278
6	Maximum Drawdown - Number of Days	4.0000
7	Maximum Drawdown - Return	-0.0745
8	Correlation with Equal Weighted Returns	0.0027

Table 1: Summary Statistics - Using In-Sample Data

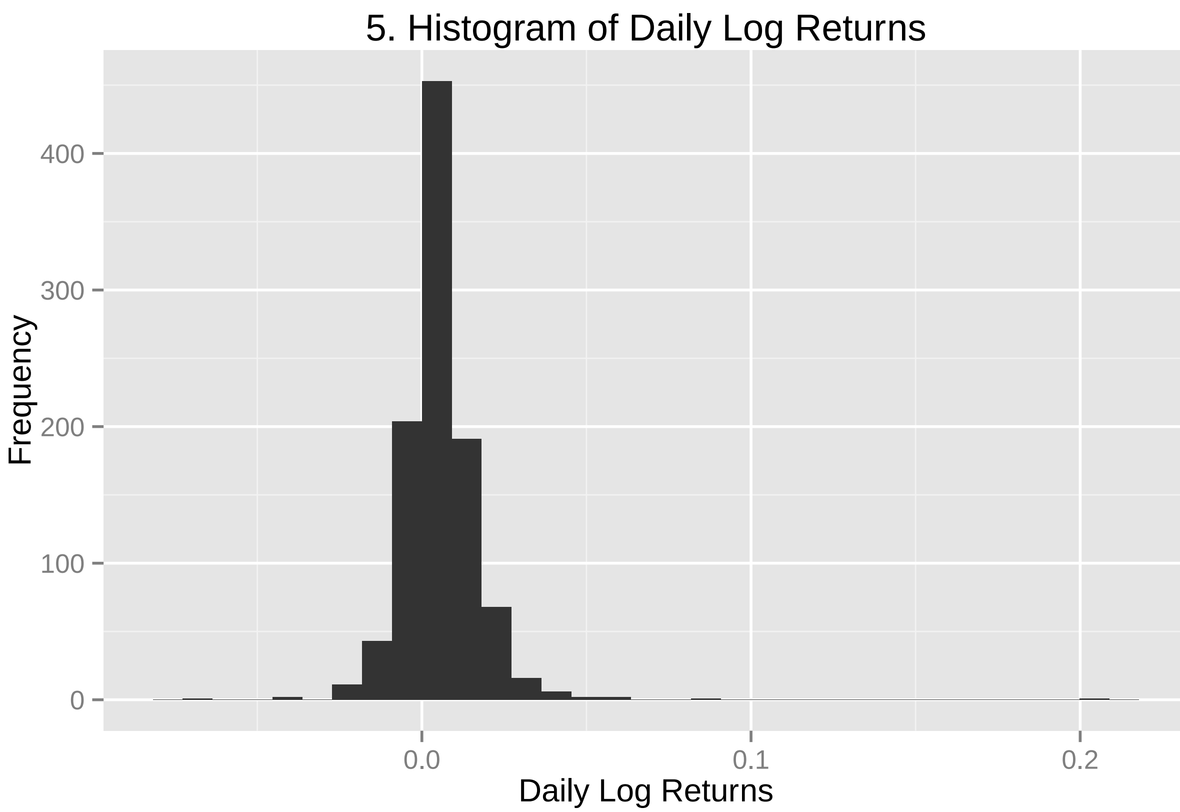
To find the parameter values, which are shown in Table 2, we tried both gradient descent and random line search algorithms in MATLAB. Since we found multiple different local minima, the problem is clearly non-convex; however all of the Sharpe ratios are fairly similar, so it is not a significant issue in this problem.

Judging from in-sample data, the strategy is performing quite well. We first observe plot 2 of cumulative returns and the annualized Sharpe ratio, both of which are significantly outperforming the strategy in part (1) and the equal weight portfolio. From a correlation of only 0.0027 with the equal weight portfolio, and observing plot 4 this strategy rarely has a portfolio direction of more than 0.2, we can see that the strategy is very market neutral. We also note the maximum drawdown for this strategy is extremely low at 4 days and -7.45%.

The only negative sign we observe from in-sample statistics is the large skew and excess kurtosis, implying extremely heavy tails. However, but looking a histogram below (plot 5), we realize that there is data point returning 20%. Removing that point, we would have a skewness of 0.675 and an excess kurtosis of 7.214

	Parameter	Value
1	a1	-0.051209
2	a2	-3.885165
3	a3	4.078397
4	a4	2.700011
5	a5	0.095185
6	a6	0.856144
7	a7	-0.966253
8	a8	-0.856446
9	a9	-0.114941
10	a10	-0.039731
11	a11	0.115225
12	a12	0.086389

Table 2: Parameter Values - Using All In-Sample Data



To check on performance of this strategy, we separately optimized a set of parameters with last 80% of the data as in-sample, and tested the performance on the first 20% as out-of-sample. We specifically avoided using the last 20% of the data as out-of-sample as the equal-weighted portfolio showed a significant upward trend, making it a terrible dataset to validate our strategy.

The summary statistics are shown in Table 3. While further validation on this strategy can be done with different out-of-sample datas, we can already be quite confident about the performance given out-of-sample Sharpe ratio of 7.24.

We also observe the new weights in Table 4, when trained only with in-sample data. While the values are quite different, we can achieve similar results. Therefore the local minima is not causing a significant problem for us in the part.

Here we observe the form of the weights is decided by a linear combination of 12 factors. All of these factors

	Names	Values
1	Average Daily Log Returns	0.0063
2	Standard Deviation of Daily Log Returns	0.0137
3	Annualized Sharpe Ratio	7.2407
4	Skewness	1.4037
5	Kurtosis	14.8634
6	Maximum Drawdown - Number of Days	2.0000
7	Maximum Drawdown - Return	-0.0733
8	Correlation with Equal Weighted Returns	0.0056

Table 3: Summary Statistics - Using Out-of-Sample Data

	Parameter	Value
1	a1	-0.669108
2	a2	-1.362873
3	a3	2.261092
4	a4	-12.114364
5	a5	0.877295
6	a6	2.004053
7	a7	-2.912369
8	a8	-2.267258
9	a9	-0.722039
10	a10	0.701638
11	a11	-0.007674
12	a12	-1.216179

Table 4: Parameter Values - Using 80% of In-Sample Data

are more less only dependent on the previous day's data. Once the parameters $\{a_1, \dots, a_{12}\}$ are fixed, we are assuming a pattern to persist between the previous day and today's data. In other words, we are assuming the conditional distribution $P(R_{OC}(t,j)|\mathcal{F}_{t-1})$ is stationary, where $\mathcal{F}_{t-1}$ is all the information up to the previous day. However, we can expect this assumption to fail in some future time such as a period of extreme volatility, especially since the in-sample data given to us look fairly stationary.