Succeeding in algorithmic trading using a simple framework: mean-reversion and momentum selection strategy analysis

Max Saparov

max.saparov@nyu.edu

ABSTRACT

This paper uses a simplified approach to create a trading algorithm through the combination of mean-reversion and momentum portfolio selection strategies. The basis of measuring the success of this strategy is by comparing its relative Sortino and Sharpe ratios to a benchmark ETF (QQQ). We find evidence that this methodology may be used to generate positive excess returns with relatively less risk when compared to the overall market.

Keywords: Stock market, Algorithmic trading, Mean Reversion, Momentum, Portfolio selection

INTRODUCTION

In a field defined by unbounded complexity, the search for a robust method to outperform markets and funds - both in absolute terms and with minimal risk - is a tedious one.

The purpose of this paper is to explore how a radically simplified approach to algorithmic trading can be used to build up a robust strategy. By beginning with a clear cut, general model it becomes easier to show its predictive power and reduces the ways in which over-fitting parameters may influence live results. The basis of the methods used in this paper is that predicting price - or even just the price direction - of any single stock is too difficult to achieve in any reasonable time frame. Instead we will focus on the general trends of a given universe of stocks. For the purpose of this paper, we will define our universe as the 100 stocks that make up the NASDAQ100 and use the QQQ ETF as a benchmark for performance.

Note that when first creating this strategy, data was split into training and testing sets in order to avoid a subconscious over-fitting of parameters and data processes based on visual results. To reflect this original methodology,this paper will follow a similar presentation style and have separate sections for the training and testing data sets. The training data set are daily candles of 83 of the stocks in the NASDAQ100 (as of Jan. 2021) from January 1, 2014 to December 31, 2018. Similarly, the testing data are the daily candles from January 1, 2019 to January 1, 2021. While it would be preferable to use all 100 stocks in our universe, the 83 stocks are arbitrary and were the ones with the best data in this time period.[1]

All back-testing and analysis is custom built in python using Pandas[2] and Alex Golec's api wrapper for TDAmeritrade's API.[3]

The strategy used consists of two main parts: ordering the stocks in our universe and then choosing a stock and a strategy based on its ranking. In order to better show the generalized methodology used to build similar simple yet robust strategies, we will explain the logical steps taken to build up this specific approach. Thus, section 1 will focus on how the ranking process works and an explanation of its predictive power. Section 2 will go in-depth into how this predictive ranking can be used to form a profitable trading strategy by picking an optimal method from a set of strategies. Finally, in section 3 we will analyze our results with commonly used metrics and show a final test on our testing data.

PREDICTIVE RANKING USING A SIMPLE CORRELATION

This ranking method was inspired by a Reddit thread(link this) on using the most simple metric possible in order to build a profitable trading strategy. Namely, the metric explored in this thread as well as in this paper is that the price change of a stock today is correlated with the price change tomorrow.

$$\text{Percent Change}_{future} \propto \text{Percent Change}_{today}$$

Using this metric we can create a ranking from 0 to 82 with 0 being the biggest loser for the day and 82 being the biggest gainer.

Showing Predictive Power

By tracking the day-to-day ranking, we can see how the ranking today is reflected in the ranking tomorrow. Figure 1 shows how the bottom four rankings will tend to stay on the extremes of the ranking - namely that tomorrow they will either stay the biggest loser or become the biggest winner. Similarly, Figure 2 shows the same relationship when we start at the other end of the ranking - the biggest winner today will either stay the biggest winner or become the biggest loser. However, in figure 3 we see that this is not the case for stocks that fall in the middle of the ranking, as their next day rankings are much more normally distributed.

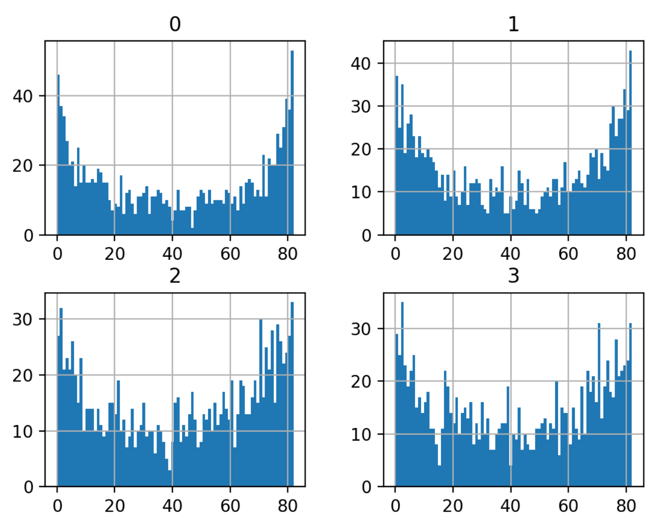

Figure 1. Relation showing how often a certain ranking followed the bottom four rankings on the next trading day

