Trading on mood?

A Work Project, presented as part of the requirements for the Award of a Master's degree in Finance from the Nova School of Business and Economics.

Trading on mood? Analysing the efficacy of sentiment and behavioural biases in predicting stock market movements and investor behaviour.

Pete E. Andrews

Work project carried out under the supervision of:

Gonçalo Sommer Ribeiro

20-May-2021

Abstract

Traditional finance theory rests on the assumption that investors are rational in aggregation. However, a wealth of behavioural finance research has shown this not to be the case. This paper examines whether biases that have been evidenced to impact individual's behaviour can be witnessed in the stock market as a whole, and whether these can be utilised as a bellwether to future price changes. The results mirror the findings of the behaviourists, evidencing a susceptibility to biases among investors, and a promising forecasting ability when incorporated into a systematic volatility trading model.

Keywords: sentiment, behavioural finance, biases, systematic trading

Introduction

GameStop. SPACs. Hertz. AMC. Over the past year, we have witnessed evidence of frothy markets. However, the question remains: are these *isolated* incidents of irrational myopic behaviour? If not, which biases and heuristics are inhibiting investors from engaging in rational thought? This paper will seek to identify which behavioural predispositions are impacting investing patterns and whether these can be utilised in a systematic trading model. Furthermore, the model will incorporate social media stock sentiment to account for investor mood. This shows initial promise; from Figure 1.1 we can see a significant drop in S&P 500 investor sentiment leading up to the selloffs in late-2018 and early-2020, as shown in Figure 1.2.

This thesis is divided into three stages. Firstly, it utilises a vector autoregressive (VAR) model run over the period 02/05/2018 – 30/04/2020, incorporating all of the specified variables (Stage 1). This seeks conclusions about investor behaviour. The next two stages take a practitioner's standpoint, looking at utilising behavioural biases and sentiment as part of a systematic trading model. The former of these takes a condensed subset of the variables which are then employed in a looping VAR engine to forecast future price movements (Stage 2). Within this model, due to the absence of some of the more data-sparse variables, we are able to test over a longer time period (16/09/2013 – 30/04/2020). Finally, due to the emergence of good accuracy for volatility movements, the final stage looks into utilising the model to trade volatility futures (Stage 3).

The results for Stage One show that investors are vulnerable to a multitude of biases. However, not all of these serve as a valid harbinger for future stock market price movements, as demonstrated in Stage Two. Conversely, when these biases are incorporated into a volatility trading model, as done in Stage Three, there is significant value-added.

Literature Review

Traditional Finance

Traditional finance theory, rooted in neoclassical economics, assumes that individuals make investment decisions based on rational process and complete information which allows them to arrive at an optimal choice. The assumption of *homo economicus* aligns itself with numerous stock pricing theories, namely the Random Walk Hypothesis and the Efficient Market Hypothesis. At the market level, similarly to the individual investor level, traditional finance assumes that prices incorporate all available information.

Random Walk Hypothesis. The theory Fama (1965) stipulates that past information cannot be used to predict stock price movements, that future fluctuations are random. This has been implied quantitatively by the likes of Cootner (1962) and Kendall (1953); each demonstrated seral correlation coefficients of consecutive price movements to be close to zero.

Efficient Market Hypothesis (EMH). The theory (Fama 1970) suggests that prices reflect all available information. This is mirrored in Fama et al. (1969); they demonstrate that new information is rapidly, although not always instantaneously, reflected in the market. As a result, the price indicates the fair value to an investor. It is unlikely that an individual can predict this value with complete accuracy. However, in keeping with the law of large numbers, the actions of abundant participants result in a randomly deviating price around its true value.

Behavioural Finance

Behavioural biases can be separated into cognitive errors (CE) and emotional biases (EB). The former corresponds to information-processing or memory errors whilst emotional biases arise from personal mood and faulty intuition. These are well documented at the individual level; however, the question to understand is whether these proliferate in a macro form. If, in aggregate, these biases cancel, then the EMH may not be violated. However, in contradiction to traditional theory, an abundance of market anomalies persist: “irrational exuberance” during bubble periods (Shiller 2000), the Halloween effect (Bouman and Jacobsen 2002), investor overreaction (De Bondt and Thaler 1985) and weather-related trading activity (Hirshleifer and Shumway 2003). These systematic deviations cannot be explained by rational asset pricing.

Conservatism Bias (CE). Conservatism bias (Edwards 1968) describes how individuals maintain prior views and inadequately incorporate new information. They overweight initial beliefs and under-react to new information. Wu et al. (2009) relate this effect to markets. They demonstrate how investors react slowly to earnings news over the medium term. In conjunction with their paper, we will utilise earnings-per-share (EPS) to determine conservatism amongst investors. This will be used to evaluate instances where price revisions lag earnings news, thus indicating conservatism bias. However, whereas Wu et al. (2009) have used quarterly EPS, we have incorporated daily change in EPS to study the bias over shorter periods. This is possible through aggregation of EPS over S&P 500 constituents. In order to smooth spikes from earnings seasons, we have taken a rolling monthly average of EPS.

Confirmation Bias (CE). Individuals are also vulnerable to confirmation bias, looking for confirming rather than disconfirming information. Pouget et al. (2017) reason that analysts are subject to confirmation bias with Bouchaud et al. (2019) indicating stickiness in forecast revision. As a result, an analyst who possesses a positive outlook will be reluctant to change that view, even after subsequent negative information flow. Building on work from Chen (2020), we will utilise the daily change in analyst targets relative to the S&P 500 level. Through a VAR model, the reactive effect of analyst targets will be evident. Specifically, in the presence of a confirmation bias, we would expect there to be a strong impact on analyst targets in response to changes in the same variable from prior periods.

Representative Bias / Base Rate Fallacy (CE). People tend to group new information based on past experiences believing these categorisations are appropriate and thus place undue weight on them. Chan et al. (2004) hypothesise that representative bias leads to firms which exhibit high (low) financial performance in the short-term being designated high (low) growth; investors thus disregard the base rate fact that only a minority of firms can achieve high long-term growth rates. Chan et al. (2004) analysed quarterly sales data and evaluated whether investors view this as representative for the company in question. Our analysis will take on two deviations from Chan et al. (2004); firstly, we will utilise shorter-term changes (daily). This is made possible through aggregation of index constituent sales. Secondly, through using an index, the nature of the bias differs. Namely, we will be analysing whether investors use strong sales growth from companies reporting on a particular day as representative for reporting in subsequent days. We have carried out the same exercise as with EPS, taking average monthly sales. Where the bias is present, we would expect investors to use sales data from prior days as representative, thus impacting subsequent returns in the same direction.

Availability Bias (CE). "What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention" (Simon 1971, p. 40). The availability bias (Tversky and Kahneman 1973) describes the mental heuristic which leads individuals to assign salient events which easily come to mind as having greater frequency. In the media, this process is worsened through "availability cascades" (Kuran and Sunstein 1999); the cascade is self-sustaining and exacerbated by "availability entrepreneurs", media outlets feeding off excessively positive or negative news flow. Gadarowski (2002) mirrored these findings; in his paper he found that stock news coverage complements book-to-market when identifying overvalued companies. Building on his findings, and in line with his suggestion, we have turned to alternative information sources to measure the bias. Through utilisation of the volume of tweets relating to the SPDR S&P 500 ETF Trust (SPY), we can measure the dissemination of market-related information through social media. Specifically, we will look at how the volume of tweets changes day-to-day, with a greater volume of tweets depicting a higher degree of availability bias, and how this impacts returns.

Familiarity Bias (CE). A further bias amongst investors is the tendency to favour stocks familiar to them, reducing the weight placed on fundamental analysis. De Vries et al. (2017) find that investors are more likely to invest in familiar corporate brands. In addition, Welch (2020) finds that retail investors on Robinhood showed a preference for larger companies and "experience" stocks during the coronavirus-induced downturn. We will employ Robinhood data as an indicator of familiarity bias. More specifically, we will look into how the ratio of FAANG (Facebook, Amazon, Apple, Netflix and Google) stock holdings relative to the overall number of holdings in S&P 500 stocks changes over time. As there tends to be a flight to safety in market downturns, we would expect the familiarity bias to accelerate in these instances.

Loss Aversion (EB). Tversky and Kahneman (1979) demonstrated in their pivotal paper that individuals are vulnerable to loss aversion; they are more sensitive to losses than gains with utility being disproportionately affected. Evidence suggests that people are about twice as sensitive to the losses (Blake et al. 2019). This effect is intensified due to the framing effect in which individuals respond more strongly to loss frames. In choosing between foregoing a gain or incurring a loss, the former is preferable (Thaler 1980); the two choices may be economically equivalent, but they are not emotionally equivalent. This leads to a concave value function in which utility diminishes faster in loss scenarios than it appreciates in gain scenarios.

Due to the asymmetry in the utility function arising from loss aversion, Low (2004) posits that the bias manifests itself in a greater responsiveness of perceived risk to negative, relative to positive, price pressure. Increases in market prices lead to a relatively subdued decrease in perceived risk compared to the opposite effect observed in falling markets; fear acts fast, yet exuberance accumulates slowly. Therefore, SPY volatility will be used as a proxy for loss aversion. We will be evaluating whether volatility sensitivity is semi-dimensional. We would expect positive changes in returns to increase volatility and thus result in a positive coefficient of returns on volatility. However, we would also expect negative returns to increase volatility which would lead to a negative coefficient. In the presence of loss aversion, we would expect this relationship to be imbalanced in favour of downside volatility, resulting in a negative coefficient as a result. In conjunction with French et al. (1987), we will utilise a rolling monthly value due to increased accuracy and precision versus longer window periods.

Overconfidence (EB). Daniel et al. (1998) demonstrate that investors are inherently overconfident. Neuroeconomics indicates that this bias is due to our innate biology; in the CFA Institute (2019, p. 19-20) book, they describe how dopamine is released in response to the expectation, and in the event of, a reward. This euphoric feeling may distract individuals from rational thought and lead to excessive risk-taking, thus providing an explanation for hubris and overreaction in the short-term. However, if the expected reward does not arise, the lack of dopamine released, and subsequent depressed state, may cause investors to become impulsive and attempt to make back their losses through further risky investing. In both cases, this biological process results in excessive trading. This is backed up by the findings of Barber and Odean (2001). They find that overconfidence leads to greater trading frequency and a lower payoff as a result. The difference is substantial. The most active quintile of participants averaged a *monthly* turnover of 9% with an average annual pre-tax return of 10% versus an average *annual* turnover of just 1% and annual return of 17.5% for the least active quintile. Therefore, in order to measure overconfidence in aggregate, this model will utilise SPY turnover, where turnover is the monetary value of shares traded per day. Specifically, we would like to identify the instances in which there is excessive trading, thus indicating overconfidence. Therefore, we will look at the deviation from the 52-week average. In line with Barber and Odean (2001), we would expect overconfidence to negatively impact returns. In addition, it will help identify which direction of returns exacerbates overconfidence to a greater degree.

Herding (EB). Humans are evidenced to not always be wholly rational; preconceptions and biases can lead to suboptimal decision making (Kahneman and Tversky 1979). The efficient market rebuttal stipulates that these biases cancel in aggregation. However, Krugman (1997)[1] describes money managers as "an extremely dangerous flock of financial sheep" raising questions of whether these errors are, in fact, systematic in nature.

The likes of Grinblatt et al. (1995), Wermers (1999) and Cipriani and Guarino (2008) demonstrate an empirical basis for the co-movement of investor behaviour. This herding implies profound effects on prices. Dornbusch and Park (1995) contend that investor behaviour is affected by past returns; stocks that have recently appreciated experience increased buying pressure with the opposite holding true for depreciating stocks. This 'positive feedback trading' culminates in herding behaviour. This is mirrored in the findings of Wermers (1999) who finds that there is a considerably higher degree of herding in stocks that have experienced substantial returns in the prior quarter; this occurs on the upside and the downside. As a result, the momentum factor premium serves as a good proxy for herding. The iShares MSCI USA Momentum Factor ETF (MTUM) will be used in place of momentum returns whereas the SPY will replicate market returns. This conforms to the findings of Choi and Skiba (2015) who find evidence that investors herd more in markets with high levels of information transparency.

However, this herding effect may not always be due to an apocryphal consensus. As Nassim Taleb writes in The Black Swan, "the members of the group can be ostracized together – which is better than being ostracized alone" (Taleb 2010, p. 94). This strength in numbers thinking was also noted by the revered John Maynard Keynes: "Worldly wisdom teaches that it is better for reputation to fail conventionally than to succeed unconventionally" (Keynes 2012, p. 158). Both authors hint at the regret aversion bias – the apprehension of going out on a limb in fear of reprimand. Koening (1999) argues that herding can be the consequence of this aversion. Firstly, investors may hold positions for an excessive period of time; they do not want to sell in case they appreciate. Secondly, during periods of depressed prices, they may incorrectly stay out the market too long. Finally, allocations may be made in popular assets to help alleviate accountability in the event that the investment turns sour. Therefore, these effects are neatly encompassed through incorporation of the momentum factor. In line with Dornbusch and Park (1995), we would expect to see higher levels of herding in rising markets.

