Quantitative Trading Strategy:

Stock Selection with Market Timing via Factor Analysis and Behavioral Finance

Project Report

MF703: Programming for Mathematical Finance

Professor Chris Kelliher

December 11th, 2021

Boston University, Questrom School of Business

Daniel Babitskiy

Dieynaba Awa Ndiaye

Taiga Schwarz

Camellia (Qiutong) Ye

Ziyi (Josie) Zhang

BOSTON UNIVERSITY

Table of Contents

Abstract

Quantitative trading has always offered a massive computational power to investors to help them access broader opportunities in the market. The timing of whether to buy or sell stocks has always been a topic of great discussion amongst traders, investors, brokers etc. The purpose of this project is to use factor analysis and behavioral finance in order to construct trading strategies to help with market decisions on the timing to buy or sell.

To construct these trading strategies, we first create a stock selection model that allows us to assess the performance of a list of stocks using logistic regression with a binary model along with chosen metrics, namely, the earnings-to-price ratio in addition to the Piotroski F-score. Then, based on our stock selection results, we create a market timing strategy using the VNSP and the NSP factor. This helps us measure the investor's willingness to sell, either due to companies' stock earnings or losses.

After backtesting, the major results of our research show that with a strategy return of 44.72%, compared to a market return of 66.32% and a stock return of 22.12%, and although our system does not beat the market, there is quite a good difference with the stock return. This indicates the usefulness of our strategy and room for future improvements.

I. Methods

a) Stock Selection

The goal of our project is to construct a trading strategy that outperforms the market in the long-run. Thus, the natural first step is to attempt to model whether a stock would outperform or underperform the market in future periods. In order to do this, we hypothesize that companies that are undervalued by the market despite being financially strong will outperform the overall market. Furthermore, it makes sense heuristically that large companies are less likely to be undervalued due to having higher analyst following and higher share-turnover. Thus, we also hypothesize that there is a significant size bias between small and large firms. The idea that "winners" can be predicted by applying a fundamental analysis strategy to undervalued stocks is covered in Piotroski (2002). However, Piotroski only set out to show that fundamental analysis can be used to improve the overall returns of the average value investor. What we set out to do is to model stock outperformance across all stocks within the S&P 500, which is a more ambitious goal and thus our hypothesis may not hold. The success of our strategy depends on the cleanliness and relevance of our data, the soundness of our modeling methodology, and whether we can draw robust conclusions from our results—whatever they may be.

i. Data

We chose the S&P 500 index to represent the overall market and choose our equity pool to be the constituents of S&P 500. We focus our analysis on the time period from the beginning of 2007 to the end of 2017 (Q1 2007 to Q2 2015) for our sample period, and 2017 to the end of 2020 (Q3 2015 to Q4 2016) for our back-test period. The factors for our stock selection strategy fall into 3 categories: value, fundamental metrics, and company size. Finally, the frequency of our data is quarterly due to the fundamental metrics being read from quarterly financial statements. We provide more detail into the factor data below, along with other data necessary for our modeling and analysis.

i.1. Value

We use Earnings/Price ratio as a proxy for the value of the company. This is calculated in the following way:

$$\text{EP R}atio = \frac{\boldsymbol{Earnings\ per\ Share}}{\boldsymbol{Price\ per\ Share}}$$

Another proxy for value is the Price/Earnings ratio, which is simply the reciprocal of the E/P ratio. This is commonly used to represent the value of the company, as one could say “Company XYZ is priced at 15 times its earnings,” which is easily understood. Clearly, a company with a high P/E ratio could be considered overvalued, while a low P/E ratio could be considered undervalued. However, the danger of using P/E ratio in our case is that if a company does not report earnings or is not profitable, then the P/E ratio would be a null value (since a 0 would be in the denominator of the ratio). This is not uncommon to see in high-growth companies, which may not be profitable at time t, but are given a high valuation because of the possibility of high future earnings. By using E/P ratio, a company with zero earnings will result in an E/P ratio of 0 instead of null. In this case, a high E/P ratio would signal that a company is being undervalued, and a low E/P ratio would signal that a company is being overvalued. Of course, whether a company is over- or under-valued is relative to the sector or industry, as some have higher P/E ratios on average (e.g. high-growth technology sector); however, we do not consider this in our project out of simplicity. Incorporating sector bias would be an interesting subject for further research.

