Quality, Volatility, and Sentiment Trading Strategy

Rachel Ahn
Management Science & Engineering
Stanford University
raahn@stanford.edu

Matthew Tan
Computer Science
Stanford University
mratan@stanford.edu

Kimberly Te
Computer Science
Stanford University
kimte@stanford.edu

Andrew Matangaidze
Computational & Mathematical Engineering
Stanford University
abmatan@stanford.edu

Jialu Sun
Computational & Mathematical Engineering
Stanford University
jlsun@stanford.edu

Abstract

The purpose of our project was to explore and integrate fundamental, volatility, and sentiment metrics extracted from to predict equity trends as a potential trading strategy. The data universe was based on the S&P 1500 provided by Quantopian. We assessed and tested combinations using over 30 features related to fundamentals, volatility, and sentiment using the Quantopian, Morningstar, SEC Analytics, PsychSignal, and SentDex resources. Fundamentals features were updated quarterly, volatility was updated monthly, and sentiment data was updated daily. In the first phase, individual modeling was done for each of the three types of features. In the second phase, combined models using features across the three groups were also explored. For features from only fundamentals and features from only sentiment, the baseline model was a simple ranking model that equally weighted each feature. Two other models were assessed: correlated weights model and linear regression model. Upon testing, performance for models varied based on time regime, where total returns was relatively low to negative. However, the volatility models and combined models had higher positive returns. Low-volatility, high-volatility with fundamentals, and sentiment volatility was explored based on the methods of Viera [1]. Overall, our results showed that the combined model with sentiment volatility and fundamental quality had the highest total returns of 228.71%. However, it had the highest maximum drawdown and a low sharpe ratio, suggesting greater risk.

1 Introduction

Trading idea generation is the holy grail of systematic investment management and a lot of research has been explored to explore the main drivers of stock market returns. For many decades, fundamental company data as been at the centre of constructing these quantitative trading strategies. Fundamental signals is comes from three financial statements: balance sheet, income statement, and cash flow statement. These statements provide an interconnected indication of the company performance. These features are used as standard practice for traditional as well as quantitative hedge funds. Rather than looking at companies individually and building discounted cash flow models or multiples valuations, we seek to apply algorithmic approaches to understand the fundamental anomalies present. In particular, we seek to understand the quality and volatility anomalies and examine to what an extent natural language processing (NLP) can help explain return premiums. Our motivation for the quality anomaly is due to the emergence of the corporate debt bubble. Since the recession from 2009-2019, corporate debt has increased worldwide from 84% of gross world product to 92%. Total U.S. corporate debt reached 47% of the U.S. economy in November 2019. This has led to an increase in "zombie" companies, who currently hold debt, and are unable to make payments and therefore repeatedly take on more debt or refinance in order to make their current payment. Interest rates have never been lower, even dropping negative for the US 1 month T-bill in

March 2020 during the corona-virus volatility, which increases the willingness to buy riskier high yield assets as it remains at historical lows and creates a mis-perception of the value of risk [2].

Over the past decade, sentiment analysis in NLP has showed promise in predicting equity trends. According to Bollen et al., sentiment analysis is the classification of emotions or moods from textual data, such as news, social media, and financial records [3]. Individual terms can be categorized into sentiment via dictionaries. The Harvard IV dictionary includes various categories of sentiment with two broad categories of positive and negative, yet it does not comprehensively cover financial terminology. The Loughran-McDonald (LM) dictionary specifically focuses on financial terms and is based on 10-k filing returns, trading volume, return volatility, fraud, material weakness, and unexpected earnings. It has six categories: positive, negative, uncertainty, litigious, modal strong (implying strong confidence and necessity), and modal weak (implying weak confidence and possibility) [4].

1.1 Objective:

Therefore, a viable trading strategy could be based on identifying alpha trading signals using a composite of volatility, quality, sentiment, and/or a combination of these features. Therefore, the purpose of our project was to predict daily, weekly, and monthly percentage change in stock price using quarterly fundamentals indicators and daily textual signals from financial reports, news articles, and social media. We explored fundamental anomalies and sentiment over correlation weighted models and linear regression.

2 Literature Review

2.1 Quality

In academic finance, many researchers have studied classical anomalies of value, volatility , size, and quality to help explain the sources of stock returns. Earlier work includes single-factor CAPM and Fama-French factor models. An interesting analysis that showed the role of behavioural bias of analysts was completed by Bouchard et al. [5]. They found that the behavioural view of systemic bias by analyst to underestimate the return of high quality firms had significant implications in explaining the quality anomaly than the long held risk view that is consistent with the efficient market hypothesis. They used a linear regression of "mistake" of analyst expectations against measure of quality. In this linear model, quality was measured by operating cash flow. In particular, they find that three main features are statistically significant (Operating cash flow, rolling volatility, and book to market). No attempt was made to combine with value or volatility anomalies, or to add text data to further explain the behavioural component to the positive returns.

Simarily, Kozlov et al., found out that there are huge diversification benefits if one combines both quality and value portfolios. They covered global equity markets across different company sizes from 1988-2012, quality portfolio had the highest shape ratio of 0.69 and was the least volatile.By combining a risk-weighted portfolio of quality and value, they found a Sharpe ratio of 0.99. Their methodology was centred on defining quality using the accruals method to come up with cash minus earnings factor. They also ran experiments with other factors such as operating cash flows to assets, Return on Assets(RoA), Return on Equity(RoE), and leverage. The cashflow-based measure produced the top returns, followed by accruals, but had the highest volatility. This study suggests the presence of the quality anomaly, in addition to other traditional anomalies such as size and premium effects but no attempt was made to incorporate volatility effect or the text data to complement financial-backed metrics [6].