Endowment Bias (EB). This emotional bias arises from people valuing an asset more when they own it versus when they do not (Samuelson and Zeckhauser 1988). This is inconsistent with traditional finance that stipulates that the price at which an individual is willing to buy and sell the good should be the same. Kahneman et al. (1991) convincingly demonstrated this to be the case; in their experiment asset owners placed a valuation on their objects about twice as high as the price at which buyers were willing to pay. However, List (2003) demonstrated that this effect is eliminated when owners view their objects as carriers of value, as is the case with experienced traders, so it will be insightful to examine in a market setting. The endowment effect only applies to traders who are long the asset; traders with zero inventory or a net short position will be void of the bias. Furche and Johnstone (2006) therefore examine how the existence of the endowment effect culminates in the average ask price exceeding the best (lowest) ask by more than the best (highest) bid price exceeds the average bid. Therefore, the endowment bias will be represented as follows; in rising markets, both the average ask and bid will increase. However, in the case of an increasing endowment bias, the average ask will rise at a faster rate and will thus be represented by an increasing bid-ask spread. The same argument holds for a static market. In falling markets, an increase in the endowment bias will be demonstrated by a fall in the ask to a smaller degree than the decrease in the bid.

Based on the findings by Furche and Johnstone (2006), we will employ the Vanguard S&P 500 ETF (VOO) average bid-ask spread as a proxy for the endowment bias. We have utilised the VOO due to SPY data collection issues. Intuitively, we would expect an increase in the bias to positively impact returns due to the asymmetrical impact on the buy side and sell side.

Sentiment Trading

Shefrin and Statman (1994) looked to develop an alternative behavioural asset pricing model with an additional stochastic discount factor incorporated into their model to reflect behavioural components. Shefrin (2008) postulates that sentiment leads to deviations from fundamental values, as determined by traditional finance, through this additional risk premium. Whereas Shefrin (2008) explores analysts' forecast dispersion as a proxy for the sentiment premium, we will seek a more dynamic measure through social media-derived data.

The likes of Rakowski et al. (2021) have demonstrated the promising correlation between Twitter content and stock movements. In addition, performing natural language processing on social media content to obtain innovative sentiment information has been conducted by the likes of Bollen et al. (2011) and Gruhl et al. (2005). Each found good predictive power, with Bollen et al. (2011) demonstrating an accuracy of 88% in predicting the daily changes in the Dow Jones Industrial Average. However, they tested a relatively short period from February to December 2008. Alongside sample size issues, this was a volatile period and thus does not evidence whether predictive power persists during normality. The periods analysed within this paper will be considerably longer. Furthermore, Bollen et al. (2011) recognise that their results "offer no information on the causative mechanisms that may connect public mood states with DJIA values" (Bollen et al. 2011, p. 7). As a result, this paper will seek to resolve this missing link through utilisation of behavioural bias proxies.

In addition, we have constructed an aggregated sentiment index (ASI). The ASI equally weights 19 various sentiment indicators, both behavioural and technical. These can be seen in Table 1.1. Bathia et al. (2016) utilise the University of Michigan Consumer Confidence Index in demonstrating the impact of sentiment on G6 countries. Furthermore, some (Huth et al. 1994) have found that sentiment measures can be indicative of future buyer behaviour. Therefore, the indicators could serve as a valuable addition for estimating market movements. Bathia et al.

(2016) found that investor surveys were the most valuable and consistent sentiment measures when forecasting future returns. As a result, numerous investor surveys are incorporated within the behavioural indicators; these indicators look at 'soft' data, canvassing consumer and industry opinions and confidence rather than activity. These measures tend to be more infrequent, predominantly being published every month or quarter. The infrequency of the 'soft' data requires the measure to be bolstered by more regular data-rich technical indicators. The technical indicators are calculated based on raw financial and economic data, i.e., the total allocation to stocks, bonds and cash, credit spreads, initial jobless claims, equity options etc. During the back end of the stock market pandemic recovery, public and industry sentiment recovered even in the face of weak economic data. Conversely, during market dips, surveys and questionnaires tend to be a lagging indicator. Therefore, the two indicator subgroups taken together provide a more cohesive and diversified measure of the public state of mind.

Across both measures, we would expect sentiment to lead returns. In addition, it will be interesting to see how sentiment interacts and impacts the other biases.

Data

Sentiment Data. In line with Rakowski et al. (2021) and Manfield et al. (2017), we will analyse public sentiment through PsychSignal. PsychSignal use natural language processing to analyse millions of Twitter posts in order to produce sentiment scores relating to each individual security. The propagation of *cashtags* (e.g., '$GOOGL'), stock-specific hashtags, on social media has allowed platforms like PsychSignal to more accurately analyse stock-related text.

Natural language processing draws from both computer science and computational linguistics to enable a computer to derive meaning from human language. It takes a volume of text, divides it up into smaller parts (tokenisation), attempts to understand the interactions and connections between these parts and then explores how these fit together in order to convey meaning. This is then used in conjunction with text analysis to perform sentiment analysis.

Sentiment analysis extracts emotional context from written language. After the text is tokenised as mentioned above, it is then cleaned, during which the stop words (i.e., "the", "was") are removed. From a pre-determined lexicon, the remaining words are then classified into whether they are positive (+1), negative (-1) or neutral (0), with the whole text receiving an overall polarity score. PsychSignal also analyse the disposition intensity in order to grade the polarity on a scale. Utilising both factors, PsychSignal assign each tweet a 'Bullish Intensity Score' and a 'Bearish Intensity Score' relaying the positive and negative polarity, respectively. These are then aggregated over the total tweets corresponding to each stock every day.

This paper will exploit the 'Bull Minus Bear' indicator, the 'Bullish Intensity Score' minus the 'Bearish Intensity Score'; a net positive value would indicate bullish mentality whereas a negative value would indicate a bearish tone. Furthermore, the daily number of stock-related tweets is utilised within the measure for availability bias. Due to the S&P 500 not being directly tradable, we employed the SPY ETF when obtaining sentiment data. The SPY ETF has been trading since 1993 and has a high daily trading volume, with an average in excess of 70,000,000. As a result, the average number of tweets per day from 16/09/2013 – 30/04/2020 containing the $SPY cashtag is 1058, totalling 2,552,239. This provides a great basis for analysis. One adjustment was made before analysis; the sentiment data is provided on every day of the week, however, the other variables within the model are only informative on trading days. In order to not lose the predictive information across the weekends, the sentiment scores from the weekend are included in those on the following Monday.

The data for the ASI measure was collected from a variety of sources (Table 1.1). Each variable was then rebased between 0 and 1 to allow for comparison across the measures[2]. After rebasing, the scores were then equally weighted and collated through summation into the ASI.

Stock Market Data. The stock market data, namely the SPY closing price, the MTUM closing price, earnings-per-share data, S&P 500 (SPX) sales growth, SPY turnover, VOO average bid-ask spread and analyst targets, were all sourced from Bloomberg Terminal. The momentum premium and SPY volatility were then calculated implicitly from these variables.

The analyst target is the 12-month index target based on the aggregation of the underlying member price targets. It is calculated by multiplying the average target price for each constituent stock in the S&P 500 by the number of shares in the index for each member, summing this value up over all stocks, and then dividing by the index divisor. Essentially, this creates an analyst index target, weighted proportionally to market capitalisation. The earnings-per-share function utilised returns the aggregate of all equity contributions[3]. For the index sales, we took the price-to-sales ratio for the SPX and divided the current market capitalisation of the SPX by this value. SPX price-to-sales is the aggregated market capitalisation of the index divided by the aggregated sales of all constituent members for the specified period. Dividing the index market capitalisation by this value therefore leaves total index sales. The average bid-ask spread collates all the bid-ask points during each day and returns the average result.

Robinhood Trading Data. The Robinhood trading data was acquired through the Robinhood application programming interface (API). Prior to them restricting access, the API allowed you to collect popularity data, the number of open trades in each stock on the platform. From this, we can ascertain our measure for familiarity bias: how the proportion of FAANG holdings to overall S&P 500 holdings changes over time.

Methodology

Vector Autoregressive (VAR) Model

A VAR model is a multivariate forecasting tool used to determine how different time series impact on each other in a multi-directional way. Each variable is modelled as a function of the past values of itself and the other variables. In a typical univariate autoregressive model, $AR(p)$, the time series is modelled as a function of its previous values:

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} + \epsilon_t \qquad (1.1)$$

In the case of a VAR model, let y_t be a vector with the value of n variables at time t:

$$y_t = [y_{1,t}, y_{2,t}, \dots, y_{n,t}]' \tag{2.1}$$

A p-order VAR process generalises a one-variable $AR(p)$ process to n variables[4]:

$$y_t = G_0 + G_1 y_{t-1} + G_2 y_{t-2} + G_3 y_{t-3} + \cdots + G_p y_{t-p} + e_t \tag{2.2}$$

Stationarity

To be estimated through a VAR model, we need to ensure the time series is stationary. A p order VAR is covariance stationary if the expected value of y_t does not depend on time:

$$E[y_t] = E[y_{t+j}] = \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix} \tag{3.1}$$

Furthermore, the covariance matrix of y_t and y_{t+j} depends on the time lapsed j and not on the time period t. In other words, the covariance between two lag periods may change if the size of the lag changes but if the lag remains the same then the covariance should not change over time:

$$E\left[(y_t - \mu)(y_{t+j} - \mu)'\right] = E\left[(y_s - \mu)(y_{s+j} - \mu)'\right] = \Gamma_j \tag{3.2}$$

Detection. There are numerous ways to test for stationarity (Dickey and Fuller 1979; Kwiatkowski et al. 1992; Phillips and Perron 1988). However, for the purposes of this thesis, we will focus on the Augmented Dickey-Fuller (Dickey and Fuller 1979) test. Each test looks to identify a unit root within the data where a unit root indicates a non-stationary time series[5].

With that in mind we run the `adfuller_test` through the `statsmodels` Python package across every variable[6]. The lag is chosen according to minimisation of the Bayesian Information Criterion (Schwarz 1978). The results for Stage One, Two and Three can be seen in Table 2.1, Table 2.2 and Table 2.3, respectively. Evidently, all variables are stationary.

Multicollinearity

Multicollinearity occurs when an independent variable is highly correlated with one or more of the other independent variables. This can lead to an underestimation of the statistical significance of the independent variable(s) (Allen 1997); the standard error of the independent variable(s) will be inflated, resulting in unstable coefficients which are also less statistically significant. Multicollinearity will be assessed by the Variance Inflation Factor (Woolridge 2016, p. 86)[7]. The rule of thumb for the severity, as indicated by the VIF value, to demand correction remains open for debate. Whereas some (Hair et al. 1995) indicate that 10 is reasonable, we will employ a relatively strict cut-off equal to 5, in accordance with Akinwande et al. (2015).

Within Python, the test was run using the `variance_inflation_factor` function from the `statsmodels` package. The results for Stage One can be seen in Table 3.1. Evidently, none of the VIF values surpass the maximum of 5. The lack of correlation between the variables in Stage One is further demonstrated in Figure 2.1. For Stages Two and Three, please consult Table 3.2 and Table 3.3, respectively. Likewise, none of the variables exhibit multicollinearity.

Order of Lags

When selecting the order of lags (p), we face a trade-off. If p is too short, the model may be poorly specified as residuals may not satisfy their desired properties, i.e., that they appear as white noise. If p is too long, then too many degrees of freedom will be lost (every lag adds n^2 coefficients) and the densely parametrised model can lead to overfitting. If this is the case, the VAR tries to estimate many coefficients with little data and the coefficients are poorly estimated; the model would fit the in-sample data well yet performs poorly out-of-sample. As a result, we can use maximum likelihood estimation across the Bayesian (Schwarz 1978), Akaike (Akaike 1974) and Hannan-Quinn (Hannan and Quinn 1979) Information Criterion and Akaike's Final Prediction Error (Akaike 1969). These tests are rooted in information-based criteria, balancing the trade-off between parsimony and a reduction in the sum of squares. However, having run sample regressions, the latter three consistently return a high number of lags and the consequences specified above are observed. Therefore, more concentration will be placed on the Bayesian Information Criterion (BIC). The full results of each order of lags for

Stage One can be seen in Table 4.1; the BIC is minimised at a lag of 2. Therefore, each equation to estimate is the dependent variable regressed on itself and all the other variables lagged by 1 and 2 trading days[8]. For the second and third stages of the experiment, a looping model will be engaged and will therefore utilise a refreshed order of lags for each repetition.

Forecasting Using VAR

In order to analyse the forecasting accuracy in the second and third stages, we need to split the data into training and testing subsets. The training data will be used to obtain the estimated coefficients[9]. These coefficients are then utilised to forecast the next day's variables, which can be compared against the true value in the testing data. Within the model engine, this is incorporated within a loop to forecast the accuracy over the whole period. For example, the first day in the forecast period is calculated using the estimated coefficients from historical data[10]. At the end of this day, the new data is incorporated into the model to estimate new coefficients and the order of lags. These are then used to forecast the subsequent day, and so on.

Checking for Serial Correlation

Before analysing our results, we want to ensure a lack of correlation amongst errors. To run the serial correlation test, we use the Durbin-Watson statistic (Durbin and Watson 1950, 1951). The value of this statistic varies between 0 and 4. The closer it is to the value 2, then the greater the evidence for no significant serial correlation. The closer to 0, the greater the positive serial correlation, and the closer it is to 4, the greater the implied negative serial correlation. The results for Stage One, Two and Three can be seen in Tables 5.1, 5.2 and 5.3 respectively. Evidently, none of the residuals demonstrate any significant autocorrelation.

