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.

DO SENTIMENT AND BEHAVIORAL HETEROGENEITY AMONG INVESTORS IMPACT ASSET PRICES IN FINANCIAL MARKETS?

LAURA NIKOLAY

Work project carried out under the supervision of:

Paulo M. M. Rodrigues

19/12/2023

Abstract

This work project adds to behavioral finance literature by examining the effect of behavioral investor sentiment and its degree of heterogeneity on asset prices. With several regression models it is differentiated between daily and monthly variables, as well as different sectors. The analysis underlines the relevance of incorporating investor sentiment into behavioral asset pricing models as it explains up to 20% of variation in market excess returns. The work project concludes that heterogeneity in investor sentiment has a negative effect on asset prices. Furthermore, investor sentiment has a bigger effect on the financial sector than on other analyzed sectors.

Keywords: Asset pricing, behavioral finance, investor sentiment, behavioral heterogeneity, US equities

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

1. Introduction

Behavioral finance is rising in popularity and especially since the 2008 financial crisis, behavioral models are replacing or complementing traditional financial theories. In the field of asset pricing, research insights are challenging the notion of efficient prices and behavioral finance provides explanations and revealing insights into the conduct of market participants. The link between market prices and investor sentiment has been a long debated subject in financial economics. This work project uses regression analysis to examine the effect of investor sentiment and furthermore, its degree of heterogeneity, on the US stock market. Thereby, sentiment is captured by commonly used variables such as the VIX (volatility index), the bullish percent index, and the put/call ratio as well as a sentiment index (Baker & Wurgler, 2022) and a tail risk index. Furthermore, a newly constructed heterogeneity index captures the degree of dispersion in investor sentiment. This work project contributes to academic research in the field of asset pricing by taking on a behavioural approach and investigating the following research question:

Do sentiment and behavioral heterogeneity among investors impact asset prices in financial markets?

The work project provides answers to this problem statement in the following format: First, a literature review synthesizes prior research, and identifies potential research gaps. Additionally, relevant definitions are provided. Second, the data which is used in the analysis is portrayed. Third, the methodology of the analysis is described in detail and the model equations are derived. Subsequently, the analysis of recent market data is performed, and the outcomes are discussed and placed within the academic context. Furthermore, it is analyzed, how the sentiment and heterogeneity model can complement traditional asset pricing factor-models. Then, the work project makes a reference to the limitations of the presented analysis as well as to suggestions for further research. Finally, a summary of the main findings and a conclusion are provided.

2. Literature Review

2.1 Relevance of Investor Sentiment in asset pricing

In the following, an overview of the current state of research in this field is given by synthesizing prior literature. Academic consensus prevails regarding the deficiencies of traditional asset pricing models that assume fully rational investors and a strong form of efficient markets (Barras 2019, Fama and French 2012). Shefrin and Statman (1994) started to build an alternative to the classical capital asset pricing model. The authors acknowledge that "noise traders" whose behavior diverges from rationality, have a significant effect on asset prices, and therefore, must be factored into models (Shefrin and Statman 1994). Thereupon, Shefrin (2008) develops a "behavioral stochastic discount factor-based" (SDF- based) asset pricing model that focuses on biases as a function of investor sentiment relative to fundamental asset prices. Here, sentiment is generally defined as "aggregate error in the market". Within the scope of this work project, sentiment is further defined as aggregated subjective beliefs, emotions, attitudes or feelings towards an asset or the market as a whole. In his analysis, in which Shefrin (2008) uses solely the dispersion of analysts' forecasts as a proxy for market sentiment, he concludes that sentiment causes a deviation from efficient prices.