2.2 Volatility

Viera et al. made a thorough study of the volatility anomaly covering both developed and developing markets across many time zones and different market sizes. They showed that the low volatility portfolio has higher return relative to high volatile portfolio, confirming the low-volatility anomaly. Volatility in their case is calculated as the standard deviation of the daily total price return for trailing 12 months. Notable in their research was the weak relationship during times of crises, especially during financial crisis period and emerging markets recently. Their study however did not touch on other anomalies or the role played by alternative data sources in the form of text data [1].

2.3 Sentiment Analysis

A potentially useful source of sentiment analysis data are financial records, like SEC filings, social media data and news articles. In their paper, Bollen et al. used commercially available tools to get a 87 percent prediction accuracy on Dow Jones Industrial Average (DJIA) stock market returns. Their methodology was centered on including specific public mood dimensions. The study analysed Twitter data towards the prediction of changes in DJIA closing values, where daily public mood can improve predictions on up and down closing values in the DJIA. They collected text from daily Twitter feeds over a 10-month period in 2008 (10 million tweets from 2.7 million users). Tweets were filtered for sentiment-expressing content, removing spam and information-oriented content. Then, they analyzed sentiment by using two mood tracking tools with lexicons for sentiment (OpinionFinder for positive vs. negative moods and Google-Profile of Mood States (GPOMS) for 6-types of mood). Models for the prediction included Self-organizing Fuzzy Neural Network (SOFNN) using DJIA values and permuted mood time series. Results also showed a reduction of mean average percentage error by 6%, implying sentiment from social media can be a useful return signal indicator [3].

In a related paper but utilising a different text data source, Lee et al. examined financial event-related features extracted from 8K documents to predict stock prices, which showed potential for short term prediction after a financial event. The 8K reports were collected from S&P 500 using the SEC's EDGAR dataset and the data was combined with analyst Earnings Per Share (EPS) consensus estimates. Financial events were extracted and categorized from the reports. Features included recent movements prior to 8k report release, volatility S&P 500 index, event category, unigrams, and unigram non-negative matrix factorization (NMF) for addressing sparsity. Random forest classifiers were used as the baseline models with unigram-based approaches. Additional features that were also explored includes sentiment, bigram, and word clustering. None boosted performance, however, possibly due to the sentiment lexicons being general and not accurately capturing financial moods as is the case with Loughran and MacDonald dictionary. Overall, the work suggests text analysis can improve stock market predictions but the accuracy was not that high to confidently develop a viable trading strategy [7].

Utilising a different approach on a different data source in the form of live article news from various sources, Ke et al (2019) developed a simple long-short trading strategy, buying top 50 sentiment score stocks and selling low sentiment bottom 50. Their strategy, which relied heavily on a white box data mining methodology of predictive screening, topic modelling and likelihood penalisation to come up with a sentiment score that does not depend on pre-existing dictionaries, beat returns from RavenPack, a commercially available trading tool based on sentiment data [8]. The paper did not explore the implications of combining text data with financial-based metrics.

Overview In this paper, we are different in the sense that we explore the innovative idea of combining quality, volatility, value and text sentiment analysis for an alpha-generating trading strategy. We introduce the concept of sentiment volatility, in addition to defining tradition volatility using a 6-month time window. Defining quality measure differently using free cash flow yield and return on invested capital, we design a composite quality-volatility- value-and sentiment quantitative trading strategy on SP 1500 stocks that explores the strength that each provides. For sentiment analysis, we combine text data from SEC filings, social media, and article news sources.

3 Data

3.1 Quantopian

To research and build our strategies, we leveraged the Quantopian platform which provides daily pricing and volume data for the US equities market. For the purposes of our project, we used historical daily close prices which are already adjusted for corporate actions such as splits, dividends, or mergers/acquisition that would incorrectly skew the perception of the historical price. This accounts for automatic dropping or adding of companies to the S&P1500 for example. This adjustment is made at the point-in-time to avoid any look-ahead bias or selection-bias. Quantopian data is available starting Jan 1, 2004 and we have tested a variety of dates in preliminary testing. However, for the results of our final testing we have chosen to consider a five-year window from Jan 1, 2015 - Dec 31, 2019.

3.2 Quality Features

For the quality features, we looked at the Morningstar dataset in Quantopian which holds hundreds of different fundamental metrics. Most of these are updated quarterly but depending on the field, some are reported daily. See Table 1 for the features that we have tested.

We tried a variety of fundamental features that we have seen through in the literature that could potentially illustrate what defines a 'quality' company. For the combined model we narrowed our definition of quality to the following formula:

$$Quality = \textit{long term debt to equity ratio} + \textit{return on capital} + \textit{cash return} + \textit{free cash flow yield}$$

3.3 Sentiment Analysis

For the sentiment features, we extracted sentiment from a combination of SEC filings via SEC Analytics suite, StockTwits via PsychSignal, and news articles via the SentDex algorithm. Sentiment features were updated daily. Features were also processed to include z-scores, simple moving averages, and percentages of documents.

The SEC Analytics Sentiment dataset contains sentiment scores using the LM lexicon and Harvard IV-4 lexicon (only negative scores) on SEC filings, namely 8k reports, since 1994 from EDGAR filings. The main features utilized from this dataset was the six LM sentiment scores (positive, negative, uncertainty, litigious, modal strong, and modal weak) and Harvard IV-4 negative sentiment scores. Sentiment scores involved counting of terms in the filing that fit within the given category.

Moreover, we also leveraged sentiment datasets from Quantopian in conjunction with the SEC sentiment scores. This was done to cover all sources of sentiment information- news, company specific news and social mood. Based on the work of Bollen et al.[1], we incorporated social media sentiment from PsychSignal, which computes trader mood based on Stocktwits and Twitter data. Unlike the SEC dataset, PsychSignnal trader mood tool provides daily bull and bear sentiments, which relate to moods that have a positive and negative impact on stock prices respectively. Specifically, we used these bull and bear sentiment scores computed from Stocktwits data both separately and in conjunction with the other sentiment sources.