Results

Stage One

Stage One of the experiment was to run a VAR model on all of the variables in order to derive insights into investor behaviour; this yielded some interesting results. Firstly, in line with the findings of De Bondt and Thaler (1985), there is evidence of initial overreaction in stock prices to new information with a subsequent correction. This is demonstrated by the highly statistically significant negative coefficient of SPY return, at a lag of 1 trading day, on itself. The coefficient, as seen in Table 6.1, is equal to -0.345 with a standard error equal to 0.0579. This therefore yields a t-statistic of -5.96 and is significant at the 99% significance level. Furthermore, the relatively large negative coefficient also demonstrates economic significance and evidence of mean reversion in stock returns. Furthermore, the highly significant lags of SPY return on itself at t-1 and t-2 provide evidence against the random walk hypothesis and the efficient markets hypothesis that stipulate that future price movements cannot be ascertained from past data. This finding is aligned with overconfidence in that investors overestimate their beliefs based on initial information leading to the subsequent correction. Interestingly, overconfidence was shown to be exacerbated by availability bias, consistent with Kliger & Kudryavtsev (2010). Within the regression results for the SPY turnover deviation from the annual average, as seen in Table 6.2, the positive coefficient at t-1 of tweet volume growth (0.00153) on SPY turnover was shown to be significant at the 95% level. With an average change in tweet volume growth over the whole period of 4.55 tweets and an average change in the SPY turnover variable of 2.21% (0.0221), this result is also economically significant. In addition, overconfidence (excessive trading) was further reinforced through analyst targets. The coefficient of analyst targets on the SPY turnover variable is 40.9 at a lag of one day, thus indicating a reinforcing effect on overconfidence from positive analyst targets. This is significant at the 95% level.

Additionally, there is evidence of familiarity bias (FAANG ownership) positively impacting SPY returns at a lag of 1 trading day, as shown in Table 6.1. With a coefficient of 2.63 and a t-statistic of 2.02, this is both economically and statistically significant (at 95% confidence). This is unsurprising given the fact that S&P 500 companies are among the largest and most well-known in the global economy.

Furthermore, there is strong evidence of the endowment bias positively impacting SPY returns. The positive coefficient of the average bid-ask spread on SPY return, lagged one period, is statistically significant at the 99% level, as shown in Table 6.1. The coefficient, at 0.312, is also economically significant and demonstrates a large impact from this particular bias. However, this result is rather expected; as explained previously, the endowment effect impacts those holding the asset but not those who don't. Therefore, an increase in the endowment bias reduces selling pressure amongst asset owners but does not affect buying pressure amongst prospective owners. However, the VAR model also demonstrates that this endowment bias is reduced in periods of increased herding amongst investors. In Table 6.3, we can see that the coefficients of the momentum factor excess return on the bid-ask spread have a negative impact when lagged over both one and two periods. With values of -0.165 (t-1) and -0.158 (t-2), the negative impact of herding on the endowment bias is evident. These are statistically significant at the 99% confidence level. Behaviourally, the herding effect indicates investors purchase recent winners, abandoning their current assets and overcoming the endowment bias.

Low (2004) postulates that loss aversion is exhibited in a greater sensitivity of perceived risk to downward price pressure. This is demonstrated with conviction in this study. As shown in Table 6.4, the coefficients of SPY return on SPY volatility at lags of 1 and 2 trading days, are statistically significant at the 99% and 95% level, respectively. With values of -0.218 (t-1) and -0.124 (t-2), the negativity demonstrates the asymmetric volatility effect arising from loss aversion. Furthermore, the degree of magnitude in both coefficients demonstrates how economically meaningful they are in reality. Measuring investor utility by the Sharpe Ratio, we can therefore see how for the same size change in return, a greater reduction in utility on the downside corresponds to a higher relative change in volatility.

With respect to availability bias, this appears to be exacerbated by falling sentiment. The coefficient of the ASI on tweet volume growth lagged two periods is -102, as shown in

Table 6.5; this is significant at the 95% significance level. Therefore, falls in sentiment have the inverse effect on the volume of tweets (availability bias). This is an interesting insight providing evidence that availability bias can be more perpetuating in market (sentiment) downturns, relative to upturns. Furthermore, the coefficients of tweet volume growth on itself at lags of 1 and 2 trading days, provide evidence of availability cascades proliferating. As shown in Table 6.5, the coefficients have values of 0.316 and 0.212, respectively. Both values are statistically significant at 99% confidence. Therefore, availability bias appears to have a self-reinforcing effect, providing further credence to the thinking of Kuran and Sunstein (1999).

The results for changes in analyst targets are shown in Table 6.6. Firstly, the results surrounding analyst targets response to EPS and sales seem paradoxical at first glance. At 99% significance, analysts respond positively to EPS growth but negatively to sales growth. However, on closer inspection, this is as expected. If we assume that EPS stays constant and sales fall then margins have been positively affected. This is deemed attractive by analysts and they respond accordingly. The reversed signs on both EPS and sales at a lag of 2 indicate an initial overreaction. With respect to the confirmation bias, adjustment to analyst expectations is exacerbated by confirmatory data, in this case prior analyst targets. At a lag of one day, the coefficient of analyst targets on itself is 0.533. At a lag of two days, this value stands at 0.254. These are both highly significant (>99%) with t-statistics of 11.8 and 5.65, respectively. Not only does this indicate confirmation bias but also a significant degree of target trend following (herding) amongst analysts. Furthermore, following the regression, a looping VAR model was implemented in order to obtain the forecast accuracy of analyst targets. Through this procedure, we obtained a forecast accuracy for positive and negative movements in analyst targets of 77.4%, offering substantial evidence for the predictability of analyst targets. The lack of added value in utilising analyst targets to predict returns in subsequent days is illustrated in Table 6.1, where none of the coefficients of analyst targets on SPY return are both positive and significant.

Finally, the data revealed excessive optimism amongst analysts, with the ratio of analyst targets to the S&P 500 never dropping below 1, even during the vertiginous drops in 2018 and 2020. This indicates ineffectuality in utilising analyst targets, especially during market downturns.

Stage Two

Stage Two was about refining the variables in order to build a trading model over the period. In order to do this, we regressed each variable separately on SPY return in a looping model in order to obtain the prediction accuracy for each. As explained earlier, this involves moving iteratively through subsequent periods. At the end, we have a series of actual values and forecasted values, avoiding any look-ahead bias by only utilising past data. These actual values can then be compared against the forecasted values to obtain a prediction accuracy. Having done this for each variable individually, it was then a case of selecting the variables which obtained a prediction accuracy significantly different from 50%, statistically speaking. The full list of accuracy data can be seen in Table 7.1. The four variables with values statistically different from 50%, as determined by the `scipy.stats.t.interval` function, are Twitter Sentiment (bull_minus_bear), Momentum Excess Return (mom-mkt), SPY Volatility (spy_vol) and the VOO Bid-Ask Spread (bid_ask_spread).

Due to no longer being restricted by the shorter-term nature of the Robinhood data, we can now run the prediction model with these specified variables over a longer period (16/09/2013 – 30/04/2020). The resulting looping VAR model produced forecasts demonstrated in Figure 3.1 with the first 100 days shown in Figure 3.2 for clarity; note, the level of the forecasts does not matter, only the positive-negative alignment with the true values. This stage produced two results of note. Firstly, the SPY return accuracy is weak, at only 51.3%. With a 95% confidence interval of (0.489, 0.537), this result is not statistically different from pure chance (50%). However, the model did produce a positive-negative forecast accuracy of 55.9%

for SPY volatility. This was consistent symmetrically, with 55% accuracy for negative moves and 57% accuracy for positive moves.

Stage Three

As a result, we implemented the model again using 1-month VIX index futures (UX1) instead of SPY returns. These results were then incorporated into a systematic trading model; if the forecasted return for UX1 in the subsequent period was negative, then the trading model goes short, and long when the forecast is positive. This yielded a cumulative return of 286% over the whole period, compared to 71.6% for the SPY and well above a UX1 buy-and-hold return of -94.8% (due to futures rolldown). The comparison of returns can be seen in Figure 4.1. However, due to the high annual volatility of 91.5%, the strategy only achieves a Sharpe ratio of 0.248 versus a SPY buy-and-hold ratio of 0.495 over the same period. However, interestingly, the strategy has a negative correlation of -0.0561 with the market.

Therefore, by providing a buffer in market downturns, the strategy is promising in terms of portfolio insurance. As shown in Table 8.1, by allocating anything between 0% and 60% of a SPY buy-and-hold portfolio to the behavioural futures trading model, the Sharpe ratio of the portfolio can be drastically improved. We can see from Figure 4.2 the difference in returns between SPY buy-and-hold, SPY +5% strategy allocation, SPY +15% strategy allocation and SPY +25% strategy allocation. Evidently, as the allocation to the UX1 strategy increases, returns increase, but in conjunction with volatility. This trade-off reaches its zenith at 15%. The blended combinations perform well over the whole period analysed, as shown by Table 8.2, which depicts the monthly returns for SPY buy-and-hold, SPY +5% strategy, SPY +15% strategy and SPY +25% strategy allocation. We can see from the data that the added insurance of the strategy provides significant protection in downturns. For example, for the downturn from February to March 2018, the SPY returned 6.65% vs. 15.7% for the blended portfolio with 15% invested in the UX1 strategy. We can also see a similar pattern in 2020. In the precipitous drop in February and March 2020, the SPY returned -19.9%; with the 15% protection, the blended portfolio returned -7.06%. These results are mostly down to the increased forecasting accuracy during the periods of market volatility. For example, during the early-2018 period, the model returned an average forecast accuracy of 62% (positive accuracy: 68%, negative accuracy: 61%). The disparity between the strategies is even more pronounced when considering the annual data demonstrated in Table 8.3. Whereas the SPY returned positive gains in 63% of the years during the period, the SPY +5%, SPY +15% and SPY +25% strategies achieved profits in 75%, 75% and 88% of the same years, respectively. Overall, the strategy returns underperform during stable periods, leaking small profits but maintaining a war chest in order to capitalise on market downturns.

Whilst the 15% allocation achieves the highest Sharpe ratio, from a practitioner's point of view there are a number of key variables to consider when analysing the validity of a strategy. These can be seen in Table 8.4. Firstly, drawdown refers to the maximum distance between the portfolios peak return and its nadir before it recovers back to the peak. Drawdowns are fundamentally important due to the asymmetric nature between the drawdown and the gain required to recover the loss. From Table 8.4, we can see that the maximum drawdown, in returns, for the blended strategies is less than the maximum drawdown for the SPY buy-and-hold. However, only the 5% allocation portfolio outperforms the SPY in terms of average drawdown.

Secondly, skewness provides insights into the distribution of returns. Positive skewness indicates frequent smaller negative returns but a small number of large positive returns. Negative skewness implies the opposite. From an investing perspective, we want these extreme returns to be positioned to the positive end of the distribution in order to avoid catastrophic risk. Therefore, the smaller negative skewness of the 5% allocation, relative to the SPY, is appealing. However, the positive values for both the 15% and 25% allocations are even more attractive.

Whereas skewness examines the differences between the tails of the distribution, kurtosis looks at the shape of the tails when taken together. Essentially, it is a measure of how 'fat' the tails are; fatter tails (higher kurtosis) exhibit more numerous extreme returns. From an investing perspective, this is a concern. From Table 8.4, we can see that both the 5% and 15% allocations exhibit kurtosis values lower than the SPY and are therefore relatively attractive.

Conclusion

Stage One demonstrated how investors are vulnerable to a wide range of behavioural biases which, in turn, affect investor behaviour. There are significant implications for individual investors, however, the findings also raise systemic concerns. Falling investor sentiment exacerbates availability bias amongst investors which culminates in an availability cascade due to a self-reinforcing effect. In market downturns, these cascades could be of particular interest due to the subsequent interactions between availability bias and overconfidence. This highlights the contrasting and important difference between dynamic biases such as the availability bias – which change over time – versus biases such as the endowment bias – which are more static in nature. The endowment bias was demonstrated to prop up SPY returns; however, it is ever present and thus loses value when constructing a forecasting model.

When the condensed variables were implemented within a systematic trading model with a view to predict market movements, there was a lack of forecasting ability; sentiment and behavioural biases are not a panacea for prediction that was initially indicated. This pays credence to efficient market proponents who argue that these biases cancel in aggregation. Nevertheless, when approached differently, the predictive power of biases and sentiment can serve as a valuable asset to trading strategies, particularly those dealing with volatility.

There is substantial evidence for the benefits of incorporating behavioural indicators into a systematic volatility trading model. VIX futures are profitable when analysed from this point of view, with the utilised variables exploiting its reputation as the "fear gauge" of Wall

Street. Not only can integrating such a strategy improve the risk-return payoff but can also positively alter the characteristics of the return distribution, thus making it more attractive to prospective investors. This result is robust across a wide range of allocations to the strategy. By introducing behavioural factors into a volatility forecasting model, we were able to significantly outperform the benchmark and construct a drastically improved portfolio, relative to the market. The predictive model was able to successfully forecast changes in volatility, particularly during turbulent periods. This was made possible through incorporation of proxies for behavioural biases, benefitting from the heightened emotion during these episodes as a result.

