Active Management Strategies

Varun Jauhar & Wyeth Thompson

Increasing the returns of an investment portfolio is the goal of individuals and institutions all across the world. Beating the market and increasing profits is considered to be a very challenging task, due to volatility, the theory of efficient markets, and other investors that will seize any available opportunity to increase their returns. Due to these perceived impediments to achieving excess returns, many equity investors have begun moving their money into passive strategies rather than paying institutional investors like hedge funds to actively manage their money. Passive fund managers generally charge lower fees than their active counterparts, and therefore, investors will choose to invest in passively managed funds if they don't believe that active managers are worth the premium that they charge. The growth of passive investing has taken off in recent years, especially due to the popularity of exchange-traded funds (ETFs), with the number of assets, for American public equity-based funds, that passive management accounts for increasing from 25% to 45% in the past 10 years.

This trend will lead to changes within the market and the economy. Although passive fund managers are surely ecstatic about their gain in popularity, many investors and economists are growing increasingly concerned with what this increase in passive investing will mean for markets. Risks that this shift may pose deal with the potential for increasing market volatility and the potential for bubbles to form due to market inefficiencies.

The market is changing rapidly and although we do not know what the long-term effects of the move towards passive investing will be, we were influenced by it to study some trading strategies, through analyzing historical stock data. We were able to find a dataset showcasing the complete daily trading history of every publicly traded equity on the New York Stock

Exchange, the NASDAQ, and the American Stock Exchange, on Kaggle. This dataset includes individual files for each company, which showcase the date, ticker, volume, open price, close price, and adjusted close price for each row within the csv files.

Active investing may be on the decline for a variety of reasons, which was our inspiration for wanted to look at the historical data that we were able to find to test commonly used active investment strategies. The first of these strategies is momentum trading. Momentum trading is a common strategy used by hedge funds and other institutions that trade public equities. This strategy relies on the intuition that stocks that are performing well will continue performing well in the future and stocks that aren't performing well will continue to do so in the future.

We tested our own version of this trading strategy using the historical stock data. We did this through creating an equally allocated portfolio of stocks that were purchased based on their performance over the past 2 days. If a stock had gone up by two percent or more over the last two days, shares of that stock would be purchased at the start of the next trading day. The hope being that there will be a sort of inertia with these stock gains and the stock will continue to go up. If the stock lost 2 percent of its daily open price within a day, the entire position in that company would be closed at the start of the following trading day. This was done to protect our portfolio's downside, with the fear that the stock's price would continue going down or would undergo a correction.

The second strategy that we decided to test was the very popular, "buy low sell high" strategy. In this strategy, we did the opposite of what we did in the momentum strategy, buying stocks that had gone down by 2% or more for two consecutive days, while closing the positions as soon as the stock went up 2% in a given day. The intuition behind this strategy is that certain stocks may become oversold and thus undervalued. Investors using this strategy would like to purchase stocks that have been on the decline and will sell them as soon as their price increases by a set amount. We chose our specific thresholds for buying and selling for two main reasons. First, because these investment points fall within the underlying thought-process of the “buy low sell high” strategy and thus testing our own version of the strategy was a way to test the intuition behind the strategy as well. Second, we intentionally made the investment buy and sell points the opposites of our momentum strategy so that we could directly compare both of these strategies and the merits of the thought-processes behind each of them.

The results of each trading strategy were compared to a benchmark, the S&P 500. This index has performed very well over the last century and is a commonly used point of comparison for investors. Daily data for the S&P 500 was collected from Yahoo Finance for the purpose of comparing its returns to the returns of our trading strategies.

We tested these strategies on our entire database, which we treated as a portfolio so that we could see the overall effects of our strategy, as used on the market over a long time period. We used the largest sample size of time and quantity of stocks that we could find in order to ensure the accuracy of our results.