Finally, we also utilized the SentDex algorithm. SentDex incorporates daily news articles for binary sentiment classification. It provides positive-negative sentiment categories on named entities in daily news articles. Quantopian SentDex provides a sentiment signal, which we use as a feature. See Table 1 for details.

4 Methodology

For our long/short trading strategy, we considered the combinations of different parameters such as market capitalisation, slippage, capital, different re-balance rates depending on the dataset and volatility, chosen volatility regime, time horizon, and time period.

In the first phase of our project, we examined each group of features individually (fundamentals, volatility, and sentiment). We tested different features within each group type to identify potential features with positive total returns. In the second phase, after gaining signal of positively correlated features, we then integrated these features into a combined model.

4.1 Models: Quality Anomaly and Sentiment Analysis

The quality-plus-sentiment factor model was analysed and tested for robustness on three alternative formulations: equally weighted, correlated weights, and value-weighted linear regression. The

Source	**Feature**
Morningstar	Market Cap
Morningstar	EBITDA Margin
Morningstar	Net Margin
Morningstar	Operating Cash Flow
Morningstar	Capital Expenditure
Morningstar	Change in Employees
Morningstar	Change in R&D
Morningstar	Book to Value Yield
Morningstar	Current Liabilities
Morningstar	Debt to Assets
Morningstar	Interest Coverage
Morningstar	Net Income Growth
Morningstar	Revenue Growth
Morningstar	Operation Ratios
Morningstar	Cash Return
Morningstar	Long Term Debt to Equity Ratio
Morningstar	Return on Capital
Morningstar	Free Cash Flow Yield
Morningstar	Price to Book Ratio
Morningstar	Sustainable Growth Rate
SEC	LM Positive Proportion of Words
SEC	LM Negative Proportion of Words
SEC	LM Uncertainty Proportion of Words
SEC	LM Litigious Proportion of Words
SEC	LM Modal Strong Proportion of Words
SEC	LM Modal Weak Proportion of Words
SEC	Harvard IV-4 Proportion of Words
PsychSignal Stocktwits	Counts of Bullish Message (Positive)
PsychSignal Stocktwits	Counts of Bearish Message (Negative)
PsychSignal Stocktwits	Difference between Number of Bullish and Bearish Messages
PsychSignal Stocktwits	Ratio of Bullish Messages to Bearish Messages
SentDex	Sentiment Signal

Table 1: Table of fundamentals and sentiment features.
Note: SentDex = number of words for a given sentiment/total number words
sustainable growth rate = shareholder's equity * (1 - dividend per share / diluted earning per share)

baseline model involved a simple ranking, where each feature receive equal weights. Combinations of different fundamental signals and sentiment was then tested with two approaches: correlated weights model and linear regression.

For the correlation models, we considered an equally weighted correlation, recalculated using the historical time series up to the re-balance period in our back-test window. Additionally, we tested an exponentially weighted correlation weighting model. The exponentially weighted correlation approach emphasizes more recent observations therefore could potentially be more sensitive to high volatility or regime shifts. We tested the exponentially weighted correlation approach with different decay parameters such as 30, 90, and 120 day half-lives.

In terms of the linear regression model, we defined our training set to be 1 year prior to the current rebalance date and our test set to be the following month. By defining our model in this way, we were able to capture market fluctuations through the backtest by retraining the model during each rebalance period. We considered different training set windows, rebalance rates, and added regularization through L_1 and L_2 techniques.

4.2 Volatility Models

For the volatility factor and its associated strategies, we simply calculated two volatility measures: 6-month stock price volatility and volatility of the sentiment scores. We then ranked the stocks in our considered investment universe from highest to lowest volatility for the duration of the re-balance period, and then long the low volatility stocks and short the high volatility stocks. For the sentiment volatility, we long high-sentiment volatility and short low sentiment volatility. We calculated the traditional stock volatility as the standard deviation of the daily total price return over a trailing time period of 6 months.

4.3 Combined Models

Our goal of the paper is to integrate quality, volatility, and sentiment analysis into a quantitative trading strategy. As all of these features are interconnected, we have evaluated different combinations and ways to incorporate these three ideologies into a few different strategies. For these strategies, we considered different investment universes, rebalance frequencies, and a constraint to trade equally across all sectors. Note that data frequency was monthly.

We chose to experiment with a high-volatility strategy, coupled with quality features. The intuition behind this strategy is that during high volatility, there would be a flight to high quality companies defined by consistent high cash flows relative to earnings or low debt to equity ratio. We proposed a two-step approach, first measure the time-series volatility to identify the top 50 volatile companies. Secondly, within the top 50 bucket, rank by their quality features and long the top 10. In addition to time-series volatility, we chose to volatility of Psych Signal's bull and bear sentiment scores, calculated as the standard deviation of the ratio of the number of bull over bear messages in Stocktwits over time.

5 Results and Discussion

5.1 Quality Anomaly Factor Strategy

Quality anomaly strategy was tested over the baseline, correlation, and linear regression models. Using the simple baseline of ranking the quality fundamental features z-scores only and applying the long/short trading strategy resulted in the highest return of 15.96%. The simple baseline was one of the few positive signals during this time period. This can be seen in Table 2.

Next, we looked at correlation models. They had overall negative returns with low Sharpe ratios over 2015-2019. When building our models, we tested in smaller time horizons with dates prior to 2014. However, the positive signal did not translate into the longer, more recent time horizon from 2015-2019 as seen by second and third rows in Table 1.