i.2. Fundamental Metrics

We use the Piotroski F-score to represent the financial strength of a company. To summarize, the Piotroski F-score is an aggregate of 9 binary signals that measure 3 areas of a firm's financial condition: profitability, financial leverage/liquidity, and operating efficiency (Piotroski 2002). The range of scores is from 0 to 9, where higher values indicate stronger financial health.

The formula is below:

$$\mathrm{F_{Score}} = F_{ROA} + F_{\Delta ROA} + F_{CFO} + F_{ACCRUAL} + F_{\Delta \mathrm{MARGIN}} + F_{\Delta TURN} + F_{\Delta LEVER} + F_{\Delta LIQUID} + \mathrm{F_{EQ\_OFFER}}$$

Profitability:

- F_{ROA} = 1 if ROA > 0; 0 otherwise
  - *ROA* (Return on Assets) is calculated as $\frac{Net\ Income}{Total\ Assets}$
- $F_{\Delta ROA}$ = 1 if ΔROA is > 0; 0 otherwise
  - *ΔROA* (change in ROA) is calculated as *ROA – prior year's ROA*
- F_{CFO} = 1 if CFO > 0; 0 otherwise

- CFO (Cash Flow from Operations) represents the amount of cash a company generates from normal business operations

- $F_{ACCRUAL}$ = 1 if CFO > ROA; 0 otherwise
  - $ACCRUAL$ is calculated as $\frac{Net\ Income - CFO}{Total\ Assets}$

Leverage/Liquidity:

- $F_{\Delta LEVER}$ = 1 if leverage ratio goes down; 0 if leverage ratio goes up
  - $\Delta LEVER$ captures changes in a firm's long-term debt levels and is calculated as the change in the ratio of total long-term debt to average total assets
- $F_{\Delta LIQUID}$ = 1 if ΔLIQUID > 0; 0 otherwise
  - $\Delta LIQUID$ measures the change in liquidity and is calculated as the change in a firm's current ratio (*Current Tot. Assets / Current Tot. Liabilities*)
- F_{EQ_OFFER} = 1 if firm did not issue common equity; 0 otherwise ***
  - A company issuing common equity is generally a negative sign that the company is in need of fund-raising. *** However, we were unable to find reliable data on this so we choose not to include this factor in our F-score metric.

Operating Efficiency:

- $F_{\Delta MARGIN}$ = 1 if ΔMARGIN > 0; 0 otherwise
  - $\Delta MARGIN$ is the change in the firm's gross margin ratio ([*Revenue – Cost of Goods Sold]/ Revenue*)
- $F_{\Delta TURN}$= 1 if ΔTURN > 0; 0 otherwise
  - $\Delta TURN$ is the change in a firm's asset turnover ratio (*Revenue / Total Assets)*

Although Piotroski chose his factors based on profitability and default risk trends he noticed in high book-to-market firms, we would like to test if the same factors could be used to identify strong financial performance in the broader market. However, one potential problem of using the F-score as a factor is that it is an aggregate of binary signals and thus we may be losing potentially significant information by discretizing the underlying financial ratios to having the value of a 1 or 0. We will go more in depth into this problem when we create our model.

i.3. Company Size

We use historical market capitalization data as a proxy for company size. We create a factor called *size* that takes the value of a 1 or a 0 based on the below formula:

$$size = \begin{cases} 1, & mkt\ cap \geq median\ mkt\ cap \\ 0, & mkt\ cap < median\ mkt\ cap \end{cases}$$

Where the median market capitalization is the median of all companies in the S&P 500 in the given quarter. We will use this size factor as a dummy variable in our regression model.

ii. Data Sources and Limitations

We use Bloomberg Terminal to find data for the E/P Ratio, fundamental metrics for the F-score, and quarterly return data. Additionally, we use the Financial Modeling Prep API for the historical market capitalization data.