These findings raise questions moving forward. List (2003) and Barber and Odean (2008) have both demonstrated the greater susceptibility of inexperienced investors to biases with Graham (1999) also confirming the lower vulnerability of practiced professionals to herding. For example, List (2003) showed that amateurs are vulnerable to the endowment bias, yet this is absent in experienced traders. In addition, Barber and Odean (2008) evidenced that professional investors are less prone to the availability bias, being more selective when responding to news. In a recent Charles Schwab (2020) study[11], they evidenced the huge growth in retail investing with 15% of all current U.S. stock market investors starting in 2020. This mirrors the Robinhood data collected for this study; the number of open positions increased from 22,329,872 on 02-05-2018 to 168,360,296 on 13-08-2020. With the prevalent rise in retail investing, coinciding with the higher susceptibility of these investors to behavioural biases, it raises the potential of incorporating such variables into a profitable trading strategy in the future.

Recommendations

Within the paper we have mentioned information (availability) cascades but with no singular measure to account for them. Rakowski et al. (2021) differentiate between information supply (tweets) and information consumption (retweets), providing intuitive insights into the development of information cascades. They demonstrate a multiplicative effect on stock price impact from the presence of information consumption. Due to data limitations, retweet information was not available but would be a beneficial inclusion to future research.

Furthermore, the focus of this study has been on U.S. markets. As mentioned by Bikhchandani and Sharma (2000), emerging markets show a greater tendency to herd, due to limitations surrounding reporting, accounting and regulation. Therefore, future research should look into the vulnerability of investors outside of the U.S. to biases, particularly those in emerging markets. Similarly, the Twitter sentiment data collected focused on English vernacular. With an international investor base participating in the U.S. stock market, a future social media sentiment measure should look to analyse a broad spectrum of languages.

A more structural recommendation for the future would be on model specification. As noted by Taleb (2010), linearity does not exist in “Extremistan”, that is in a world where extreme outliers can have a disproportionate impact, the home of black swans. During market turbulence, the linearity-constrained VAR model is inappropriate, and a recommendation would be to consider non-linear models, like those described by Kock and Teräsvirta (2011).

Given the almost direct relationship between market volatility and market returns, we are left with the question of why the prediction model does not perform well in forecasting stock market changes yet does for volatility. The logical explanation ought to lie in the differences between historical volatility (derived from price action) and implied volatility (derived from expectations, and utilised in VIX futures), and how changes in sentiment and behavioural biases interact differently with the latter; however, this remains a query for further research. Furthermore, the viability of this model rests on the potential impact from both implicit and explicit costs from trading. Due to the relatively high tick size of 0.05 for VIX futures, this issue is worthy of additional investigation.

References

Akaike, Hirotugu. 1969. "Fitting Autoregressive Models for Prediction." *Annals of the Institute of Statistical Mathematics*. 21: 243-247.

Akaike, Hirotugu. 1974. "A new look at the statistical model identification." *IEEE Transactions on Automatic Control.* 19(6): 716-723.

Akinwande, Michael Olusegun, Hussaini Garba Dikko and Agboola Samson. 2015. "Variance Inflation Factor: As a Condition for the Inclusion of Suppressor Variable(s) in Regression Analysis." *Open Journal of Statistics.* 5: 754-767.

Allen, Michael Patrick. 1997. *Understanding Regression Analysis*. New York: Springer US.

Barber, Brad M. and Terrance Odean. 2001. "Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment." *The Quarterly Journal of Economics*. 116(1): 261–292.

Barber, Brad M. and Terrance Odean. 2008. "All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors." *The Review of Financial Studies*. 21(2): 785-818.

Bathia, Deven, Don Bredin and Dirk Nitzsche. 2016. "International Sentiment Spillovers in Equity Returns." *International Journal of Finance & Economics*. 21(4): 332-359.

Bikhchandani, Sushil and Sunil Sharma. 2000. "Herd Behavior in Financial Markets." *IMF Staff Papers.* 47(3): 279-310.

Blake, David P., Edmund S. Cannon and Douglas Wright. 2019. "Quantifying Loss Aversion: Evidence from a UK Population Survey." *Pensions Institute Discussion Paper PI-1912.*

Bollen, Johan, Huina Mao and Xiaojun Zeng. 2011. "Twitter mood predicts the stock market." *Journal of Computational Science*. 2(1): 1-8.

Bouchaud, Jean-Philippe, Philipp Krüger, Augustin Landier and David Thesmar. 2019. “Sticky Expectations and The Profitability Anomaly.” *The Journal of Finance*. 74(2): 639-674.

Bouman, Sven and Ben Jacobsen. 2002. “The Halloween Indicator, ‘Sell in May and Go Away’: Another Puzzle.” *The American Economic Review*. 92(5): 1618-1635.

CFA Institute. 2019. *CFA Program Curriculum 2020: Level III, Volume 2*. Hoboken: Wiley.

Chan, Wesley S., Richard Frankel and S.P. Kothari. 2004. “Testing behavioral finance theories using trends and consistency in financial performance.” *Journal of Accounting and Economics.* 38: 3-50.

Chen, Keqi. 2020. “A Closer Look at Analyst Expectations: Stickiness and Confirmation Bias.” *Journal of Applied Finance & Banking*. 10(5): 281-298.

Choi, Nicole and Hilla Skiba. 2015. “Institutional herding in international markets.” *Journal of Banking & Finance.* 55: 246-259.

Cipriani, Marco and Antonio Guarino. 2008. “Herd Behavior and Contagion in Financial Markets.” *B.E. Journal of Theoretical Economics*. 8(1): Article 24.

Cootner, Paul H. 1962. “Stock Prices: Random vs. Systematic Changes.” *Industrial Management Review*. 3(2): 24-45.

Daniel, Kent, David Hirshleifer and Avanidhar Subrahmanyam. 1998. “Investor Psychology and Security Market Under- and Overreactions.” *The Journal of Finance*. 53(6): 1839-1885.

De Bondt, Werner F. M. and Richard Thaler. 1985. “Does the Stock Market Overreact?” *The Journal of Finance*. 40(3): 793-805.

De Vries, Annalien, Pierre D. Erasmus and Charlene Gerber. 2017. “The familiar versus the unfamiliar: Familiarity bias amongst individual investors.” *Acta Commercii*. 17(1): 10 pages.

Dickey, David A. and Wayne A. Fuller. 1979. “Distribution of the Estimators for Autoregressive Time Series With a Unit Root.” *Journal of the American Statistical Association.* 74(366): 427-431.

Dornbusch, Rudiger and Yung Chul Park. 1995. *Financial Opening: Policy Lessons for Korea*. San Francisco: International Center for Economic Growth.

Durbin, James and Geoffrey. S. Watson. 1950. “Testing for Serial Correlation in Least Squares Regression: I.” *Biometrika*. 37(3/4): 409-428.

Durbin, James and Geoffrey S. Watson. 1951. “Testing for Serial Correlation in Least Squares Regression. II.” *Biometrika*. 38(1/2): 159-177.

Edwards, Ward. 1968. “Conservatism in human information processing.” In *Formal representation of human judgment*, edited by B Kleinmuntz. 17-52. New York: Wiley.

Fama, Eugene F. 1965. “The Behaviour of Stock Market Prices.” *The Journal of Business*. 38(1): 34-105.

Fama, Eugene F. 1970. “Efficient Capital Markets: A Review of Theory and Empirical Work.” *The Journal of Finance*. 25(2): 383-417.

Fama, Eugene F., Lawrence Fisher, Michael C. Jensen and Richard Roll. 1969. “The Adjustment of Stock Prices to New Information.” *International Economic Review*. 10(1): 1-21.

French, Kenneth R., George William Schwert and Robert F. Stambaugh. 1987. “Expected stock returns and volatility.” *Journal of Financial Economics*. 19(1): 3-29.

Furche, Andreas and David Johnstone. 2006. “Evidence of the Endowment Effect in Stock Market Order Placement.” *Journal of Behavioral Finance*. 7(3): 145–154.

Gadarowski, Christopher. 2002. “Financial Press Coverage and Expected Stock Returns.” *Unpublished*.

Graham, John R. 1999. “Herding among Investment Newsletters: Theory and Evidence.” *The Journal of Finance.* 54: 237-268.

Grinblatt, Mark, Sheridan Titman and Russ Wermers. 1995. “Momentum Investment Strategies, Portfolio Performance, and Herding: A Study of Mutual Fund Behavior.” *The American Economic Review*. 85(5): 1088-1105.

Gruhl, Daniel, R. Guha, Ravi Kumar, Jasmine Novak and Andrew Tomkins. 2005. “The predictive power of online chatter.” In *KDD '05: Proceeding of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, 78-87. New York: ACM Press.

Hair, Joseph F., Rolph E. Anderson, Ronald L. Tatham and William C. Black. 1995. *Multivariate Data Analysis*. New York: Macmillan Publishing Company.

Hannan, Edward J. and Barry G. Quinn. 1979. “The Determination of the Order of an Autoregression.” *Journal of the Royal Statistical Society*. Series B (Methodological). 41(2): 190-95.

Hirshleifer, David and Tyler Shumway. 2003. “Good Day Sunshine: Stock Returns and the Weather.” *The Journal of Finance*. 58(3): 1009-1032.

Huth, William L., David R. Eppright and Paul M. Taube. 1994. “The indexes of consumer sentiment and confidence: Leading or misleading guides to future buyer behavior.” *Journal of Business Research.* 29(3): 199-206.

Kahneman, Daniel and Amos Tversky. 1979. “Prospect Theory: An Analysis of Decision under Risk.” *Econometrica*. 47(2): 263-291.

Kahneman, Daniel, Jack L. Knetsch and Richard H. Thaler. 1991. “Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias.” *The Journal of Economic Perspectives.* 5(1): 193-206.

Kendall, Maurice G. 1953. "The Analysis of Economic Time-Series, Part I: Prices." *Journal of the Royal Statistical Society.* Series A (General). 116(1): 11-34.

Keynes, John Maynard. 2012. *The General Theory of Employment, Interest and Money*. New York: Cambridge University Press.

Kliger, Doron and Andrey Kudryavtsev. 2010. "The Availability Heuristic and Investors' Reaction to Company-Specific Events." *Journal of Behavioral Finance*. 11(1): 50-65.

Kock, Anders Bredahl and Timo Teräsvirta. 2011. "Forecasting with Nonlinear Time Series Models." In *The Oxford Handbook on Economic Forecasting*, edited by M.P. Clements and D.F. Hendry. New York: Oxford University Press.

Koening, J. 1999. "Behavioral Finance: Examining Thought Processes for Better Investing." *Trust and Investments*. 69: 17-23.

Kuran, Timur and Cass R. Sunstein. 1999. "Availability Cascades and Risk Regulation." *Stanford Law Review*. 51(683): 683:768.

Kwiatkowski, Denis, Peter C.B. Phillips, Peter Schmidt and Yongcheol Shin. 1992. "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?" *Journal of Econometrics*. 54(1–3): 159-178.

List, John A. 2003. "Does Market Experience Eliminate Market Anomalies?" *The Quarterly Journal of Economics*. 118(1): 41–71.

Low, Cheekiat. 2004. "The Fear and Exuberance from Implied Volatility of S&P 100 Index Options." *The Journal of Business*. 77(3): 527-546.

Manfield, Jonathan, Derek Lukacsko and Thársis T. P. Souza. 2017. "Bull Bear Balance: A Cluster Analysis of Socially Informed Financial Volatility." Presented at *2017 IEEE Computing Conference, London, UK.*

Phillips, Peter C. B. and Pierre Perron. 1988. “Testing for a Unit Root in Time Series Regression.” *Biometrika*. 75(2): 335-346.

Pouget, Sebastien, Julien Sauvagnat and Stephane Villeneuve. 2017. “A Mind Is a Terrible Thing to Change: Confirmatory Bias in Financial Markets.” *The Review of Financial Studies*. 30(6): 2066–2109.

Rakowski, David, Sara E. Shirley and Jeffrey R. Stark. 2021. “Twitter activity, investor attention, and the diffusion of information.” *Financial Management*. 50(1): 3-46.

Samuelson, William and Richard Zeckhauser. 1988. “Status quo bias in decision making.” *Journal of Risk and Uncertainty*. 1(1): 7–59.

Schwarz, Gideon. 1978. “Estimating the Dimension of a Model.” *The Annals of Statistics*. 6(2): 461-464.

Shefrin, Hersh and Meir Statman. 1994. “Behavioral Capital Asset Pricing Theory.” *The Journal of Financial and Quantitative Analysis.* 29(3): 323-349.

Shefrin, Hersh. 2008. *A Behavioral Approach to Asset Pricing*. Boston: Elsevier.

Shiller, Robert J. 2000. *Irrational Exuberance*. Princeton, N.J.: Princeton University Press.

Simon Herbert A. 1971. “Designing organizations for an information rich world.” In *Computers, communications, and the public interest*, edited by Martin Greenberger, 37-72. Baltimore: Johns Hopkins Press.

Taleb, Nassim Nicholas. 2010. *The Black Swan*. Harlow: Penguin Books.

Thaler, Richard. 1980. “Toward a positive theory of consumer choice.” *Journal of Economic Behavior & Organization.* 1(1): 39-60.

Tversky, Amos and Daniel Kahneman. 1973. “Availability: A heuristic for judging frequency and probability.” *Cognitive Psychology*. 5(2): 207-232.

Welch, Ivo. 2020. “The Wisdom of the Robinhood Crowd.” *National Bureau of Economic Research Working Paper Series 27866.*

Wermers, Russ. 1999. “Mutual Fund Herding and the Impact on Stock Prices.” *The Journal of Finance.* 54(2): 581-622.