For the long/short linear regression model, when testing with a one year training window, and monthly prediction windows over the backtest period of Jan 3, 2010 - July 31, 2014 , the trading strategy generated a low beta of 0.7, Shape ratio of 0.84 and returns of 36.57% over the period. However, the strategy on a 5-year window from 2015-2019 returned -12.57% with a sharpe ratio of -0.41. This is consistent with the traditional market hypothesis view that one can not consistently beat the market. Of the quality strategies using only fundamental data, the long/short linear regression performed the worst. This can be seen in row six of Table 2. The quality anomaly only long/short strategy significantly under-performed in the more recent time period possibly as a result of investors exploring new alternative datasets to identify alpha generating trading ideas. In 2010, digitization was not yet as ripe as it is today and only a few systematic asset managers explored the idea of natural language processing to seek alpha. Another plausible explanation relates to changing market regimes. 2015-2019 has been generally boom and capital deployed to seek alpha in stock markets has being going up as investors recovered from global financial crisis, hence reallocating their portfolios to have more exposure to stocks.

In addition to a long/short strategy, we looked at the possibility of generating alpha using

Model Class	Brief description	Universe	Long Freq	Total Returns	Beta	Sharpe	Max Drawdown
Fundamental	Long top 10 fundamental, short bottom 10	S&P1500	Monthly	**15.96%**	-0.25	0.58	-35.84%
Fundamental	Fundamental, equally weighted correlation, long/short strategy	S&P1500	Monthly	-2.04%	0.07	0.02	-12.63%
Fundamental	Fundamental, exponentially weighted correlation, long/short strategy	S&P1500	Monthly	-7.96%	0.03	-0.2	-11.58%
Fundamental	Linear regression long only, train annually, predictions quarterly	S&P1500	Quarterly	5.96%	-0.02	0.41	-4.67%
Fundamental	Linear regression short only, train annually, predictions quarterly	S&P1500	Quarterly	-10.89%	0.02	-0.8	-15.45%
Fundamental	Linear regression long short, train annually, predictions quarterly	S&P1500	Quarterly	-12.57%	0.7	-0.41	-19.90%
Volatility	Longs the top 10 low volume stocks	S&P500	Quarterly	**71.26%**	0.7	1.2	-15.35%
Combined	From top 50 high vol companies. Long top 10, short bottom 10 based on quality	S&P1500	Monthly	9.92%	0.03	0.48	-7.19%
Combined	Get high vol companies per sector, then get the top 10 quality per sector and long	S&P1500	Monthly	29.65%	0.43	0.43	-8.79%
Combined*	PsychSignal, bull_bear_msg_ratio, sentiment volatility, and fundamentals	S&P1500	Daily	**228.71%**	5.57	0.81	-80.26%
Sentiment	Correlation model using LM positive proportion, LM negative proportion, LM modal strong proportion, sum of LM positive z-score and LM modal strong z-score, Harvard IV-4 negative proportion, PsychSignal Percentage of Bullish Messages, and SentDex sentiment signal	S&P1500	Daily	-0.41%	0	-8.03	-0.42%
Sentiment	Correlation model using LM positive proportion, LM negative proportion, LM modal weak proportion, LM modal strong proportion, Harvarid IV-4 negative proportion, and PsychSignal bullish messages percent	S&P1500	Daily	**9.28%**	0.06	0.84	-4.86%
Sentiment	Correlation model using LM positive proportion, LM negative proportion, LM modal weak proportion, LM modal strong proportion, Harvarid IV-4 negative proportion, sum of z-scores of LM litigious, uncertain, modal strong, and modal weak, and PsychSignal bullish messages percent	S&P1500	Daily	8.80%	0.02	0.63	-3.99%
Sentiment	Linear regression using LM positive proportion, LM negative proportion, LM modal strong proportion, Harvard IV-4 negative proportion, sum of LM uncertainty z-score and LM litigious z-score, and SentDex sentiment signal	S&P1500	Daily	-2.47%	-0.01	-0.14	-7.71
Sentiment	Linear regression using SentDex sentiment signal and simple moving average of LM litigous proportion over 7-day window	S&P1500	Daily	-1.74%	0.01	-0.12	-7.9

Table 2: Results for highest performing models across fundamentals, volatility, and sentiment. Bolded results indicate the model with the highest total returns in each category (fundamental, volatility, sentiment, and combined). Combined model with sentiment volatility was the highest overall model across all groups.

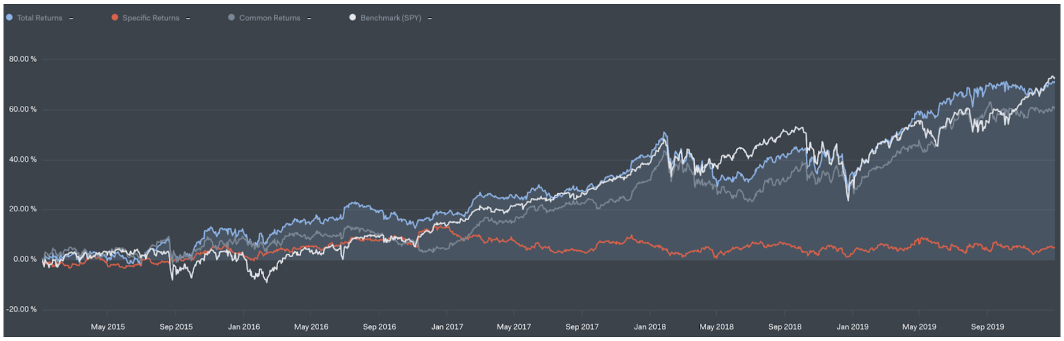

Figure 1: Results of returns for low volatility model from January 1, 2015 - December 31, 2019. Total returns (blue) gradually increased over time, where returns passed the benchmark during months in 2016 and 2019.

long only or short only strategies over the 2015-2019 time period. The long only strategies had insignificant positive returns over the period but had a highest shape ratio.