We had a couple important limitations when it came to finding data. One major issue was finding data that corresponds to the list of all historical constituents of the S&P during our time frame. This would have been ideal to include because not having delisted companies in our dataset would allow survivorship bias to affect our results. We were successful in scraping a list of tickers for all historical constituents of the S&P in our time frame, however we were not able to figure out how to source the financial metrics that correspond to each ticker. As a result, we used only the current companies in the S&P 500 to train our model.

iii. Models

Methodology:

We are attempting to model market outperformance in the next period, which we decide to be the next quarter. As our dependent variable is binary, we can let outperformance in the next quarter to be denoted as a 1 and underperformance to be denoted as a 0. The binary nature of our dependent variable leads us to choose logistic regression as our model of choice. To begin, we first construct a stacked data frame with a multi-index of date ("year_quarter") and ticker, and columns for each of the explanatory factors and dependent variable (1 if observed outperformance; 0 otherwise). Then we conduct an 80-20 train-test split and train our logistic regression model on the training set. After making necessary adjustments, we then validate our model on the test set, draw conclusions about our model, and finally implement our model on an out-of-sample set in order to provide a list of stocks to invest in for our market-timing back-test.

iii.1. Model 1: Initial Hypothesis

Our first model is using E/P Ratio, F-score, and size interactions with E/P Ratio and F-score as predictors. Our model is below:

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 * \text{EP Ratio} + \beta_2 * \text{F\_Score} + \beta_3 * (\text{EP Ratio} * \text{Size}) + \beta_4 * (\text{F\_Score} * \text{Size})$$

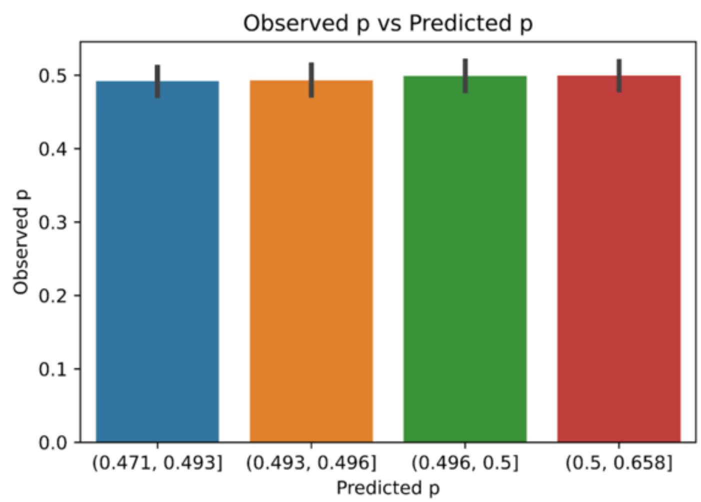
Note that *p* is the probability of outperformance and *ln(p/1-p)* is called the log-odds.

The results of training our model are below:

Logit Regression Results

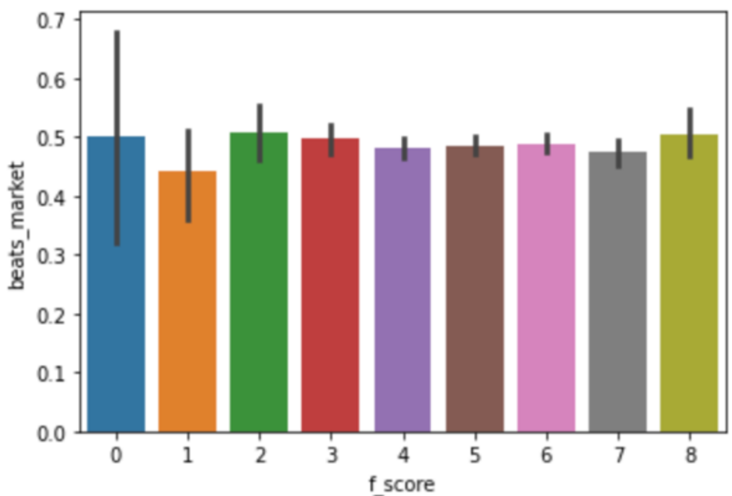
Dep. Variable:	beats_market	No. Observations:	9500
Model:	Logit	Df Residuals:	9495
Method:	MLE	Df Model:	4
Date:	Tue, 07 Dec 2021	Pseudo R-squ.:	0.0001504
Time:	00:38:18	Log-Likelihood:	-6583.6
converged:	True	LL-Null:	-6584.6
Covariance Type:	nonrobust	LLR p-value:	0.7393