2.2 Heterogeneity among analysts' expectations

More research, particularly on the effect of disagreement among analysts on asset prices shows that investors seem to earn a risk premium on assets towards which the market has a very insecure or heterogeneous sentiment (Carlin, Lingstaff and Matoba 2014). Carlin et al. (2014) confine themselves to the level of disagreement regarding prepayment speeds in the US mortgage-backed securities market and find that there is a positive risk premium for higher levels of disagreement. Cornea-Madeira and Madeira (2020) analyze intraday dynamics of behavioral heterogeneity in stock prices. They restrict their homogeneity analysis to a bidimensional differentiation between investors that assume market efficiency and investors with momentum beliefs. Cornea-Madeira and Madeira (2020) conclude that an increase in momentum traders leads to higher realizedvolatility . Doukas et al. (2004) investigate the effects of investor disagreement and find that there is a higher degree of analyst disagreement for value stocks than for growth stocks. Thereupon the authors argue that the return advantage of value strategies is a reward for the higher level of disagreement about the stocks' future growth (Doukas, Kim and Pantzalis 2004).

2.3 Sentiment and systematic risk

Furthermore, research documents a statistical relationship between sentiment and systematic risk (Barone-Adesi, Mancini and Hersh 2011). The authors find that systematic risk builds over time in response to variations in market sentiment. As sentiment changes direction, a reduction in optimism and confidence gives rise to a substantial increase in systematic risk and a sharp fall in prices. This finding underlines the relevance of incorporating investor sentiment into asset pricing models as it has a significant effect on price developments.

2.4 Sentiment and herding behavior

Babalos et al. (2013) show that investor sentiment and herding behavior are closely related. Herding behavior is a behavioral bias and relates to the tendency of investors to follow the crowds when markets fluctuate (Nguyen 2022). Therefore, herding behavior can further reinforce market sentiment in either direction. However, studies have found that the reinforcement effect of herding behavior on market sentiment tends to be stronger for pessimistic investors (Ren and Wu 2020), and for negative macroeconomic shocks (Huang and Wang 2017).

2.5 Low predictive power of conventional equity risk premium models

A study by Welch and Goyal (2008) on the empirical performance of equity premium prediction highlights the need for further research in this field. The authors analyze the predictive power of most commonly used variables that include the dividend price ratio and dividend yield, the earnings price ratio, dividend-earnings (payout) ratio, various interest rates and spreads, the inflation rates, and the book-to-market ratio. The main conclusion of their analysis is that the most prominent variables explored in the literature, such as equity risk premium predictors, have little predictive power and would not have helped investors to outperform the market (Welch and Goyal 2008). Therefore, more research must be done that includes other than the mentioned variables.

2.6 The Fama French 3-Factor Model (FFM3) and its extensions with other risk factors

The FFM3 model extends the traditional capital asset pricing model by adding two additional risk factors which are firm size and book-to-market ratio (Fama and French 1993). The model considers the finding that value and small-cap stocks tend to outperform when simply considering market risk. The model has been expanded by several other risk factors like momentum or reinvested earnings. Recently, researchers have also analyzed the effect of sentiment on the model and found that sentiment plays an important role in explaining the Fama French factors (Habibah, Bhayo and Shahid Iqbal 2021). The authors further note that only limited research has been done to develop an asset pricing factor model that incorporates investor sentiment.

2.7 Summary of the literature review and outlook on the analysis

The results of this literature review can be summarized as follows: prior studies have shown that more research is required in the field of behavioral asset pricing and that traditional models which neglect aspects of investor behavior are lacking predictive power. The literature further agrees that investor sentiment plays a significant role in determining asset prices and should therefore, be included in behavioral asset pricing models. However, there is discordance about how to measure market sentiment and which proxies to use. Furthermore, many prior studies focus only on a relatively short time span of up to 10 years. Besides, the focus of prior research lies on market sentiment alone whereas the degree of its heterogeneity among investors has attracted little attention by researchers. This work project will close these gaps in the literature by performing a comprehensive analysis of the effect of sentiment and the degree of investor heterogeneity on market prices. Thereby, up to seven proxies will be incorporated into several regressions over a timespan of 18 years with daily as well as monthly data.