Wooldridge, Jeffrey M. 2016. *Introductory Econometrics: A Modern Approach*. Mason: Cengage Learning.

Wu, Chen-Hui, Wu Chin-Shun and Victor Liu. 2009. “The conservatism bias in an emerging stock market: Evidence from Taiwan.” *Pacific-Basin Finance Journal*. 17(4). 494-505.

Footnotes

[1] Krugman, Paul. 1997. "Seven Habits of Highly Defective Investors." *Forbes*. December 29. https://archive.fortune.com/magazines/fortune/fortune_archive/1997/12/29/235886/index.htm

[2] To re-base, we used a rolling function to bound each sentiment indicator value between 0 and 1:

$$\frac{X_t - \min_{0,t}(X)}{\max_{0,t}(X) - \min_{0,t}(X)} \tag{4.1}$$

Where X is the value of the indicator.

I.e., the ratio gives the location of the current value within the total range. If the ratio gave a value of 0.5 then the current value is halfway between the historical minimum and maximum.

[3] Each equity contribution is calculated as:

$$\frac{\textit{Basic Earnings Per Share} \cdot \textit{Respective Number of Shares in Index}}{\textit{Index Divisor} \cdot \textit{Coverage Factor}} \tag{5.1}$$

Where the coverage factor is defined as:

$$1 - \frac{\textit{Unused Market Capitalisation}}{\textit{Index Market Capitalisation}} \tag{5.2}$$

Where unused market capitalisation is the sum of securities that do not have an equity contribution to the index value.

The index divisor is simply a value chosen at index inception used for standardisation in order to compute the nominal value of a price-weighted index. The number can be altered over time and ensures that corporate events, such as stock splits and buybacks, do not alter the underlying index.

[4] Generalised VAR model information:

$G_0 = (n \times 1)\ vector\ of\ constants$ $\qquad E[e_t] = 0$

$G_j = (n \times n)\ matrix\ of\ coefficients$ $\qquad E[e_t e_\tau'] = \begin{cases} \Omega, if\ t = \tau \\ 0\ otherwise \end{cases}$

$e_t = (n \times 1)\ vector\ of\ white\ noise\ innovations$

The innovations should be uncorrelated with their lagged values, i.e., no serial autocorrelation.

[5] To explain this concept, let us consider an $AR(1)$ model:

$$a_t = \phi a_{t-1} + \varepsilon_t \qquad (6.1)$$

$$(E(a_t) = \phi E(a_{t-1}) = \phi^2 E(a_{t-2}) = \cdots = \phi^t a_0, \qquad as\ E(\varepsilon_t) = 0)$$

The variance of a_t can be written as follows:

$$var(a_t) = \sigma^2\left[\phi^0 + \phi^2 + \phi^4 + \cdots + \phi^{2(t-1)}\right] \qquad (6.2)$$

As,

$$a_t = \phi a_{t-1} + \varepsilon_t = \phi^t a_0 + \textstyle\sum_{k=0}^{t-1} \phi^k \varepsilon_{t-k} \qquad (6.3)$$

$\phi^t a_0$ is a constant and therefore has a variance of zero:

$$var(\varepsilon) = \sigma^2 \qquad (6.4)$$

The ϕ is squared as we are looking at variance.

Now, let's look at three scenarios: $|\phi| < 1,\ \ |\phi| = 1\ and\ |\phi| > 1.$

When $|\phi| < 1$:

$$E(a_t) \to 0\ as\ t \to \infty \qquad (7.1)$$

$$var(a_t) \to \frac{\sigma^2}{1-\phi^2} as\ t \to \infty * \qquad (7.2)$$

Therefore, in this case, the time series is stationary, the expected value does not depend on the time period and the variance also does not depend on the time period.

*This is possible as $\left[\phi^0 + \phi^2 + \phi^4 + \cdots + \phi^{2(t-1)}\right]$ is a geometric series.

At each stage we are multiplying by the common ratio ϕ^2; as $t \to \infty$, this becomes an infinite geometric series, and we can sum it up as the common ratio ϕ^2 is less than one in absolute value.

In an infinite geometric series,

$$\sum_{k=0}^{\infty} ar^k = \frac{a}{1-r}, when\ |r| < 1 \tag{7.3}$$

Proof:

$$\sum_{k=0}^{\infty} ar^k = S_\infty = ar^0 + ar^1 + ar^2 + \cdots \tag{7.4}$$

$$r \cdot S_\infty = ar^1 + ar^2 + ar^3 + \cdots \tag{7.5}$$

$$S_\infty - r \cdot S_\infty = ar^0 = a \tag{7.6}$$

$$S_\infty(1-r) = a \tag{7.7}$$

$$S_\infty = \frac{a}{(1-r)} \tag{7.8}$$

We can therefore go from equation 6.2 to equation 7.2, where $a = \sigma^2$ and $r = \phi^2$.

When $|\phi| > 1$:

$$E(a_t) \to |\infty|\ as\ t \to \infty \tag{8.1}$$

Thus, the time series is non-stationary, as this is a direct violation of our stationarity conditions.

When $|\phi| = 1$:

$$E(a_t) = 0 \tag{9.1}$$

This does not violate stationarity, however, if we look at the variance:

$$var(a_t) = t\sigma^2 \tag{9.2}$$

$$var(a_t) \to \infty\ as\ t \to \infty \tag{9.3}$$

Therefore, the time series is non-stationary in this case.

[6] Augmented Dickey-Fuller (ADF) test. We have an autoregressive model as follows:

$$y_t = \mu + \phi_1 y_{t-1} + \varepsilon_t \tag{10.1}$$

$$H_0: \phi_1 = 1 \ (Unit\ root)$$

$$H_1: \phi_1 < 1 \ (No\ unit\ root)$$

$$y_t - y_{t-1} = \mu + (\phi_1 - 1) y_{t-1} + \varepsilon_t \tag{10.2}$$

$$\Delta y_t = \mu + \delta y_{t-1} + \varepsilon_t \tag{10.3}$$

$$H_0: \delta = 0$$

$$H_1: \delta < 0$$

$$t_{\widehat{\delta}} = \frac{\widehat{\delta}}{s.e.(\widehat{\delta})} \tag{10.4}$$

More generally,

$$y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \varepsilon_t \tag{10.5}$$

$$\Delta y_t = \mu + \delta y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t \tag{10.6}$$

$$H_0: \delta = 0$$

$$H_1: \delta < 0$$

$$t_{\widehat{\delta}} = \frac{\widehat{\delta}}{s.e.(\widehat{\delta})} \tag{10.7}$$

We then compare $t_{\widehat{\delta}}$ with the Dickey-Fuller distribution:

$$t_{\widehat{\delta}} < DF_{critical} \rightarrow Reject\ H_0\ (Stationary)$$

$$t_{\widehat{\delta}} > DF_{critical} \rightarrow Do\ not\ reject\ H_0\ (Non-Stationary)$$

For β_i:

$$t_{\widehat{\delta}} = \frac{\widehat{\beta}_i}{s.e.(\widehat{\beta}_i)} \tag{10.8}$$

[7] The test is run as follows:

First, assuming $i = 1$, we run the variable in question on all other explanatory variables:

$$X_1 = \alpha_0 + \alpha_2 X_2 + \alpha_3 X_3 + \cdots + \alpha_k X_k + e \quad (11.1)$$

Where α_0 is the constant and e is the error term.

Next, we calculate the Variance Inflation Factor for $\hat{\beta}_i$:

$$VIF_i = \frac{1}{1-R_i^2} \quad (11.2)$$

Where R_i^2 is the coefficient of determination* from equation 11.1.

*The coefficient of determination can be used to explain how much of the variation in a particular variable can be explained by the variation in other variable(s). This is also known as "goodness of fit" and is bounded between 0 and 1.

$$R^2 = 1 - \frac{RSS}{TSS} \quad (11.3)$$

R^2: Coefficient of Determination

RSS: Sum of Squares of Residuals

TSS: Total Sum of Squares

[8] For example, the SPY Return equation to estimate would take the form:

$$SPY\_Return_t = \alpha_0 + \beta_1 \cdot SPY\_Return_{t-1} + \beta_2 \cdot SPY\_Return_{t-2} + \beta_3 \cdot Twitter\_Sentiment_{t-1} + \beta_4 \cdot Twitter\_Sentiment_{t-2} + \beta_5 \cdot ASI_{t-1} + \beta_6 \cdot ASI_{t-2} + \beta_7 \cdot FAANG\_\%_{t-1} + \beta_8 \cdot FAANG\_\%_{t-2} + \beta_9 \cdot MTUM_{t-1} + \beta_{10} \cdot MTUM_{t-2} + \beta_{11} \cdot Analyst_{t-1} + \beta_{12} \cdot Analyst_{t-2} + \beta_{13} \cdot SPY\_vol_{t-1} + \beta_{14} \cdot SPY\_vol_{t-2} + \beta_{15} \cdot SPY\_turnover_{t-1} + \beta_{16} \cdot SPY\_turnover_{t-2} + \beta_{17} \cdot EPS\_growth_{t-1} + \beta_{18} \cdot EPS\_growth_{t-2} + \beta_{19} \cdot Sales\_growth_{t-1} + \beta_{20} \cdot Sales\_growth_{t-2} + \beta_{21} \cdot Tweet\_growth_{t-1} + \beta_{22} \cdot Tweet\_growth_{t-2} + \beta_{23} \cdot Bid\_Ask_{t-1} + \beta_{24} \cdot Bid\_Ask_{t-2} + \varepsilon_t \quad (12.1)$$

[9] Let Y_{t-1} be a matrix containing all the information available up to time t (before realizations of e_t are known):

$$Y_{t-1} = (y_{t-1}, y_{t-2}, \dots, y_{t-T}) \tag{13.1}$$

Then:

$$E[y_t|Y_{t-1}] = \hat{G}_0 + \hat{G}_1 y_{t-1} + \hat{G}_2 y_{t-2} + \cdots + \hat{G}_p y_{t-p} \tag{13.2}$$

[10] Note: the first forecasted day is the 101st day in the input data. Therefore, the first forecasted day uses the data from the first 100 days in order to calculate the coefficients and lags for the forecast. For the 102nd day, the first 101 days are utilised, and so on.

[11] The Charles Schwab Corporation. 2020. "The Rise of the Investor Generation." *Charles Schwab*. https://www.aboutschwab.com/generation-investor-study-2021. Accessed on 12th May 2021.

Tables

Table 1.1: *Range of Sentiment Indicators Incorporated into the Aggregated Sentiment Index (ASI)*

Source	**Behavioural Indicators**	**Technical Indicators**	Source
American Association of Individual Investors	**AAII Bull Ratio**	**AAII Asset Allocation – Stocks**	American Association of Individual Investors
University of Michigan	**UoM Consumer Confidence**	**AAII Asset Allocation – Bonds**	American Association of Individual Investors
The Conference Board	**Conference Board Consumer Confidence**	**Citigroup Macro Risk Index**	Citi
Langer Research Associates	**Langer Consumer Comfort Index**	**Bloomberg US Economic Surprise Index**	Bloomberg
National Federation of Independent Business	**NFIB Small Business Optimism**	**Citi Economic Surprise Index – US**	Citi
Federal Reserve Bank of San Francisco	**Daily News Sentiment Index**	**NAAIM Exposure Index**	National Association of Active Investment Managers
IHS Markit Ltd	**U.S. Markit Composite PMI**	**AAII Asset Allocation – Cash**	American Association of Individual Investors
		University of Michigan – Stocks	University of Michigan
		CSFB Fear Barometer	Credit Suisse
		CBOE Equity Put / Call Ratio	CBOE
		Westpac Risk Aversion Index	Westpac
		UBS G10 Carry Risk Index Plus	UBS

Table 2.1: *Augmented Dickey-Fuller Test Results for Stage One*

Variable	Lags Chosen	Test Statistic	Critical Value (1%)	P-Value	Result
SPY Return	8	-6.2492	-3.444	0.00	Stationary
Twitter Sentiment	3	-7.3418	-3.444	0.00	Stationary
Aggregated Sentiment Measure	0	-23.5093	-3.443	0.00	Stationary
FAANG / S&P 500 %	4	-4.8342	-3.444	0.00	Stationary
Momentum Excess Return	0	-21.8159	-3.443	0.00	Stationary
Analyst Targets	2	-3.6806	-3.443	0.00	Stationary
SPY Volatility	1	-8.8494	-3.443	0.00	Stationary
SPY Turnover	1	-8.7354	-3.443	0.00	Stationary
S&P 500 EPS Growth	1	-3.7338	-3.443	0.00	Stationary
S&P 500 Sales Growth	0	-3.9624	-3.443	0.00	Stationary
Tweet Volume Growth	2	-6.0172	-3.443	0.00	Stationary
VOO Bid-Ask Spread	0	-31.1862	-3.443	0.00	Stationary

Note: Null Hypothesis: Data has unit root. Non-Stationary.

Table 2.2: *Augmented Dickey-Fuller Test Results for Stage Two*

Variable	Lags Chosen	Test Statistic	Critical Value (1%)	P-Value	Result
SPY Return	8	-13.2267	-3.434	0.00	Stationary
Twitter Sentiment	3	-11.2799	-3.434	0.00	Stationary
Momentum Excess Return	7	-17.4741	-3.434	0.00	Stationary
SPY Volatility	22	-11.9714	-3.434	0.00	Stationary
VOO Bid-Ask Spread	2	-32.5108	-3.434	0.00	Stationary

Note: Null Hypothesis: Data has unit root. Non-Stationary.