	coef	std err	z	P>|z|	[0.025	0.975]
EP_RATIO	0.0329	0.468	0.070	0.944	-0.885	0.950
EP_ratio_size	0.6355	0.733	0.867	0.386	-0.800	2.071
F_score_size	-0.0838	0.088	-0.948	0.343	-0.257	0.089
f_score	-0.0520	0.124	-0.419	0.675	-0.295	0.191
intercept	0.0233	0.083	0.281	0.778	-0.139	0.185



Analyzing our model, we see that the predictors (including the intercept) are not significant. We also plot the observed values of *p* against the predicted values of *p* grouped by quantile and see that the trend is flat. A model that is well-trained would show a positive trend in the plot, which makes logical sense because we would expect the upper quantiles of predicted *p* to correspond to higher values of observed *p*. In this case, our model does as well as 50-50 random selection.

To begin, we expected to see that F-score has some correlation with outperformance. However, looking at the plot below, we see that there clearly is not.



Is there a possibility that we are losing information by discretizing the fundamental ratios that comprise the F-score? To test this theory, we train a second model in which the F-score is replaced by its significant components.

iii.2. Model 2: Replacing F-score with its Significant Components

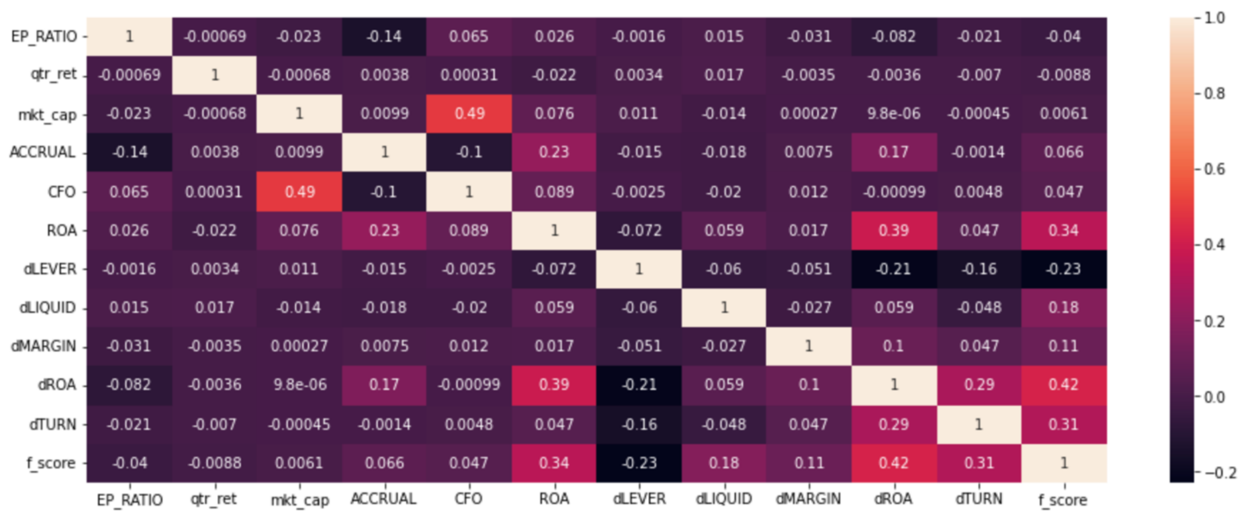
We first create a model that has every component of F-score included and then train the model. Our results are below:

Logit Regression Results

Dep. Variable:	beats_market	No. Observations:	9500
Model:	Logit	Df Residuals:	9487
Method:	MLE	Df Model:	12
Date:	Wed, 08 Dec 2021	Pseudo R-squ.:	0.001433
Time:	19:08:01	Log-Likelihood:	-6575.1
converged:	True	LL-Null:	-6584.6
Covariance Type:	nonrobust	LLR p-value:	0.09175

	coef	std err	z	P>|z|	[0.025	0.975]
EP_RATIO	0.0665	0.429	0.155	0.877	-0.774	0.907
ACCRUAL	-3.1976	1.559	-2.051	0.040	-6.254	-0.141
CFO	-19.6329	8.309	-2.363	0.018	-35.918	-3.348
ROA	-0.1134	0.553	-0.205	0.838	-1.198	0.971
dLEVER	-0.8596	0.776	-1.108	0.268	-2.381	0.662
dLIQUID	-0.1380	0.556	-0.248	0.804	-1.228	0.953
dMARGIN	-9.9723	5.695	-1.751	0.080	-21.134	1.189
dROA	-0.3429	0.533	-0.644	0.520	-1.387	0.701
dTURN	-0.0346	0.467	-0.074	0.941	-0.950	0.880
CFO_size	19.9496	8.251	2.418	0.016	3.779	36.120
ROA_size	-0.3091	0.367	-0.843	0.399	-1.028	0.410
ROA_dROA	1.0396	0.848	1.226	0.220	-0.622	2.701
intercept	11.3846	5.087	2.238	0.025	1.414	21.355

We see that CFO, CFO_size, ACCRUAL, and the intercept are statistically significant under a 95% confidence level. dMargin has a 0.08 p-value, so we will keep it in the model. We will drop the rest of the variables from the model for simplicity. We also include a size interaction term between CFO and company size. This is due to the relatively high correlation between mkt-cap and CFO, which we can see in the correlation matrix below.



Thus, our second model is:

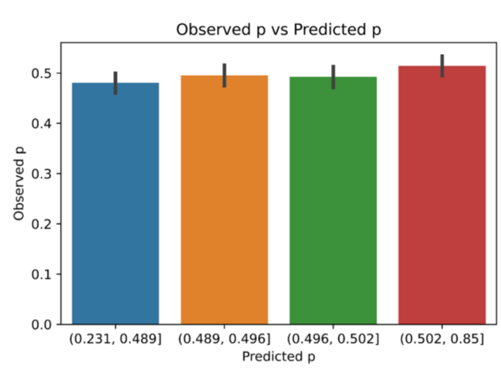
$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 * \mathrm{CFO} + \beta_2 * \mathrm{Accrual} + \beta_3 * \Delta\mathrm{Margin} + \beta_4 * (\mathrm{CFO} * \mathrm{Size})$$

The results of training our model are below:

Logit Regression Results

Dep. Variable:	beats_market	No. Observations:	9500
Model:	Logit	Df Residuals:	9495
Method:	MLE	Df Model:	4
Date:	Tue, 07 Dec 2021	Pseudo R-squ.:	0.001062
Time:	17:18:25	Log-Likelihood:	-6577.6
converged:	True	LL-Null:	-6584.6
Covariance Type:	nonrobust	LLR p-value:	0.007346

	coef	std err	z	P>|z|	[0.025	0.975]
CFO	-17.6219	7.898	-2.231	0.026	-33.101	-2.143
CFO_size	17.6884	7.782	2.273	0.023	2.436	32.941
ACCRUAL	-3.8136	1.457	-2.618	0.009	-6.669	-0.959
dMARGIN	-9.6374	5.526	-1.744	0.081	-20.467	1.192
intercept	11.0882	4.894	2.266	0.023	1.495	20.681