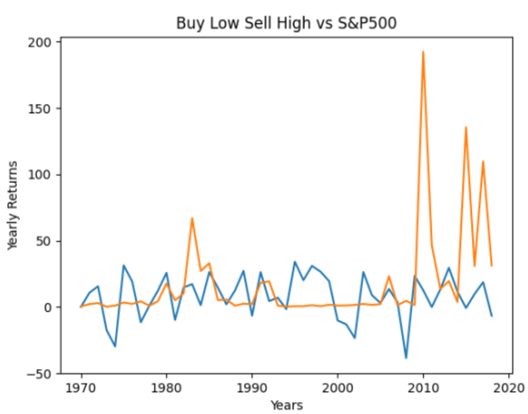
For our methods we used many different commands and packages that we have learned throughout DS2000 class. The nature of our project involved using large csv files with multiple columns, so we found that pandas data frames were the most effective way in dealing with the large files. When looking at our code, the first step we took was iterating through all of our company stock information and making buy and sell decisions based on the daily gains. By looking at the all_company_data function in the all_company_iterator.py file you can see this operation. For this function we used for loops to iterate through the individual stock information to find buy cases and sell cases. After purchasing ten dollars’ worth of a company stock based on the buy case, we would hold that stock in a list until the sell criteria was met. At the time of the sale we would then append the sell date, investment dollar gain, and investment amount to a dictionary which could then be converted into a data frame. One method we found very useful in this operation was keeping track of the investment amount. This meant that our function could purchase stock for multiple days in a row and it would keep track of the number of purchases in that individual transaction. After iterating through each company's stock information, we would then append the data frames into one larger data frame which contained all of the investment information of the entire strategy. Another helpful method that helped us perform this function was a list of ticker names provided by the data source, all_symbols.txt. This allowed us to make sure we accessed all of the company specific csv files in the dataset.

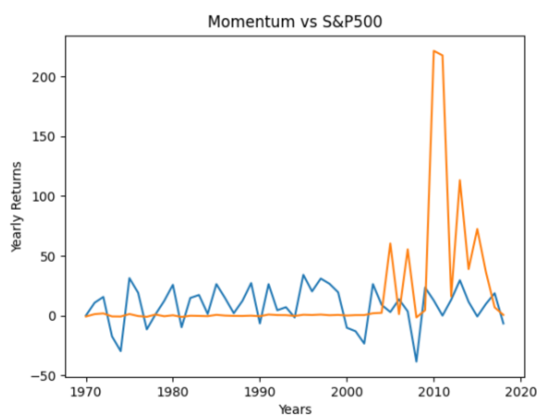
While iterating through the large folder of company data, we were confronted with an index error which we struggled with for several hours. After testing, we found that if our function tried to buy a stock at the very end of the dataset it would create this error because the purchase date would lie one location further than the length of the data frame. Our solution to this problem was to decrease the length of the iteration by two values, which would make the function stop before it reached the end. We found that this was the most logical way to approach this problem as we felt that the alternative of selling all of our held stock on the last day would contradict our investment strategy.

The function ran for about two hours for both strategies, after which we downloaded the data frames as csv files. CompanyReturns.csv is the file for our initial momentum strategy and BuyLowReturns.csv is the name of the file for the second strategy.

After running our investment strategies, we had to figure out how to present this data in a way that we could compare our results to S&P performance and aggregate the returns to the entire portfolio. Our method for this step was to use the portfolio_yearly_returns function in the DS2000ProjectCode file. In this function, we were able to use the filter command in pandas to extract each year containing all of the investment information for all the companies in the given year. This command proved to be very helpful to us as it allowed us to create sub data frames and compute the yearly percentage return based on the sum of the investment amount of all the companies. Combining this step with a for loop we were able to extract yearly gains from each year in our measurement timeframe.

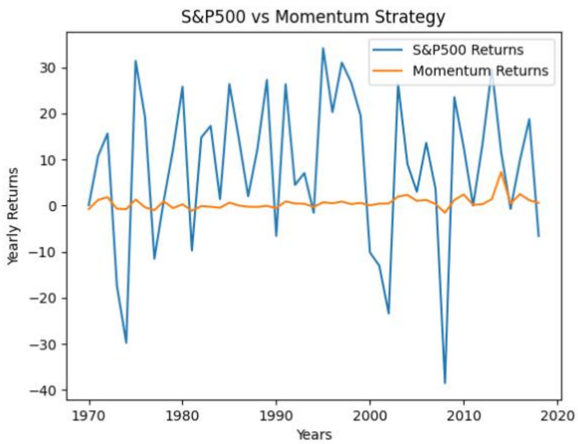
We also used the filter command on our S&P data in the sp_yearly_returns function which allowed us to measure yearly percentage gains based on the first open and last close price of that year. Both the S&P and portfolio strategy yearly returns were then converted into data frames and used in our plots.