Table 2.3: *Augmented Dickey-Fuller Test Results for Stage Three*

Variable	Lags Chosen	Test Statistic	Critical Value (1%)	P-Value	Result
UX1 VIX Futures Return	0	-43.1899	-3.434	0.00	Stationary
Twitter Sentiment	3	-11.2799	-3.434	0.00	Stationary
Momentum Excess Return	7	-17.4741	-3.434	0.00	Stationary
SPY Volatility	22	-11.9714	-3.434	0.00	Stationary
VOO Bid-Ask Spread	2	-32.5108	-3.434	0.00	Stationary

Note: Null Hypothesis: Data has unit root. Non-Stationary.

Table 3.1: *Variance Inflation Factor Values for All Variables in Stage One*

Variable	VIF
SPY Return	1.60
Twitter Sentiment	1.43
Aggregated Sentiment Measure (ASI)	1.29
FAANG / S&P 500 %	1.16
Momentum Excess Return	1.07
Analyst Targets	1.42
SPY Volatility	1.68
SPY Turnover	1.90
S&P 500 EPS Growth	1.50
S&P 500 Sales Growth	1.16
Tweet Volume Growth	2.03
VOO Bid-Ask Spread	1.20

Table 3.2: *Variance Inflation Factor Values for Condensed Variables in Stage Two*

Variable	VIF
SPY Return	1.464
Twitter Sentiment	1.136
Momentum Excess Return	1.259
SPY Volatility	1.023
VOO Bid-Ask Spread	1.064

Table 3.3: *Variance Inflation Factor Values for Condensed Variables in Stage Three*

Variable	VIF
UX1 VIX Futures Return	1.274
Twitter Sentiment	1.110
Momentum Excess Return	1.072
SPY Volatility	1.092
VOO Bid-Ask Spread	1.042

Table 4.1: *Order of Lag Selection for Stage One (* indicates minimised lag)*

Lag	AIC	BIC	FPE	HQIC
0	-100.7	-100.6	1.86E-44	-100.7
1	-107.8	-105.4	1.53E-47	-107.3*
2	-108.0	-106.5*	1.29E-47	-107.0
3	-108.0	-104.2	1.25E-47	-106.5
4	-108.1	-103.1	1.13e-47*	-106.1
5	-108.1	-101.8	1.14E-47	-105.6
6	-108.0	-100.5	1.26E-47	-105.1
7	-108.1*	-99.36	1.16E-47	-104.7
8	-108.0	-98.05	1.28E-47	-104.1
9	-108.0	-96.76	1.38E-47	-103.6
10	-108.0	-95.49	1.47E-47	-103.1
11	-107.9	-94.24	1.57E-47	-102.6
12	-108.0	-93.07	1.54E-47	-102.1

Table 5.1: *Durbin-Watson Statistic Values for All Variables in Stage One*

Variable	D-W Statistic
SPY Return	2.04
Twitter Sentiment	2.03
Aggregated Sentiment Measure (ASI)	2.03
FAANG / S&P 500 %	2.04
Momentum Excess Return	2.00
Analyst Targets	2.13
SPY Volatility	2.02
SPY Turnover	2.01
S&P 500 EPS Growth	1.99
S&P 500 Sales Growth	1.98
Tweet Volume Growth	2.08
VOO Bid-Ask Spread	2.07

Table 5.2: *Durbin-Watson Statistic Values for Condensed Variables in Stage Two*

Variable	D-W Statistic
SPY Return	2.01
Twitter Sentiment	2.07
Momentum Excess Return	2
SPY Volatility	2.04
VOO Bid-Ask Spread	2.06

Table 5.3: *Durbin-Watson Statistic Values for Condensed Variables in Stage Three*

Variable	D-W Statistic
UX1 VIX Futures Return	2
Twitter Sentiment	2.08
Momentum Excess Return	2
SPY Volatility	2.02
VOO Bid-Ask Spread	2.05

Table 6.1: *Regression Results for Daily Percentage SPY Return (spy_return) in Stage One*

Constant / Variable	Coefficient	Standard Error	T-Statistic	P-Value
Constant	0.001238	0.000851	1.455	0.146
L1.spy_return	-0.345257	0.057905	-5.962	0.000
L1.bull_minus_bear	0.003403	0.005330	0.638	0.523
L1.sentiment_measure	-0.097962	0.030319	-3.231	0.001
L1.faang_sap	2.627923	1.302764	2.017	0.044
L1.mom-mkt	0.245035	0.133790	1.831	0.067
L1.analyst_target	-2.288774	0.837924	-2.731	0.006
L1.spy_vol	0.050684	0.054162	0.936	0.349
L1.spy_turnover	-0.003120	0.002598	-1.201	0.230
L1.eps_growth	1.158057	0.765292	1.513	0.130
L1.spx_sales_growth	-3.264586	5.650391	-0.578	0.563
L1.tweet_volume_growth	-0.000034	0.000032	-1.088	0.277
L1.bid_ask_spread	0.311672	0.113919	2.736	0.006
L2.spy_return	0.144340	0.057944	2.491	0.013
L2.bull_minus_bear	0.001881	0.005226	0.360	0.719
L2.sentiment_measure	0.013123	0.029062	0.452	0.652
L2.faang_sap	0.191732	1.285116	0.149	0.881
L2.mom-mkt	0.058855	0.134997	0.436	0.663
L2.analyst_target	1.163016	0.834685	1.393	0.164
L2.spy_vol	-0.027590	0.053180	-0.519	0.604
L2.spy_turnover	-0.000467	0.002592	-0.180	0.857
L2.eps_growth	-1.563213	0.751321	-2.081	0.037
L2.spx_sales_growth	4.876679	5.689830	0.857	0.391
L2.tweet_volume_growth	-0.000039	0.000032	-1.239	0.216
L2.bid_ask_spread	0.202565	0.108840	1.861	0.063

Table 6.2: *Regression Results for SPY Turnover Proximity to Annual Average (spy_turnover) in Stage One*

Constant / Variable	Coefficient	Standard Error	T-Statistic	P-Value
Constant	0.010832	0.018699	0.579	0.562
L1.spy_return	-0.309896	1.272971	-0.243	0.808
L1.bull_minus_bear	0.180938	0.117175	1.544	0.123
L1.sentiment_measure	1.003793	0.666534	1.506	0.132
L1.faang_sap	-5.892519	28.639583	-0.206	0.837
L1.mom-mkt	-1.683070	2.941206	-0.572	0.567
L1.analyst_target	40.906019	18.420666	2.221	0.026
L1.spy_vol	0.722397	1.190690	0.607	0.544
L1.spy_turnover	0.407959	0.057120	7.142	0.000
L1.eps_growth	-5.153869	16.823948	-0.306	0.759
L1.spx_sales_growth	-204.128574	124.216514	-1.643	0.100
L1.tweet_volume_growth	0.001526	0.000695	2.194	0.028
L1.bid_ask_spread	3.028830	2.504358	1.209	0.227
L2.spy_return	-0.064946	1.273832	-0.051	0.959
L2.bull_minus_bear	-0.159333	0.114892	-1.387	0.166
L2.sentiment_measure	-0.121214	0.638898	-0.190	0.850
L2.faang_sap	7.582154	28.251612	0.268	0.788
L2.mom-mkt	-3.208174	2.967728	-1.081	0.280
L2.analyst_target	-23.096527	18.349459	-1.259	0.208
L2.spy_vol	-0.808938	1.169091	-0.692	0.489
L2.spy_turnover	0.146360	0.056974	2.569	0.010
L2.eps_growth	11.152552	16.516822	0.675	0.500
L2.spx_sales_growth	167.647360	125.083530	1.340	0.180
L2.tweet_volume_growth	-0.000057	0.000699	-0.081	0.935
L2.bid_ask_spread	0.825696	2.392699	0.345	0.730

Table 6.3: *Regression Results for Daily Change in Average Bid Ask Spread (bid_ask_spread) in Stage One*

Constant / Variable	**Coefficient**	**Standard Error**	**T-Statistic**	**P-Value**
Constant	-0.000036	0.000379	-0.094	0.925
L1.spy_return	-0.076167	0.025779	-2.955	0.003
L1.bull_minus_bear	0.002986	0.002373	1.258	0.208
L1.sentiment_measure	0.008089	0.013498	0.599	0.549
L1.faang_sap	-0.043271	0.579989	-0.075	0.941
L1.mom-mkt	-0.164737	0.059563	-2.766	0.006
L1.analyst_target	0.456590	0.373043	1.224	0.221
L1.spy_vol	0.047478	0.024113	1.969	0.049
L1.spy_turnover	0.000505	0.001157	0.436	0.663
L1.eps_growth	0.103400	0.340707	0.303	0.762
L1.spx_sales_growth	1.872650	2.515548	0.744	0.457
L1.tweet_volume_growth	-0.000003	0.000014	-0.248	0.804
L1.bid_ask_spread	-0.450189	0.050717	-8.877	0.000
L2.spy_return	-0.029426	0.025797	-1.141	0.254
L2.bull_minus_bear	-0.002400	0.002327	-1.032	0.302
L2.sentiment_measure	0.004744	0.012939	0.367	0.714
L2.faang_sap	0.352752	0.572132	0.617	0.538
L2.mom-mkt	-0.157557	0.060100	-2.622	0.009
L2.analyst_target	0.307694	0.371601	0.828	0.408
L2.spy_vol	-0.011385	0.023676	-0.481	0.631
L2.spy_turnover	-0.000414	0.001154	-0.358	0.720
L2.eps_growth	-0.179060	0.334487	-0.535	0.592
L2.spx_sales_growth	-2.208884	2.533106	-0.872	0.383
L2.tweet_volume_growth	0.000027	0.000014	1.889	0.059
L2.bid_ask_spread	-0.164997	0.048455	-3.405	0.001

Table 6.4: *Regression Results for Daily Percentage Change in SPY Volatility (spy_vol) in Stage One*

Constant / Variable	Coefficient	Standard Error	T-Statistic	P-Value
Constant	0.000482	0.000774	0.623	0.533
L1.spy_return	-0.218329	0.052691	-4.144	0.000
L1.bull_minus_bear	0.004811	0.004850	0.992	0.321
L1.sentiment_measure	-0.003295	0.027589	-0.119	0.905
L1.faang_sap	-1.307623	1.185450	-1.103	0.270
L1.mom-mkt	-0.292288	0.121742	-2.401	0.016
L1.analyst_target	-1.015889	0.762469	-1.332	0.183
L1.spy_vol	0.098084	0.049285	1.990	0.047
L1.spy_turnover	-0.001263	0.002364	-0.534	0.593
L1.eps_growth	1.098264	0.696377	1.577	0.115
L1.spx_sales_growth	-5.960475	5.141572	-1.159	0.246
L1.tweet_volume_growth	0.000057	0.000029	1.987	0.047
L1.bid_ask_spread	0.204778	0.103660	1.975	0.048
L2.spy_return	-0.123982	0.052726	-2.351	0.019
L2.bull_minus_bear	-0.002662	0.004756	-0.560	0.576
L2.sentiment_measure	-0.022767	0.026445	-0.861	0.389
L2.faang_sap	-0.931823	1.169391	-0.797	0.426
L2.mom-mkt	0.119415	0.122840	0.972	0.331
L2.analyst_target	0.215998	0.759521	0.284	0.776
L2.spy_vol	0.192269	0.048391	3.973	0.000
L2.spy_turnover	0.002243	0.002358	0.951	0.342
L2.eps_growth	0.035626	0.683665	0.052	0.958
L2.spx_sales_growth	4.706301	5.177460	0.909	0.363
L2.tweet_volume_growth	0.000017	0.000029	0.590	0.555
L2.bid_ask_spread	0.013842	0.099039	0.140	0.889

Table 6.5: *Regression Results for Daily Change in Average Monthly Tweet Volume Growth(tweet_volume_growth) in Stage One*

Constant / Variable	Coefficient	Standard Error	T-Statistic	P-Value
Constant	-0.392365	1.514259	-0.259	0.796
L1.spy_return	-75.701428	103.087376	-0.734	0.463
L1.bull_minus_bear	-1.691501	9.489006	-0.178	0.859
L1.sentiment_measure	-53.938237	53.977076	-0.999	0.318
L1.faang_sap	-3374.867232	2319.282303	-1.455	0.146
L1.mom-mkt	-135.060572	238.183881	-0.567	0.571
L1.analyst_target	2256.248680	1491.736979	1.512	0.130
L1.spy_vol	28.868394	96.424094	0.299	0.765
L1.spy_turnover	8.371584	4.625662	1.810	0.070
L1.eps_growth	-459.760494	1362.431962	-0.337	0.736
L1.spx_sales_growth	-1393.831535	10059.265121	-0.139	0.890
L1.tweet_volume_growth	0.316265	0.056314	5.616	0.000
L1.bid_ask_spread	123.593584	202.807156	0.609	0.542
L2.spy_return	-75.268231	103.157077	-0.730	0.466
L2.bull_minus_bear	-15.942654	9.304153	-1.713	0.087
L2.sentiment_measure	-101.808551	51.739083	-1.968	0.049
L2.faang_sap	-1599.376195	2287.863734	-0.699	0.485
L2.mom-mkt	261.535355	240.331713	1.088	0.276
L2.analyst_target	264.681343	1485.970498	0.178	0.859
L2.spy_vol	46.874548	94.675012	0.495	0.621
L2.spy_turnover	5.156212	4.613820	1.118	0.264
L2.eps_growth	391.078871	1337.560448	0.292	0.770
L2.spx_sales_growth	2870.235285	10129.477552	0.283	0.777
L2.tweet_volume_growth	0.211902	0.056644	3.741	0.000
L2.bid_ask_spread	15.847702	193.764880	0.082	0.935