5.2 Sentiment Factor Strategy

For the correlation-based model, we tested varying combinations of features across SEC sentiment, PsychSignal, and SentDex. Initial testing was for the period January 1, 2018 to December 31, 2019 to narrow the search space over the various possible combinations of sentiment features. Positive results with total returns ranging from 2.41% to 5.61% were generated over this limited time window. However, as in the case of quality anomaly model, the early results from these combinations did not translate to the larger time regime of 2015-2019. Most combinations yielded negative returns. However, some features consistently generated improved but still negative returns. Features that generated these improved negative returns include LM positive proportion, LM negative proportion, and Harvard IV-4 proportion. While their z-scores and simple moving averages did not perform pretty well, the raw sentiment proportions did slightly better. It may be that information is lost when layering additional computation. The highest performing model yielded 9.29% total returns, -0.46 sharpe ratio, -4.85% maximum drawdown. This was a model that utilised correlation weights using LM positive proportion, LM negative proportion, LM modal weak proportion, LM modal strong proportion, Harvard IV-4 negative proportion, and PsychSignal Percentage of Bullish Messages as sentiment features. It utilized a combination of sentiment data from all three sources. The positive returns may be due to weighting methodology that used correlation and that various sources captured all information that was predictive of the stocks' prices. Moreover, the slightly negative shape ratio and comparable maximum draw-down shows that the strategy could be risky. We have included the top five performing sentiment models in the results as seen in the last five rows in Table 2.

5.3 Stock Volatility

As a baseline to confirm the low-volatility anomaly, we simply ranked the stocks in the S&P 1500 investment universe by volatility and bought the top 10 stocks with the lowest volatility. This yielded a 71.26% total return, 0.7 beta, 1.2 sharpe ratio, and -15.35% maximum drawdown over the given time regime in Table 2. The total returns seemed to have relatively linearly increased over time (Figure 1).

5.4 Combined Models

Stock Volatility Plus Overall Quality Model The model selected top 50 high stock return volatility companies. Then, of the top 50 high volatile stocks, it took a long position in the top 10 high quality stocks and shorts bottom 10(low quality stocks), generating a positive total return of 9.92% (Table 2) and sharpe ratio of 0.48. This shows a useful signal from filtering stock volatility by quality (Figure 2). Amidst chaos, there are gems! In the not so good high stock return volatile stocks, a strategy that picks the high quality but volatile stocks wins. This is intuitive in the sense that, the strategy profits from inconsistencies here- high volatile in high quality firms(high cash flow relative to earnings, high profitability growth, low leverage stocks). This inconsistency won't last long as stock prices tend to rivert to the mean, hence correcting the anomaly in the process. This strategy sported this inconsistency and reaped the profits, although it does not beat the market .

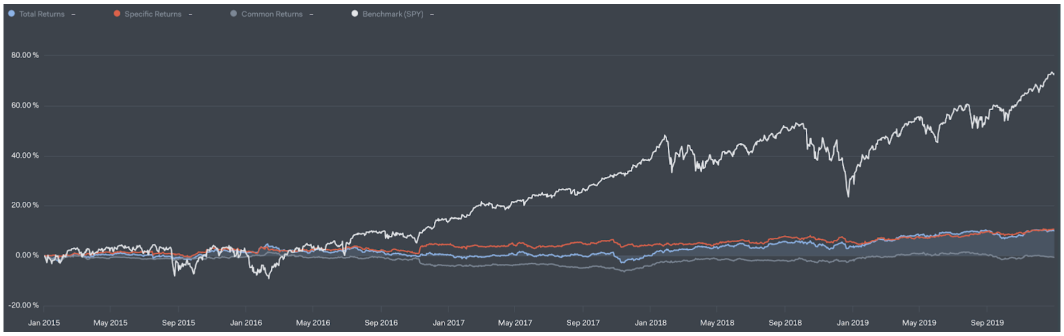

Figure 2: Results of returns for high volatility and quality model from January 1, 2015 - December 31, 2019. Total returns (blue) had similar or lower performance to specific returns (red) over time.

Stock Volatility plus Quality-by-Sector Model Next, high volatility plus quality by sector strategy took the top 50 companies with high stock volatility. We then longed the top 10 and shorted the bottom 10 based on quality per sector (Table 2). This yielded higher total returns of 29.65 percent than high stock volatility and overall quality factor model above. However, it had a smaller shape ratio since sector based strategies expose an investor to sector-specific risk. As explained in the overall quality case above, as prices mean rivets, the strategy reaps the profits. The higher total returns could be explained by different mean reversion rates for different sectors. Information efficiency is different for different sectors, company sizes and markets. When the strategy is deployed at sector level, this differences in speed to information-efficient prices per each each stock in different sectors is exploited by the strategy.

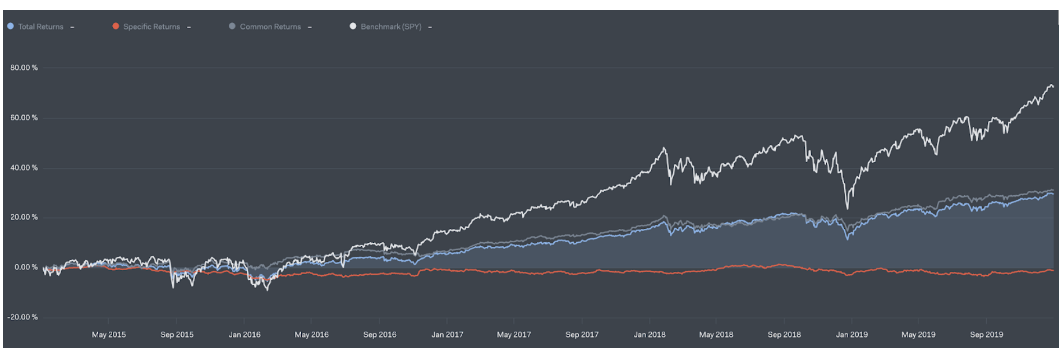

Figure 3: Results of returns for high volatility and quality-by-sector model from January 1, 2015 - December 31, 2019. Total returns (blue) gradually increased over time, but was relatively comparable to common returns.