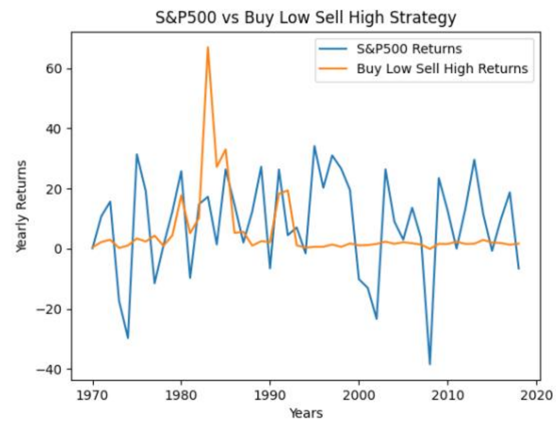


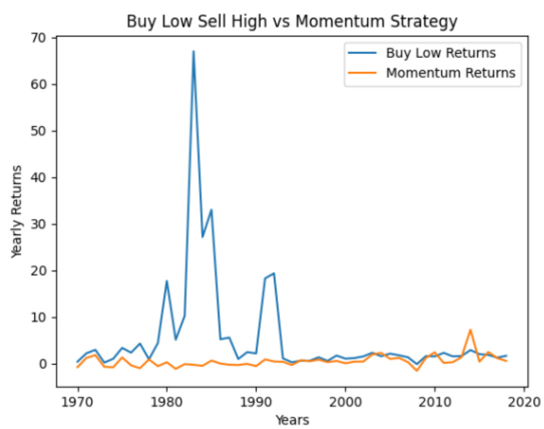


The above figures show the plots of each strategy compared to the s&p500 (blue). As you can see, we experienced major jumps in the 2010s which we later found to be irregularities in our basic company stock information which caused massive percentage gains on $10 stop investments. These results were obviously incorrect so we decided eliminating outliers would be the best method for dealing with these erroneous data points. Our method for dealing with the outliers would be to use the scipy package and eliminate outliers using z scores. Inputting our large full portfolio investment results into our outliers function we could assign z-scores to all of the percentage returns for the individual stock sales. We then only kept rows with z-scores of less than 3 and appended them to a new data frame to be graphed again.

The resulting plots show our final results without outliers:







Find the original png files in our project submission.

The results of our project are shown above. Both of the strategies varied much less in terms of yearly returns compared to the S&P500. For our momentum strategy, we attribute this to our selling decision which is intended to protect our downside when stocks are plunging into the negative. Otherwise, we think that the relatively low variance is caused by our use of so many stocks, some of which worked with the strategy while others did not. The result was a return close to 0 for many of the years measured. While this is not necessarily what we were hoping for when beginning this project, it is an interesting result, as each strategy has shown to vary less than the market which can be useful when managing risk in today's investment environment. Another significant detail of both of the graphs are spikes occurring in the momentum strategy during the mid 2010s and to a higher degree for the buy low strategy in the 1980s.

These results show that our investment strategies can be productive when compared to the S&P during particular time periods. While our project did not include analyzing the cause of these particular instances due to time constraints, we predict that during these two periods our strategies could have been novel and produced higher returns in the market. However, given the way the market adopts to new information and new investment techniques, one could understand how after these strategies are used on a widescale they will likely produce less and less returns on average. Based on the simple nature of our buy/sell decision, it is clear that not enough information could be taken into account to consistently beat the market using such a large range of companies.

Some strengths of our analysis were that we used a large dataset, which included daily information for almost 50 years for thousands of stocks, and we were able to consolidate this information into a line plot to visualize our results against a benchmark. Using this large of a sample size, while also taking outliers into account, we believe that our results were relatively accurate and reflective of the efficacy of each strategy. The large timeframe of results allowed us to see the effects of our strategies in many different time periods. We were also effective in manipulating the data that we collected, which allowed us to visualize such a large amount of data at once. Additionally, we were sure to always stay consistent with the numbers that we were comparing in our analysis, which adds to the integrity of our results.

There were many weaknesses in our project, many of them due to the time constraints of completing it. We did not provide any statistical analysis of the significance of any of our results, with the exception of the exclusion of outliers from our results. Additionally, our results are contingent on us having been able to invest exactly 10 dollars each time we made an investment, which is not actually plausible in the real world. Further, we did not account for any additional costs on top of our investments, such as bid-ask spreads and fees. Our investment strategy does not support intraday trading and relies on investment decisions made once a day based on historical data, which is definitely a limitation to the upside of our results and is not reflective of the full-range of decisions that investors have at their disposal.

If we had more time to work on this project, we would have liked to use our strategies to compare their results in different situations. Analyses on how these strategies fared in different industries, on companies with different cap-sizes, in different time periods, and during certain market conditions, would be helpful for us to gain a better idea of situations where these strategies may work better or worse We would have also like to find companies that were particularly successful with either strategy and worked to find the causes of the good performance. Finally, we would have liked to perform more statistical analysis on our yearly returns to learn more characteristics about our results.

Link to dataset: https://www.kaggle.com/qks1lver/amex-nyse-nasdaq-stock-histories

https://www.cnbc.com/2019/03/19/passive-investing-now-controls-nearly-half-the-us-stock-market.html

https://corpgov.law.harvard.edu/2018/11/29/the-shift-from-active-to-passive-investing-potential-risks-to-financial-stability/

https://finance.yahoo.com/quote/%5EGSPC/history?p=%5EGSPC