Table 6.6: *Regression Results for the Daily Percentage Change in Analyst Targets Relative to the S&P 500 (analyst_target) in Stage One*

Constant / Variable	Coefficient	Standard Error	T-Statistic	P-Value
Constant	0.000054	0.000046	1.185	0.236
L1.spy_return	0.002136	0.003120	0.685	0.494
L1.bull_minus_bear	0.000432	0.000287	1.505	0.132
L1.sentiment_measure	0.001753	0.001634	1.073	0.283
L1.faang_sap	0.175230	0.070193	2.496	0.013
L1.mom-mkt	0.006570	0.007209	0.911	0.362
L1.analyst_target	0.532883	0.045147	11.803	0.000
L1.spy_vol	-0.003733	0.002918	-1.279	0.201
L1.spy_turnover	-0.000012	0.000140	-0.085	0.933
L1.eps_growth	0.143267	0.041234	3.474	0.001
L1.spx_sales_growth	-1.074952	0.304444	-3.531	0.000
L1.tweet_volume_growth	-0.000001	0.000002	-0.482	0.630
L1.bid_ask_spread	0.001703	0.006138	0.278	0.781
L2.spy_return	0.003898	0.003122	1.248	0.212
L2.bull_minus_bear	0.000433	0.000282	1.537	0.124
L2.sentiment_measure	0.001383	0.001566	0.883	0.377
L2.faang_sap	0.035618	0.069242	0.514	0.607
L2.mom-mkt	-0.006849	0.007274	-0.942	0.346
L2.analyst_target	0.254242	0.044973	5.653	0.000
L2.spy_vol	-0.003702	0.002865	-1.292	0.196
L2.spy_turnover	0.000007	0.000140	0.053	0.958
L2.eps_growth	-0.145941	0.040481	-3.605	0.000
L2.spx_sales_growth	1.379599	0.306569	4.500	0.000
L2.tweet_volume_growth	0.000001	0.000002	0.825	0.409
L2.bid_ask_spread	0.000904	0.005864	0.154	0.878

Table 7.1: *Individual Variable Forecasting Accuracies for Stage Two*

Variable	Prediction Accuracy	95% Significance
Twitter Sentiment	56.33%	Significant
Aggregated Sentiment Measure (ASI)	54.85%	Not Significant
FAANG / S&P 500 %	54.09%	Not Significant
Momentum Excess Return	55.83%	Significant
Analyst Targets	54.84%	Not Significant
SPY Volatility	55.34%	Significant
SPY Turnover	54.09%	Not Significant
S&P 500 EPS Growth	51.36%	Not Significant
S&P 500 Sales Growth	54.84%	Not Significant
Tweet Volume Growth	54.09%	Not Significant
VOO Bid-Ask Spread	55.36%	Significant

Note: Null Hypothesis: Mean prediction accuracy is equal to 50%

Table 8.1: Sharpe Ratio for each SPY / UX1 Strategy Combination

SPY Allocation (%)	UX1 Allocation (%)	Sharpe Ratio
100%	0%	0.495
95%	5%	0.688
90%	10%	0.804
85%	15%	0.848
80%	20%	0.841
75%	25%	0.809
70%	30%	0.767
65%	35%	0.722
60%	40%	0.678
55%	45%	0.635
50%	50%	0.594
45%	55%	0.554
40%	60%	0.516
35%	65%	0.479
30%	70%	0.444
25%	75%	0.409
20%	80%	0.376
15%	85%	0.343
10%	90%	0.311
5%	95%	0.280
0%	100%	0.2476

Table 8.2: *Monthly Returns for Strategies in Stage Three*

Month	SPY Return	SPY +5% Strategy Return	SPY +15% Strategy Return	SPY +25% Strategy Return
92013	-0.78%	-0.65%	-0.41%	-0.17%
102013	4.63%	3.89%	2.38%	0.85%
112013	2.96%	3.12%	3.44%	3.74%
122013	2.04%	2.18%	2.46%	2.72%
12014	-3.52%	-3.68%	-4.01%	-4.37%
22014	4.55%	4.92%	5.63%	6.31%
32014	0.39%	0.20%	-0.21%	-0.65%
42014	0.70%	0.44%	-0.08%	-0.61%
52014	2.32%	2.65%	3.31%	3.96%
62014	1.58%	1.11%	0.18%	-0.76%
72014	-1.34%	-1.69%	-2.40%	-3.13%
82014	3.95%	3.10%	1.39%	-0.33%
92014	-1.84%	-1.33%	-0.32%	0.67%
102014	2.36%	2.10%	1.46%	0.67%
112014	2.75%	2.88%	3.13%	3.36%
122014	-0.80%	-1.80%	-3.84%	-5.93%
12015	-2.96%	-1.47%	1.53%	4.53%
22015	5.62%	5.32%	4.71%	4.06%
32015	-2.01%	-0.84%	1.53%	3.93%
42015	0.98%	2.78%	6.44%	10.22%
52015	1.29%	2.74%	5.71%	8.75%
62015	-2.51%	-3.08%	-4.27%	-5.51%
72015	2.26%	3.57%	6.17%	8.75%
82015	-6.10%	-3.76%	0.97%	5.77%
92015	-3.06%	-3.78%	-5.30%	-6.90%
102015	8.51%	10.39%	14.22%	18.11%
112015	0.37%	2.17%	5.83%	9.56%
122015	-2.31%	-3.96%	-7.25%	-10.53%
12016	-4.98%	-3.45%	-0.40%	2.64%
22016	-0.08%	0.05%	0.28%	0.49%
32016	6.18%	5.50%	4.12%	2.73%
42016	0.39%	1.24%	2.92%	4.59%
52016	1.70%	2.99%	5.60%	8.23%
62016	-0.17%	1.86%	5.82%	9.65%
72016	3.65%	3.92%	4.44%	4.95%
82016	0.12%	0.55%	1.40%	2.25%
92016	-0.50%	-2.40%	-6.16%	-9.86%
102016	-1.73%	-2.96%	-5.39%	-7.77%
112016	3.68%	3.78%	3.96%	4.08%

122016	1.43%	2.31%	4.08%	5.85%
12017	1.79%	1.91%	2.15%	2.37%
22017	3.93%	3.70%	3.25%	2.78%
32017	-0.31%	-0.42%	-0.66%	-0.90%
42017	0.99%	0.54%	-0.38%	-1.33%
52017	1.41%	1.12%	0.52%	-0.13%
62017	0.15%	0.18%	0.24%	0.29%
72017	2.06%	2.59%	3.66%	4.74%
82017	0.29%	-1.24%	-4.34%	-7.50%
92017	1.51%	1.95%	2.83%	3.70%
102017	2.36%	2.68%	3.31%	3.94%
112017	3.06%	2.91%	2.62%	2.32%
122017	0.70%	1.38%	2.76%	4.15%
12018	5.64%	4.37%	1.87%	-0.60%
22018	-3.64%	2.88%	16.01%	29.20%
32018	-3.13%	-2.16%	-0.25%	1.64%
42018	0.52%	0.24%	-0.35%	-0.98%
52018	2.43%	1.77%	0.43%	-0.91%
62018	0.13%	-0.77%	-2.56%	-4.35%
72018	3.70%	3.28%	2.41%	1.52%
82018	3.19%	1.53%	-1.74%	-4.94%
92018	0.14%	-0.31%	-1.23%	-2.14%
102018	-6.91%	-7.73%	-9.43%	-11.19%
112018	1.85%	1.58%	1.00%	0.37%
122018	-9.33%	-8.00%	-5.30%	-2.60%
12019	8.01%	8.06%	8.12%	8.13%
22019	3.24%	3.85%	5.05%	6.27%
32019	1.36%	0.90%	-0.05%	-1.03%
42019	4.09%	2.97%	0.76%	-1.41%
52019	-6.38%	-6.24%	-6.06%	-5.98%
62019	6.44%	6.82%	7.58%	8.33%
72019	1.51%	1.19%	0.53%	-0.15%
82019	-1.67%	-1.26%	-0.56%	-0.04%
92019	1.48%	2.25%	3.79%	5.33%
102019	2.21%	1.96%	1.43%	0.83%
112019	3.62%	3.61%	3.59%	3.55%
122019	2.40%	3.12%	4.53%	5.91%
12020	-0.04%	-0.85%	-2.49%	-4.14%
22020	-7.92%	-6.95%	-5.07%	-3.26%
32020	-13.00%	-9.36%	-2.09%	5.09%
42020	12.70%	11.84%	10.07%	8.25%
Total	**71.55%**	**103.65%**	**174.64%**	**250.05%**

Table 8.3: *Annual Returns for Strategies in Stage Three*

Year	**SPY Return**	**SPY +5% Strategy Return**	**SPY +15% Strategy Return**	**SPY +25% Strategy Return**
2013	9.07%	8.75%	8.06%	7.29%
2014	11.29%	8.89%	3.85%	-1.44%
2015	-0.81%	9.41%	32.31%	58.73%
2016	9.64%	13.73%	21.70%	29.28%
2017	19.38%	18.62%	16.82%	14.68%
2018	-6.35%	-4.14%	-1.02%	0.44%
2019	28.79%	29.96%	31.75%	32.76%
2020	-9.75%	-6.48%	-0.24%	5.49%
Total	**71.55%**	**103.65%**	**174.64%**	**250.05%**
Positive Years	5	6	6	7
Negative Years	3	2	2	1
% Positive	63%	75%	75%	88%

Table 8.4: *Key Variable Comparison for Stage Three*

Variable	**SPY Return**	**SPY +5% Strategy Return**	**SPY +15% Strategy Return**	**SPY +25% Strategy Return**
Alpha	0.00%	-0.01%	0.00%	0.01%
Beta	1	0.989	0.625	0.303
Annualised Return	8.50%	11.34%	16.49%	20.84%
Annualised Volatility	17.15%	16.67%	19.45%	25.61%
Sharpe Ratio	0.495	0.680	0.848	0.814
Max Drawdown (Days)	417	294	504	608
Max Drawdown (Return)	-34.1%	-30.5%	-25%	-28%
Average Drawdown (Days)	66	39	81	145
Average Drawdown (Return)	-3.00%	-2.62%	-4.0%	-7.0%
Skewness	-0.677	-0.459	0.011	1.352
Kurtosis	20.31	15.25	16.23	39.30
Positive Months	54	54	51	48
Negative Months	26	26	29	32
% Positive (Months)	**68%**	**68%**	**64%**	**60%**