While it is possible to say that the results depicted in figures 1 and 2 may be normally distributed around the extremes as if the chart had been wrapped around itself, the two most extreme rankings have clear peaks at the edges that show a high likelihood for the stocks in the two most extreme rankings to stay as either the biggest winner or become the biggest loser in the following day.

Thus, we have shown that while this ranking can not predict the direction of stocks by itself, it does have a predictive power over the following day ranking which we are able to exploit due to an inefficiency in the market that we will describe in the following section.

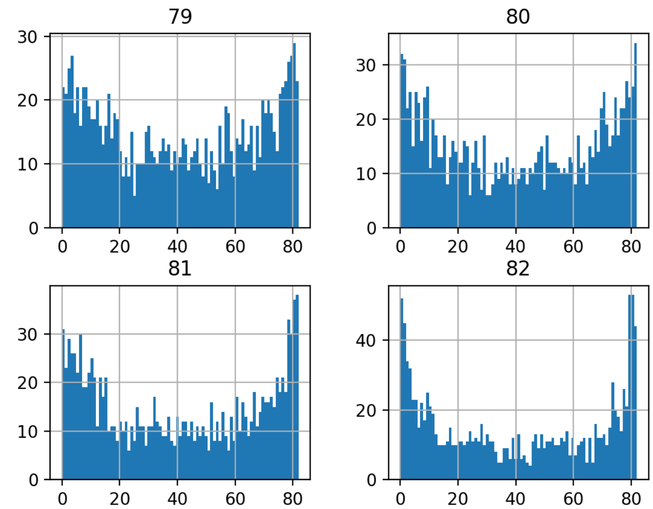

Figure 2. Relation showing how often a certain ranking followed the top four rankings on the next trading day

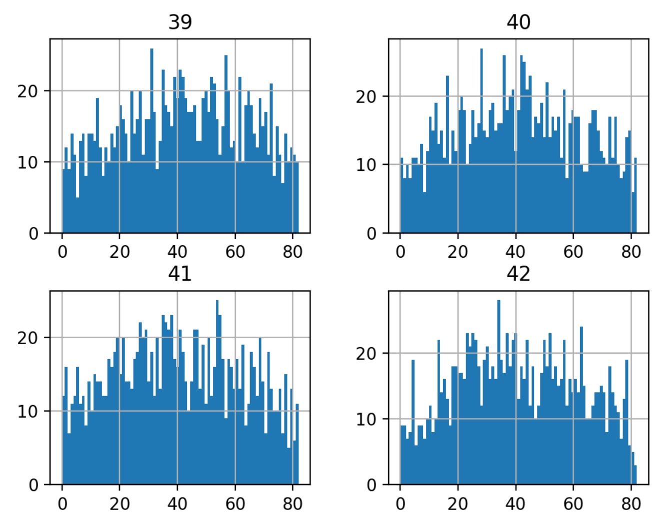

Figure 3. Relation showing how often a certain ranking followed the middle four rankings on the next trading day. Much more normal and evenly distributed

OPTIMIZING A PROFITABLE STRATEGY

In the world of perfect efficiency it would not be possible to predict the direction of these extreme picks. However, we can exploit the market's tendency to engage in crowd behavior as described by the Adaptive Market Hypothesis (Lo, 2004) - or AMH for short - to show that the direction of this swing is periodic. The AMH combines ideas of behavioral finance to provide a model for traders that recognizes that most market decisions are by humans rather than perfectly rational decision makers. Lo (2004) goes on to explain that one of the biases of human traders is to flock due to the belief that other traders are smarter or have access to better information.