High Sentiment Volatility Plus Quality-by-Sector Factor Model In this setting, the model utilized volatility of sentiment based on PsychSignal's ratio of number of bull messages over bear messages (Table 2) in conjunction with fundamentals quality grouped by sector. The strategy involves taking a long position in the high sentiment volatility stocks and shorting the low volatile ones. We believe that the positive returns could be attributable to the time it takes to process text data , especially considering increased information flow. Analysts generally have to make sense of the text data and if the information arrival is quicker and comes with volatile sentiment-packed connotations, this creates opportunity to make money by trading on its mood volatility. The value is even more prevalent in smaller stocks that have little analyst coverage. This creates exposure to sector-specific risks and returns, creaming off some diversification benefits that comes with taking a position in the broader economy. This model yielded the highest total returns of 228.71% , way above all models considered. However, it also had the highest maximum drawdown, high beta of 5.57 and a relatively low sharpe ratio of 0.81, suggesting greater risk. The risk comes from sector-specifics and information arrival, which is generally not regular given significant company events happen at times that can not be determined in advance. In short, number and significance of events that is contained in text data happen at irregular times with unknown probability. This stochastic nature of news arrival and events importance drives a lot of the volatility associated with this trading strategy, and hence high volatility of the portfolio formed that drove down the sharpe ratio. News arrival differs per sector, with some sectors having exposure to media coverage than others. For example, the Tech and financial sectors are always on the spotlight with public policy makers, considering global financial crisis and DotCom bust that drove global economies into recession, hence are susceptible to quicker information flow.In short, high portfolio re-balancing frequency, as shown by high draw down, made the portfolio more risky and costly as well.

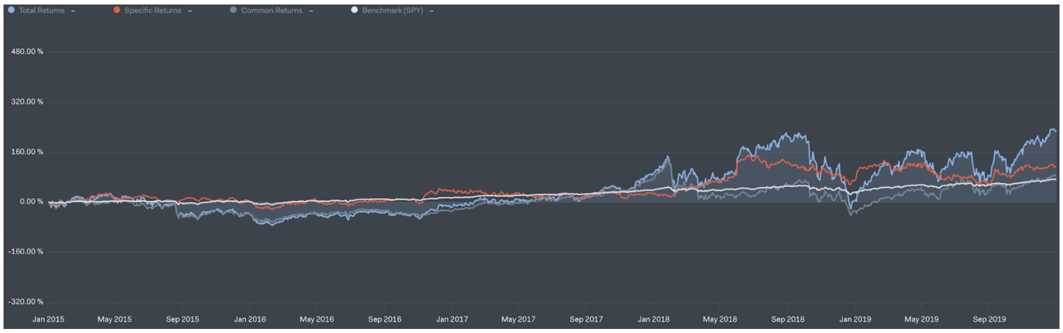

Figure 4: Results of returns for sentiment volatility and quality by sector model from January 1, 2015 - December 31, 2019. Total returns (blue) was experienced a notable increase between 2017 and 2018, passing the benchmark.

5.5 Sector Attribution

As seen from Figures 3 and 4, we can see that total returns for quality-by-sector had higher returns in the combined models. Since the high sentiment volatility plus quality-by-sector strategy outperforms all other models, further analysis on performance attribution might be very valuable. We investigate the sector attributions of our high volatility strategy by analyzing the time-varying sector exposures.

Stock Volatility plus Quality Factor Performance Attribution Figures 5, 6 show the high stock volatility plus overall quality strategy's exposures to utility sector, industrial sector, and consumer defensive sectors respectively. We can see that the exposures to these three sectors vary greatly during the back-test period. In contrast, as shown in Figure 8, the exposures to other sectors including basic materials, health care, technology, communication, real estate, and energy was negligible over the whole backtesting period

Utilities and consumer companies, especially consumer staples are generally quasi-defensive stocks, hence less volatile when compared with industrial stocks. Defensive stocks' beta is lower than industrial stocks'. Industrial sector is generally cyclic and moves with the business cycle. Companies in the consumer defensive sector produce products that are essential for everyday use, which includes food and household and personal products. Similar to the utilities, defensive consumption sector tend to have a low beta. Due to low market risk, utilities and consumer defensive sectors would both perform better than the broader market during recessions. Such an overall sector exposure structure secures a stable portfolio performance and avoids large fluctuations. However, a large loading on low-beta sectors would largely constrain the profitability of this strategy as well, as it limits upside potential during boom times.

Generally, the 2015-2019 period has been a boom market. The stock market had fully recovered from 2009-2010 financial crisis, so a trading strategy that had positive exposure to cyclical stocks in the industrial sector produced high returns but at a higher volatility.Utilities are generally low volatile stocks, but consumer stocks have both components of being cyclic and defensive. With consumer income recently increasing, high returns were obtainable from this sector, hence explains the muted positive exposure between 0-20% allocation as recommended by the algorithm. Utilities are generally considered to perform as value stocks, and consumer stocks showing both growth/value attributes, this provided the diversification benefits that were well documented in research by Kozlov et al. [6].