Figures

Figure 1.1: *Smoothed SPY Sentiment*

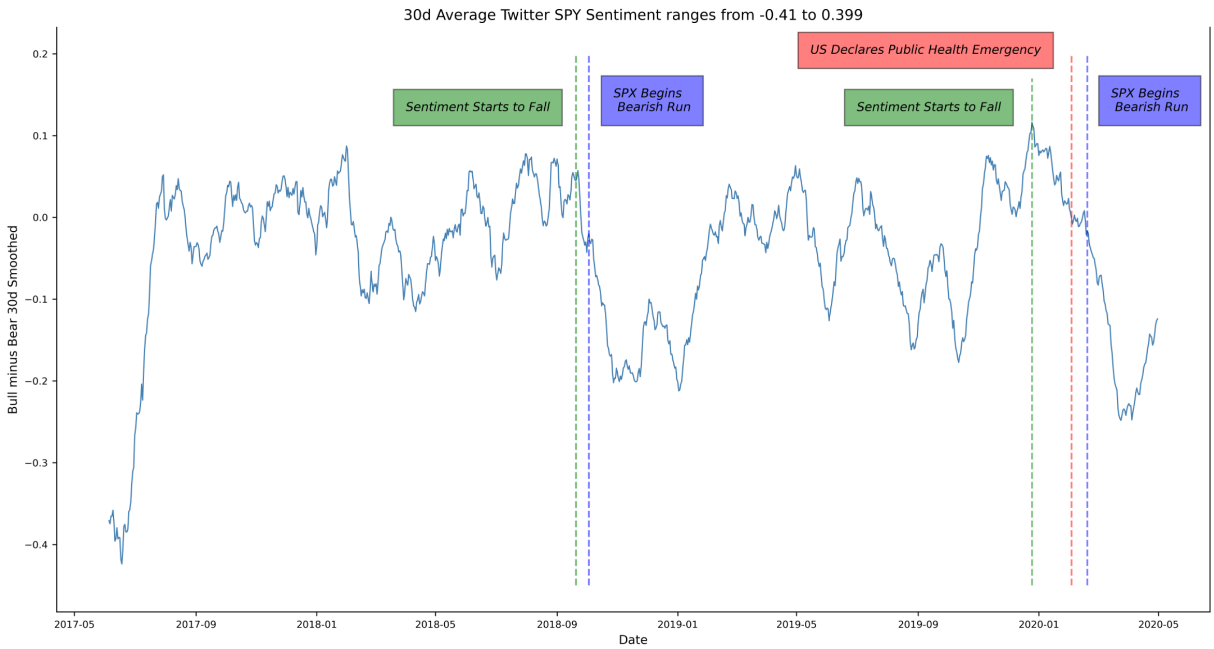

Figure 1.2: *SPY Level*

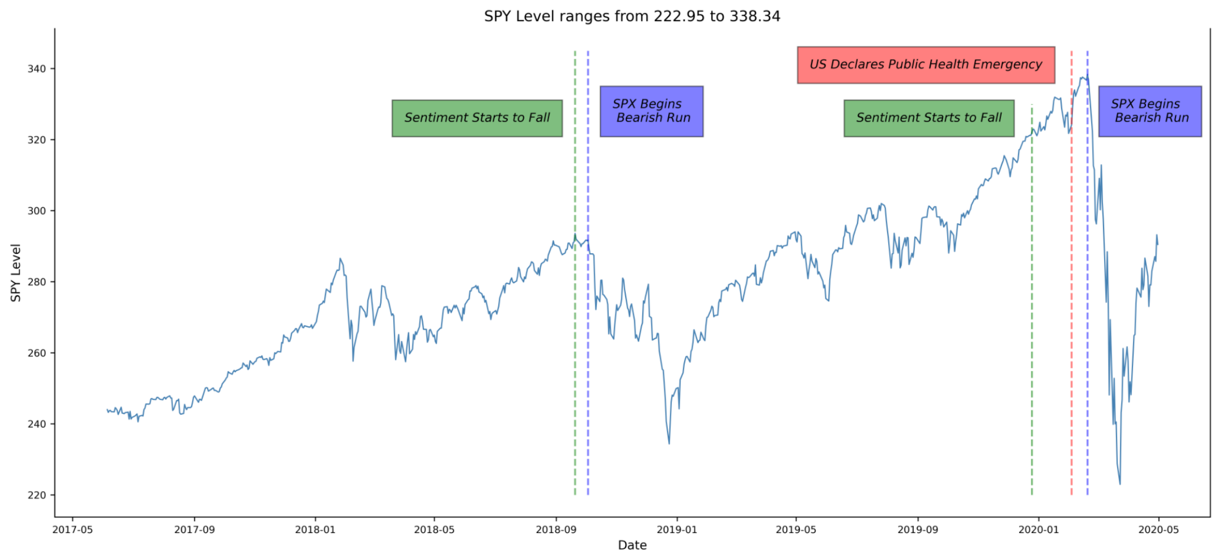

Figure 2.1: *Cross-Correlation Across All Variables in Stage One*

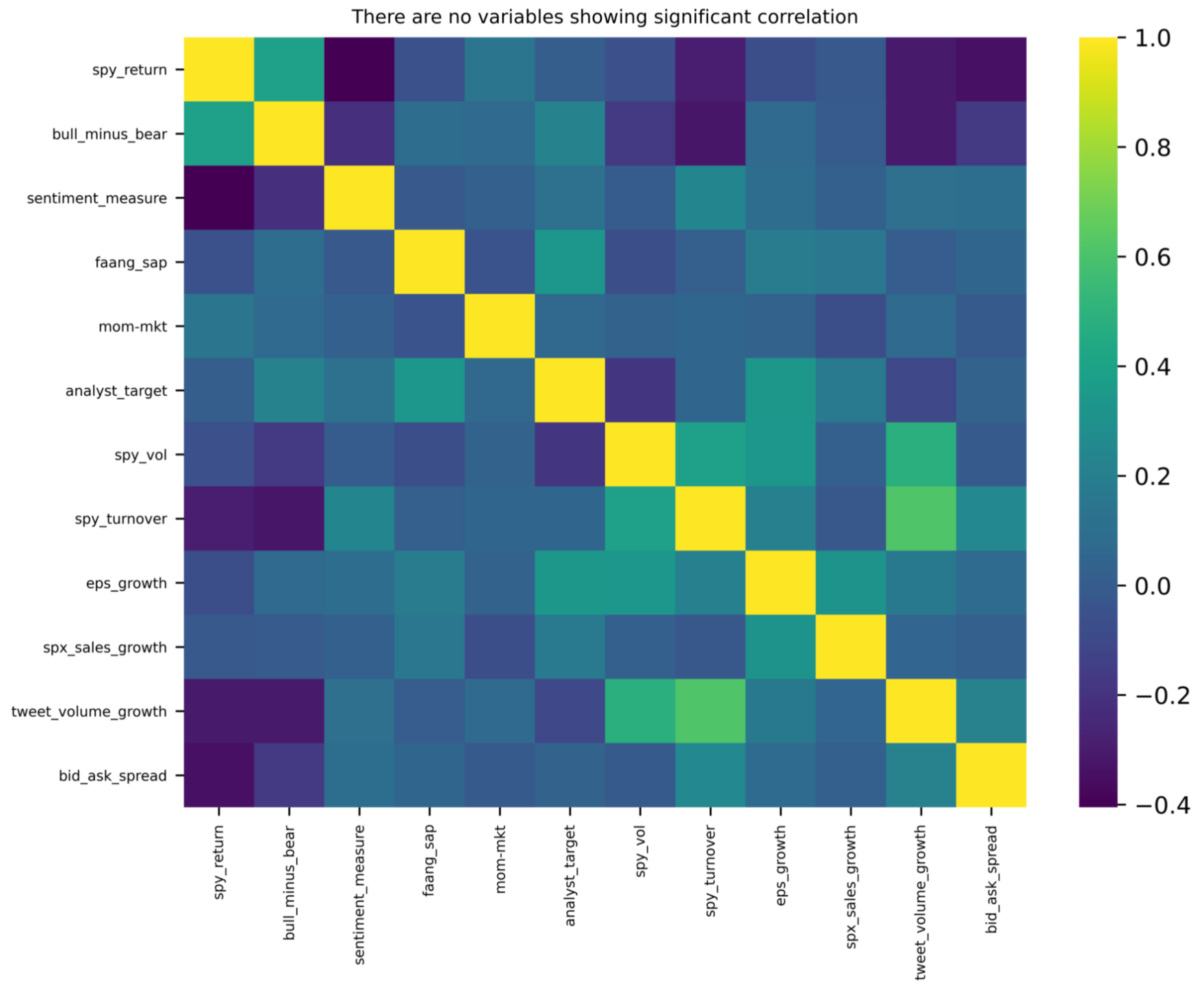

Figure 3.1: *Forecasted vs. Actual Daily SPY Return Values for Stage Two*

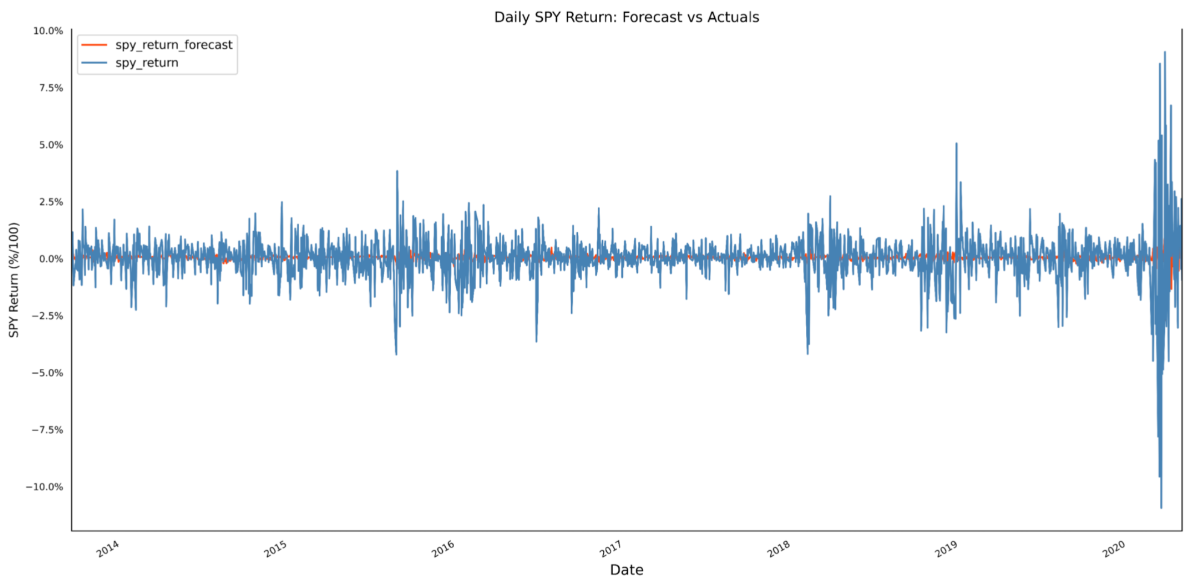

Figure 3.2: *First 100 Forecasted vs. Actual Daily SPY Return Values for Stage Two*

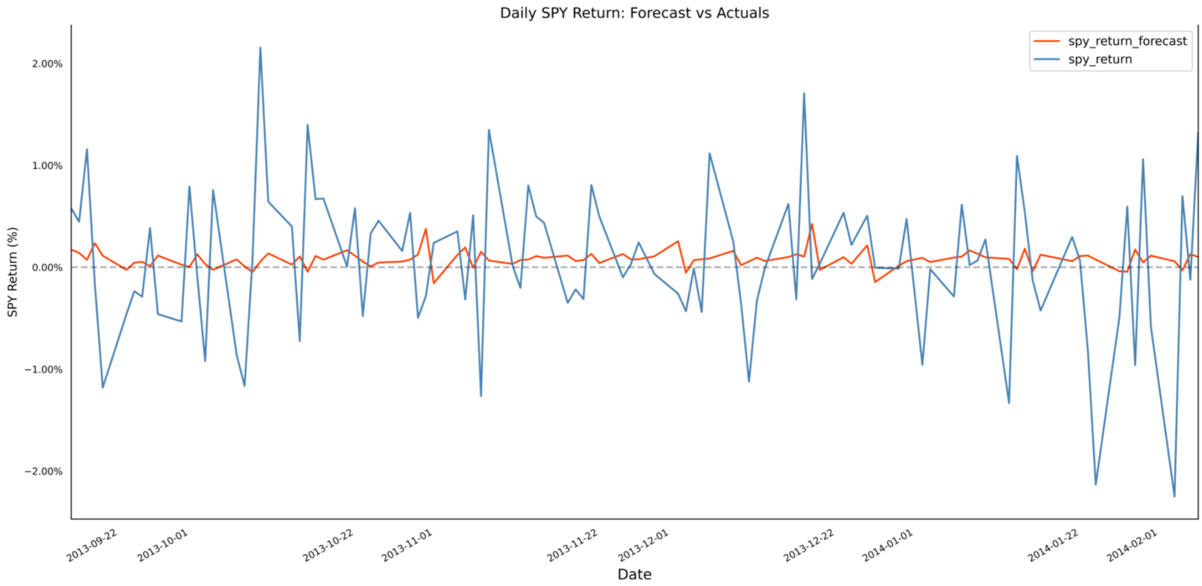

Figure 4.1: *Cumulative Return Comparison for Stage Three*

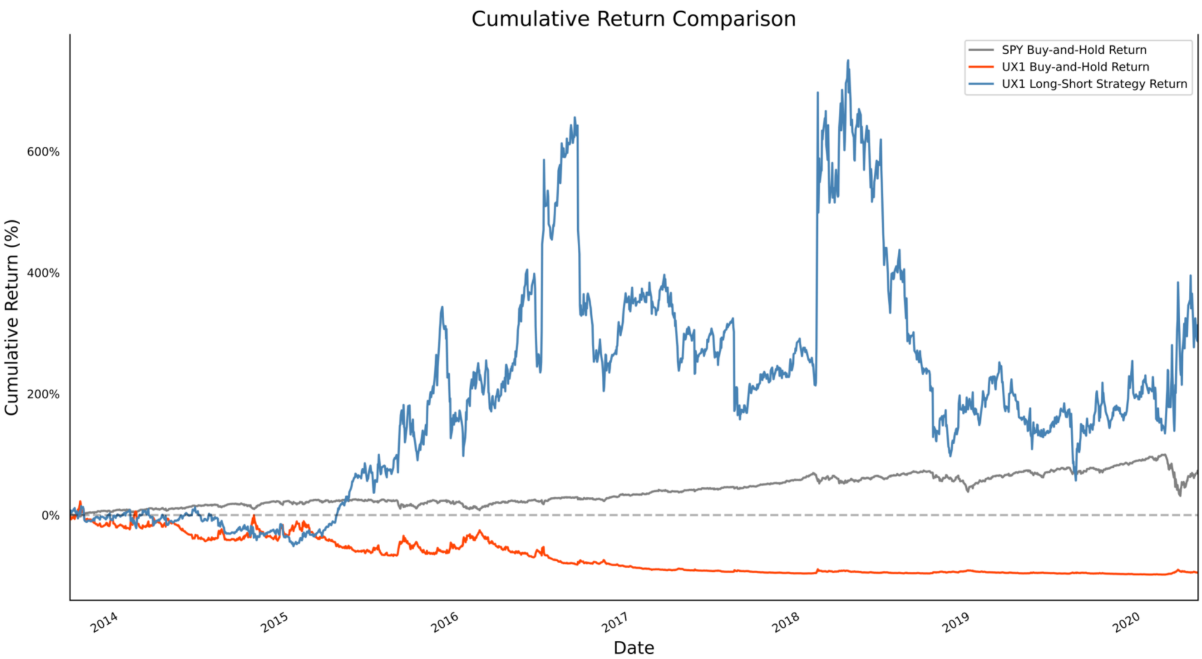

Figure 4.2: *Cumulative Return Comparison for Blended Combinations in Stage Three*