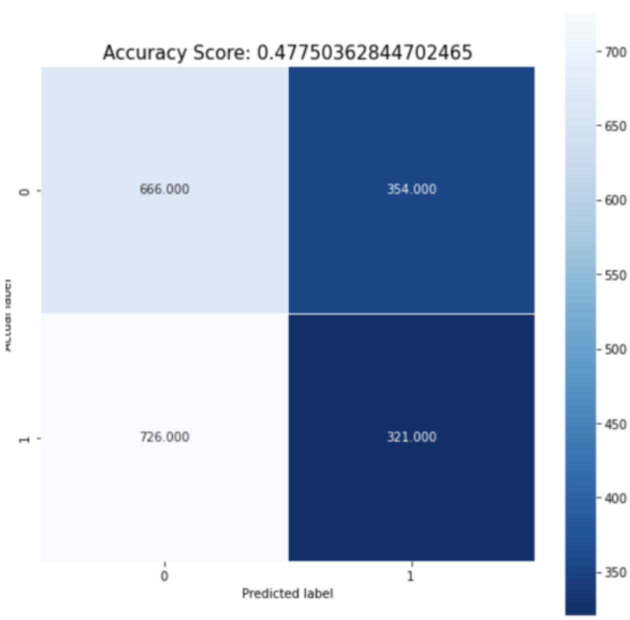


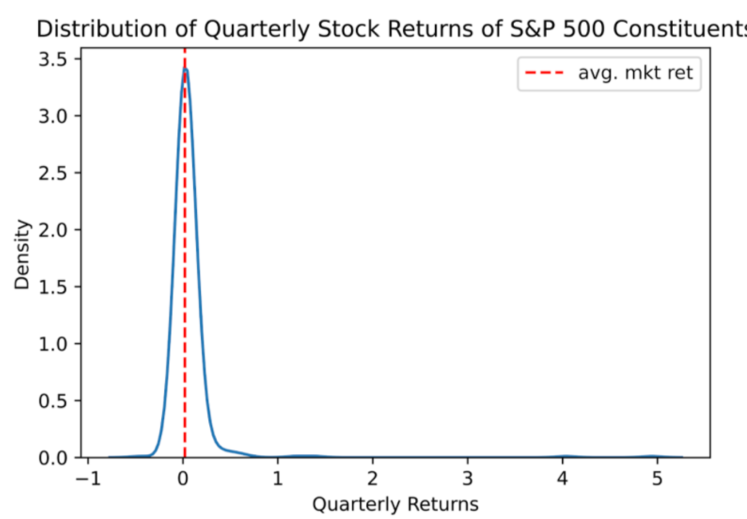
We now have statistically significant coefficients and also there is a slight improvement in predictive power, as we notice more of a positive trend in the quantile plot. However, the coefficient values are negative for some of the financial ratios, which does not make sense from an economics perspective. This suggests that our model is overfitting.

iv. Test.

We can confirm this by testing our model on the test set. The results of testing the second model are below.

The observed overall average of beating the market in the test sample is 48.6%. We see that the distribution of returns is approximately symmetric about the average in the plot to the right. However, the model does slightly worse than picking 50-50 despite having performed slightly better on the training set. This tells us that just because a model is better on the training set, it does not mean that it will generalize to a test sample. This is a very important insight that we can derive from this process.





v. Implementation.

Nevertheless, we need a set of stocks to test our market-timing strategy for our back-test period of 2017 through the end of 2020. We can still derive insightful conclusions about our market timing method by comparing the difference in performance between using market timing on our selected stocks, against simply holding the stocks selected without applying market timing. In order to provide a list of stocks to invest in for each quarter, we use our logistic regression model to output the stock with the highest probability to outperform the market in the next quarter. The list of stocks is on the right-hand side:

qtr	ticker
2017_1	CHTR
2017_2	CHTR
2017_3	CHTR
2017_4	PENN
2018_1	RCL
2018_2	CHTR
2018_3	FANG
2018_4	KHC
2019_1	CHTR
2019_2	CHTR
2019_3	NCLH
2019_4	NCLH
2020_1	NCLH
2020_2	NCLH
2020_3	CCL

There is an interesting commonality between the tickers. Most are cruise liners or energy/petroleum holdings companies. These companies also have relatively low P/E ratios, generate lots of cash flow, and have limited growth. Although our model does not accurately predict outperformance (which is a difficult task to begin with), it does find companies with interesting commonality that may be worth further research.