Analyzing General Mean-Reversion and Momentum Strategies

We can visualize the trends of the extreme picks by creating four simple strategies where we simply hold two stocks for a single trading day:

1. Always going long on the biggest loser (ranking 0). Expecting the stocks to revert to their mean and increase in price.
2. Always going short on the biggest loser. Expecting the stocks to have momentum to decrease in price.
3. Always going long on the biggest winner (ranking 82). Expecting the stocks to have momentum and continue increasing in price.
4. Always going short on the biggest winner. Expecting the stocks to revert to their mean and decrease in price.

Note that these four strategies fall into two simple predictions - expecting the price to revert to its mean or expecting the price change to have momentum and continuing in the same direction. Figure 4 shows how all four strategies perform on their own in our training period with none of them performing consistently well on their own.

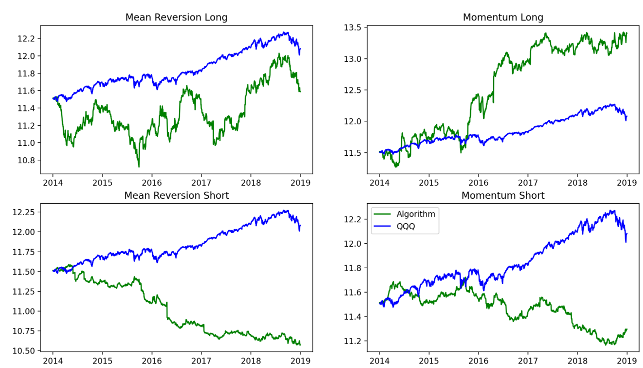

Figure 4. Four Mean-Reversion and Momentum and strategies compared relative to the QQQ ETF. Plotted with respect to log of returns with an initial balance of $100,000.

It is interesting to notice how these strategies appear to trend - as if all of them periodically work, whether for a month or for a year. When the Momentum Long strategy (betting on the two biggest winners) seems to stop being profitable in early 2017, the Momentum Short strategy begins to have consistent returns; all the while the QQQ benchmark continues to be in a bull run.

This is most likely an inefficiency in the market caused by a combination of effects. Firstly, any trading strategy can be generalized to either being a mean reversion or a momentum strategy. Regardless of the methods used, the trader either believes an equity is overvalued and the price will decrease or it is undervalued and the price will increase. Combining this with the Adaptive Market Hypothesis expectation that traders will flock - we can predict that certain general strategies will become more popular for periods of time, even without direct collaboration between traders.

Picking an Optimal Strategy

Assuming that at least one of these simple strategies is consistently profitable for a moderate time period we can begin to consider methods to pick the one that has the best performance in a chosen look back period. In continuing to keep the used methods as simple and robust as possible, we will explore using the median return of each strategy in our look back period. That is, taking the returns for each day in the last n days and using the median return as our expected future return for the day. Then we can simply use the strategy that has the best median returns.

Figure 5 shows an example in our testing period with a look back period of 15 days with its respective metrics shown in table 1. When comparing the performance of our strategy we will always show its metrics relative to the QQQ (our benchmark) in the same period - as tying specific values to the ratios would be fairly arbitrary and it is more useful to see how our strategy did relative to its benchmark.

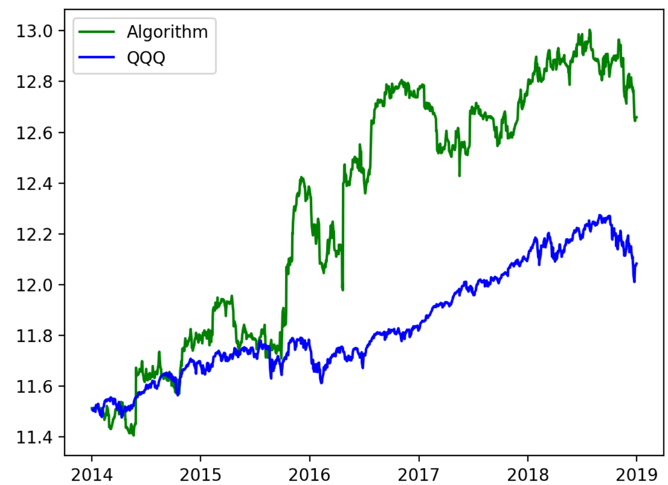

Figure 5. Algorithm that uses one strategy of the four described earlier by picking the one with the highest median returns with a look back period of 15 trading days. Shown with respect to the log of returns with an initial balance of \$100,000