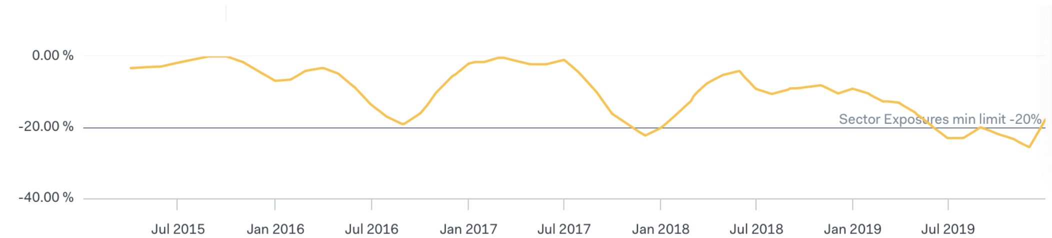

Figure 5: Utilities sector exposure of the high-volatility-quality strategy. Exposure decreased over time.

Figure 6: Industrial sector exposure of the high-volatility-quality strategy. Exposure remained positive for a majority of time.

Figure 7: Consumer defense sector exposure of the high-volatility-quality strategy. Exposure gradually decreased over time.

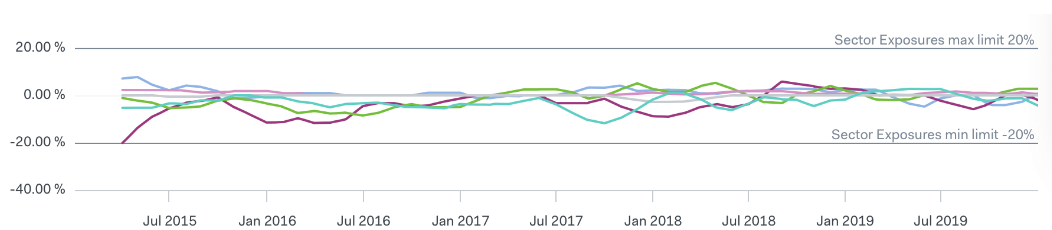

Figure 8: Strategy exposure to other sectors

High Sentiment Volatility Plus Quality-by-Sector Factor Performance Attribution For the high sentiment volatility plus quality-by-sector model, we can see from 9 that this strategy has positive exposures on every sector with comparable magnitudes ranging from 2% to 5%. It means that the sector strategy is taking advantage of the heterogeneous sector information when making trading decisions. As explained above, different sectors differences in digitisation stages, information flow and public attention, hence differences in the mood volatility as information reaches the digital platforms. Moreover, such a mixed structure in sector exposures also partly mitigates sector risks by diversifying among sectors, which keeps the maximum draw-downs relatively smaller while greatly boosting portfolio returns.

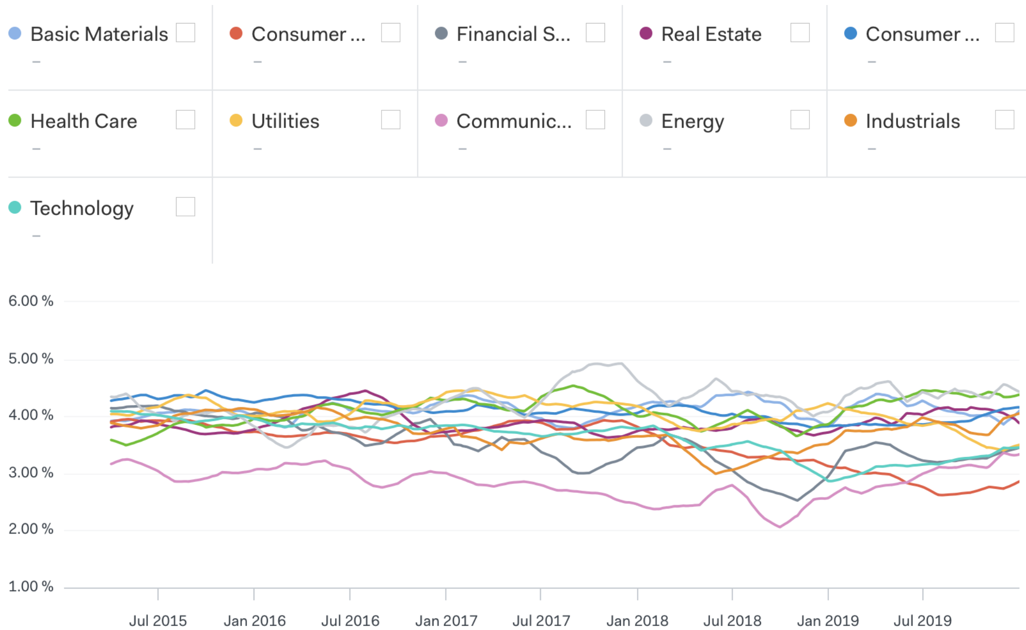

Figure 9: Strategy exposure to other sectors.

5.6 Effect of Capital, Slippage, and Cost per Share

Trading cost influences portfolio's performance. In addition, more capital deployed in one strategy tend to make the market move against your strategy when you try to build a position in a stock, especially when the slippage is large. Here, we extensively investigate how levels of capital, commission per share, and slippage influence our volatility strategies' performance. We vary the capital level from 10 million to 1 billion, the commission per share from zero to 50 basis points, and the slippage from zero to 250 basis points. Again, our back-test time period is from January 2015 to December 2019. The table 3 summarizes the results of our investigation with the portfolio returns and maximum drawdowns in the brackets. For the strategy longing top 10 high quality but volatile stocks,