b) Market Timing

i. Disposition Effect

The traditional disposition effect is based on the investors' behavior which emphasizes that when some stocks become profitable, investors tend to sell their stocks for real gains. The greater the profit margin is, the stronger the investor's willingness is to sell. When investors are at loss, they prefer risk and in order to avoid inevitable losses, they would rather bear additional risks. The disposition effect can be explained from the perspective of behavioral finance with mean regression theory, mental accounting and prospect theory. Mean regression theory states that every stock has its theoretical reasonable value, and so stock prices fluctuate and return to this value. In investors' mental accounting, they believe they always make the right decision so that when stock performs poorly, they fail to admit the wrong decision and tend to wait until the price goes up again. In prospect theory, when the stock price is higher than the investor's buying price, that is, when the investor is in a profitable state, the value function is a concave function, indicating that investors' behavior is risk averse. The value drop caused by the decline in income is greater than the increase in value caused by the rise in income. When the investor is at a loss, the value function is a convex function, indicating that investors' behavior is risk preferred. That is when the stock price is lower than the investor's buying price, in order to avoid the uncertainty caused by price drops, investors tend to sell the stocks they hold. In conclusion, combining the three theories, Investors have the tendency to sell well-performed and hold loss-making stocks and wait for the price to rise.

The V-shaped disposition effect was introduced later by Ben-David and Hirshleifer; from their research result, it shows that investors' selling propensity is actually V-shaped. Unlike the traditional disposition effect, investors tend to sell and stop loss rather than keep holding the stock when the stock is in a loss position.

ii. Gain/ Loss Factors

To measure the investors' tendency of buying and selling, we use gain factor and loss factor. The computation introduced in Li An's paper is as follows:

$$Gain_t = \sum_{n=1}^{\infty} \omega_{t-n} gain_{t-n}$$

$$gain_{t-n} = \frac{P_t - P_{t-n}}{P_t} \cdot \mathbf{1}_{\{P_{t-n} \leq P_t\}}$$

$$\omega_{t-n} = \frac{1}{k} V_{t-n} \prod_{i=1}^{n-1} [1 - V_{t-n+i}]$$

$$Loss_t = \sum_{n=1}^{\infty} \omega_{t-n} loss_{t-n}$$

$$loss_{t-n} = \frac{P_t - P_{t-n}}{P_t} \cdot \mathbf{1}_{\{P_{t-n} > P_t\}}$$

$$\omega_{t-n} = \frac{1}{k} V_{t-n} \prod_{i=1}^{n-1} [1 - V_{t-n+i}]$$

We start with computing the gain and loss factor. First compute the average stock price of the past n days, then times with the indicator of whether the current price is less than n days ago. The second step is to compute the weight by the turnover rate. The last step is to multiply the weights and gain/loss factor to get the weighted gain/loss factor.

iii. NSP and VNSP (V-shaped Net Selling Propensity)

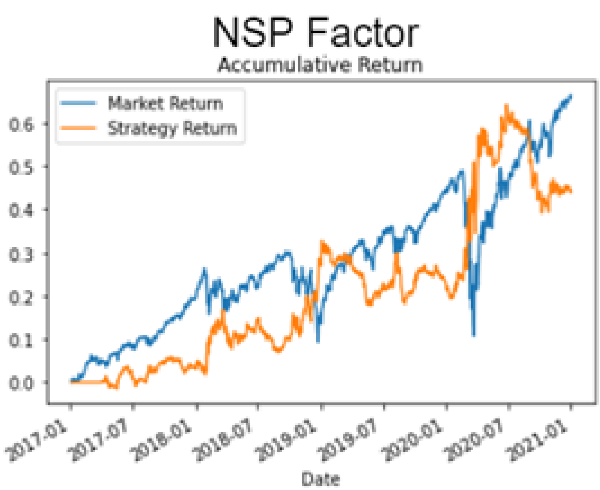
Both the NSP and VNSP factors consist of Gain and Loss factors which indicate investors' willingness to sell the stocks. According to the Disposition Effect, the Gain factor is always positive while the Loss factor is negative. Instead of simply combining these two factors, we attribute a constant δ to the Loss factor and set it as 0.5, assuming that the Gain Effect is twice as strong as the Loss effect.