Metric	Algorithm	QQQ
Sortino Ratio	2.488	1.192
Sharpe Ratio	0.837	0.765
Daily Success Rate	52.104%	55.564%
Max Drawdown	35.976%	23.155%
Beta	0.046	1
R^2	0.014	1

Table 1

This specific strategy appears to outperform the QQQ benchmark when using the Sortino and Sharpe metrics with only slightly worse daily success rate and max drawdown. We can test for overfitting by running this same test with varying look-back periods and we see that for periods of ± 1 day we get similar results. However, the balance between having enough data for our median to be a significant indicator and not having too much noise seems to be a very tight one - do note that we are using a very rudimentary signal in this analysis and more sophisticated methods would likely prove to be more effective.

Both our R^2 value and Beta show that our strategy is uncorrelated with market conditions. This low correlation may show promise for consistent results in the future, however this is difficult to say as it is unclear what conditions are necessary for this specific algorithm to be successful. The maximum drawdown of this algorithm is significantly larger than that of our benchmark - this is most likely due to the strategy inherently betting on the most volatile stock of the day which can cause larger dips. However, both the Sharpe and Sortino Ratios adjust returns for portfolio volatility - both of which our algorithm outperformed the benchmark in.

TESTING RESULTS

To confirm our results, it is important to run the final version of our algorithm on unseen data without changing any parameters. This will allow us to see the robustness of our strategy as well as see what occurs during the market crash in the spring of 2020.

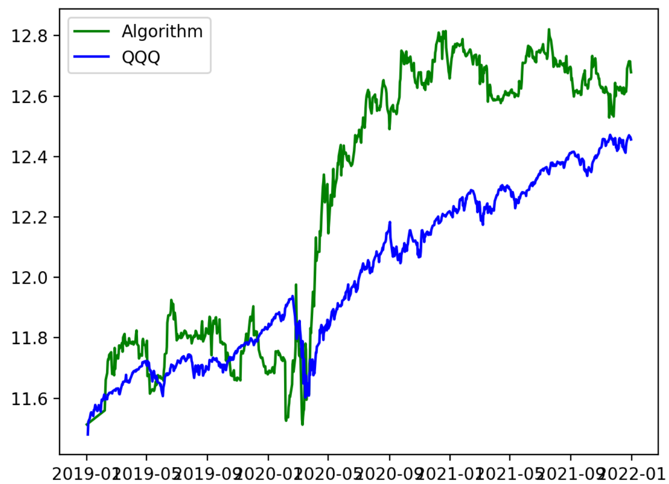

Figure 6. Previously used algorithm with a new timeframe from January 1, 2019 to December 31, 2021. Shown with respect to the log of returns with an initial balance of $100,000

Metric	Algorithm	QQQ
Sortino Ratio	2.898	2.151
Sharpe Ratio	1.123	1.387
Daily Success Rate	54.830%	59.313%
Max Drawdown	37.120%	28.559%
Beta	0.055	1
R^2	0.013	1

Table 2

We once again see that our strategy outperforms the QQQ index using the Sortino Ratio metric, however it no longer outperforms using the Sharpe. The rest of our metrics stay fairly consistent, which likely represents the stability of the algorithm, but more testing is recommended.

Concluding Remarks

This strategy does appear to show an inefficiency in market pricing - albeit a small one. This inefficiency probably exists in the short interday time period - high frequency traders have certainly made intraday trading highly efficient and longer term traders are subject to more macro, unpredictable events - instead, the time frame of this strategy allows for it to find an inefficiency in the human psyche, giving the possibility for more psychologically founded strategies to have an edge over the market. As trading becomes more prevalent in many US households, it is possible for this edge to become more influential before it is better understood. Thus, by continuing to look into similar inefficiencies we may gain a better insight into the clash between human nature and financial economics.

REFERENCES

Lo, A. (2004). The adaptive markets hypothesis: Market efficiency from an evolutionary perspective.

1. https://github.com/Vorapas/Mean-Momentum-Research/blob/master/stock-list-2014
2. https://github.com/Vorapas/Mean-Momentum-Research
3. https://github.com/alexgolec/tda-api