High-Volatility + Quality			
Capital = 10 Million	CPS = 0	CPS = 25 bps	CPS = 50 bps
Slippage= 0	9.92% (7.19%)	9.81% (7.22%)	9.71% (7.25%)
Slippage= 25 bps	9.86% (7.21%)	9.76% (7.23%)	9.65% (7.27%)
Slippage= 250 bps	9.38% (7.34%)	9.28% (7.37%)	9.17% (7.39%)
Capital = 100 Million	CPS = 0	CPS = 25 bps	CPS = 50 bps
Slippage= 0	9.92% (7.19%)	9.82% (7.23%)	9.71% (7.25%)
Slippage= 25 bps	9.87% (7.21%)	9.76% (7.24%)	9.65% (7.27%)
Slippage= 250 bps	9.39% (7.34%)	9.28% (7.37%)	9.17% (7.40%)
Capital = 1000 Million	CPS = 0	CPS = 25 bps	CPS = 50 bps
Slippage= 0	9.92% (7.19%)	9.82% (7.22%)	9.71% (7.26%)
Slippage= 25 bps	9.87% (7.21%)	9.76% (7.24%)	9.65% (7.27%)
Slippage= 250 bps	9.38% (7.34%)	9.28% (7.37%)	9.17% (7.40%)

Table 3: Analysis on total returns and maximum drawdowns based on slippage and capital

we can see that the level of capital has negligible influence on the portfolio performance. It seems that this strategy has no constant returns to scale. While, the increase in commission per share and slippage both worsen the portfolio performance by slightly decreasing the total return and increasing the maximum drawdown. Notice that the variation in the portfolio performance is small even with great changes of slippage level and commission level. Also, the Sharpe ratio of our strategy is also every stable - ranging from 0.45-0.47. All of the metrics imply that our strategy has a very stable performance.

High-Volatility + Quality by Sector			
Capital = 10 Million	CPS = 0	CPS = 25 bps	CPS = 50 bps
Slippage= 0	29.65% (8.79%)	29.58% (8.79%)	29.51%. (8.79%)
Slippage= 25 bps	29.62% (8.79%)	29.54%. (8.79%)	29.47% (8.8%)
Slippage= 250 bps	29.28% (8.8%)	29.21% (8.81%)	29.14% (8.81%)
Capital = 100 Million	CPS = 0	CPS = 25 bps	CPS = 50 bps
Slippage= 0	29.65% (8.79%)	29.59% (8.79%)	29.51%. (8.79%)
Slippage= 25 bps	29.63% (8.79%)	29.55% (8.79%)	29.48% (8.79%)
Slippage= 250 bps	29.29% (8.8%)	29.22% (8.8%)	29.14% (8.80%)
Capital = 1000 Million	CPS = 0	CPS = 25 bps	CPS = 50 bps
Slippage= 0	29.65% (8.79%)	29.59% (8.79%)	29.52% (8.79%)
Slippage= 25 bps	29.63% (8.79%)	29.55% (8.79%)	29.48% (8.79%)
Slippage= 250 bps	29.29% (8.8%)	29.22% (8.8%)	29.14% (8.8%)

Table 4: Analysis on total returns and maximum drawdowns based on slippage and capital by sector

For the strategy that longs the top high quality but high sentiment volatility stocks per sector, the level of capital again has little influence on the portfolio performance. The increase in the commission fee and slippage levels both cause a smaller total return and a larger maximum drawdown, which is quite intuitive. It is worth noticing that increasing trading cost and slippage have smaller influence on the portfolio performance for this strategy compared to the strategy without considering sectors. Also, the Sharpe ratio of this strategy stabilizes in a range from 1.09 to 1.1 despite the large variations in capital, trading cost, and slippage. As explained above, the portfolio metrics show that the volatility and quality strategy by sector performs even more stable than the one ignoring sector effects.

6 Conclusion and Future Work

We presented baseline one factor only models and composite-style factor models. Overall, our highest performing model utilized a high-volatility based on sentiment plus quality-by-sector method. This sentiment volatility model had over two-fold total returns. However, it's important to note that the strategy generates high beta, hence susceptible in market crisis. Notably, much of our preliminary testing in shorter or earlier time regimes did not translate well to the 2015-2019 time regime. This makes sense given temporal financial events and differences in economic regimes that are normally defined by different economic fundamentals. Quality anomaly only and sentiment only models generated low to negative total returns. However, the stock volatility only model and composite models yielded modest to high positive returns, especially the one based on sentiment volatility. With sentiment volatility, we combined it with the standard stock returns volatility.

Further research to build on this work would benefit if focus is made on curating own dataset and utilizing more proprietary measures of quality and sentiment scores. Beyond the desire to build more advanced strategies, it would help to perform additional risk analysis and performance attribution to understand the risk return decomposition between sectors, time periods, and market capitalization. Additionally, we'd like to have a better understanding of trading this strategy in real-time high frequency trading environment, given different risk scenarios and market regimes.

References

[1] Maria Vieira and Hugh Genin. Volatility, returns and the shape ratio in domestic and global equity markets. 2019.

[2] Lund. Are we in a corporate debt bubble? In *Project Syndicate, June*, 2018.

[3] Johan Bollen, Huina Mao, Xiao-Jun, and Zeng. Twitter mood predicts the stock market. In *Journal of Computational Science*, 2010.

[4] Tim Loughran and Bill McDonald. When is a liability not a liability? textual analysis, dictionaries, and 10-ks. In *Journal of Finance*, 2009.

[5] Augustin Landier Guillaume Simon Jean-Philippe Bouchaurd, Stefano Ciliberti and David Thesmar4. The excess returns of "quality" stocks: A behavioral anomaly. 2016.

[6] Max Kozlov and Antti Petajisto. Global returns premiums on earnings quality, value and size. 2012.

[7] Heeyoung Lee, Mihai Surdeanu, Bill MacCartney, and Dan Jurafsky. On the importance of text analysis for stock price prediction. In *Stanford CS 224N*, 2014.

[8] Bryan Kelly Zheng Tracy Ke and Dacheng Xiu. Predicting returns with text data. 2019.