$$NSP = Gain + \delta\, Loss \qquad VNSP = Gain - \delta\, Loss$$

II. Back-Test

Comparison of Traditional And V-Shaped Disposition Effect

To figure out which disposition effect fits the US stock market better, we used the S&P 500 from 2017 to 2020 as sample data to apply our timing strategy. Both factors measure the mass tendency to sell the stocks and higher NSP or VNSP value indicates that investors are more willing to take a short position in stocks in the following day. Therefore, we compare the NSP or VNSP factor of day *t-1* and day *t-2* in order to decide whether to sell ($Factor_{t-1} > Factor_{t-2}$) or buy ($Factor_{t-1} < Factor_{t-2}$) the target asset on day t .

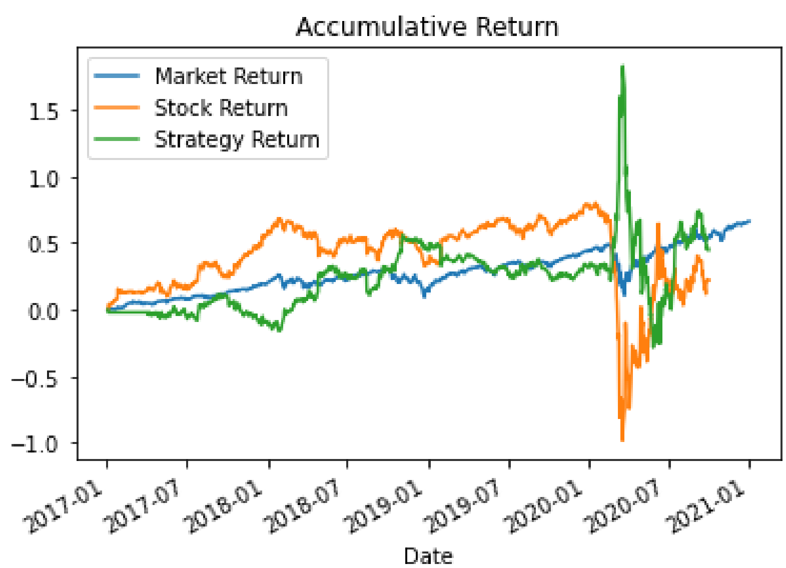


Apparently, the traditional Disposition Effect performs better than the V-shaped Disposition Effect and thus we should use the NSP factor as an indicator in our market timing strategy.

Application of Disposition Effect on Target Stocks

We then apply the trading strategy to the selected stocks from 2017 to 2020.

market return	stock return	stock std	stock sharpe ratio	strategy return	strategy std	strategy sharpe ratio
66.32%	22.12%	0.0432	-10.2357	44.72%	0.0426	-5.0675



III. Interpretation

After applying the strategy, we gain a final return of 44.72% with a standard deviation of 0.0426 over four years. According to the figure above, we notice that our strategy works well as a hedging method against the selected stocks. Although the strategy couldn't beat the market, we still reach the conclusion that it is better to implement the strategy than simply holding the portfolio for four years.

IV. Conclusion

Determining the timing of whether or not to sell or buy stocks is basically one of the most important tasks not only in the financial market, but in mathematical finance at large. This is the reason why our team created a project focused on stock selection with market timing using factor analysis and behavioral finance, in order to find a strategy that could be helpful for the future.

Based on our results, with a strategy return of 44.72%, compared to a market return of 66.32% and a stock return of 22.12%, even though our system doesn't beat the market, there is quite a good difference with the stock return. This indicates the usefulness of our strategy and room for future improvements.

References:

An, L. (2013). *The V-shaped disposition effect.* Working Paper.

Piotroski, Joseph D. "Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers." *Journal of Accounting Research*, vol. 38, 2000, p. 1., https://doi.org/10.2307/2672906.