3. Data[1]

3.1 Daily Data

The SPX daily closing prices (S&P 500) from 07.07. 2006 until 25.09.2023 in US$ were used to calculate the return data that served as proxy for market return. This time span amounts to 4310 observations per variable (see *Appendix 1* for a complete overview of the daily data). Subsequently, the daily risk-free rate was deducted to get a proxy for excess market return. As explanatory variables, first, a sentiment index (retrieved from Morgan Stanley) is used. It is a level indicator, with zero implying that there is neutral market sentiment. Second, the CBOE

[1] The complete data set as well as sources and description of the variables can also be found in the attached excel files.

volatility index (VIX) is included as regressor because it provides information about the market's expectations of near-term price changes of the S&P 500 and thereby, serves as an indicator for risk, or insecurity in the market. It is derived from prices of close-to expiry SPX index options (Bloomberg 2023). The CBOE put/call ratio provides additional information about market sentiment. It is the volume ratio of CBOE put options and CBOE call options. A high put/call ratio indicates a bearish sentiment as investors demand more put options than call options to hedge against falling market prices. Lastly, the CBOE S&P500 tail risk index returns are included. This index tracks the performance of a hypothetical risk-management strategy that consists of holding the S&P 500 portfolio and buying out-of-the-money SPX puts (CBOE Global Indices 2021). Thereby, the strategy hedges downside risk and outperforms in cases of black swan /left-tail events. The data analysis for this work project shows that the tail risk index returns are too highly correlated with market returns. This implies that index does not add significantly to the explanatory power of the model and undermines statistical significance of other variables. Therefore, the tail risk index has been excluded in the course of the analysis.

Appendix 2 displays the time-series plots for the described variables. The two most striking periods are the 2008-2009 financial crisis and the 2020 Covid-19 crisis. Here, spikes in the VIX come along with drastic falls in market excess returns. Furthermore, the lowest value of the sentiment index coincides with the most critical time of the Covid-19 crisis in 2020. It becomes also visible that the sentiment level is much more volatile than the VIX or market excess returns. The put/call ratio spikes in December 2022 which was the year in which Wall Street had the biggest annual percentage drop since 2008 (Reuters 2022). The market sentiment at the end of 2022 was quite bearish as investors were uncertain regarding the US economy's near-term outlook and recession fears (abc News 2022).

3.2 Monthly Data

The time span analyzed ranges from March 2004 until June 2022, corresponding to 220 observations per variable. Data was mainly retrieved from Bloomberg. Additional data sources were the NYU Department of Finance's website[2] and the website of Prof. Kenneth R. French[3]. As a proxy for market return, the excess return of the S&P 500 over the FED effective rate was used in the analysis. Several additional explanatory variables are added to the ones which were already described for the daily model. First, the equity market sentiment index from Baker & Wurgler (2022) was included. This index was initially constructed in 2006 and lastly updated in 2022 (Wurgler 2022). For the construction of the Baker and Wurgler (2006) sentiment index, the authors acknowledge the fact that there is not one perfect proxy for investor sentiment and therefore, consider several proxies that were suggested in previous work. Among those proxies are dividend premium, first-day returns of IPOs, or closed-end fund discounts. To uncouple those proxies from systematic risk and corresponding market movements, they have been orthogonalized to several macroeconomic conditions (Baker and Wurgler 2006). The (orthogonalized) sentiment index and all sentiment proxies can be found in the 'Sentiment B&W' sheet of the enclosed data file.

Furthermore, the bullish percent Index from the American Association of Individual Investors was included in this paper's analysis. This index is constructed based on survey results and provides insights into investor sentiment as well as its degree of dispersion.

The put/call ratio index reflects the monthly middle of the CBOE equity put option to call option volume ratio and provides additional information about investor's market outlooks. If investors are buying more puts than calls, it signals a bearish sentiment. A put/call ratio of one suggests an equal number of buyers for calls and puts. However, there is normally a higher demand for

[2] https://pages.stern.nyu.edu/~jwurgler/

[3] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Breakpoints

calls than for puts, therefore, a put/call ratio of 0.7 is considered average (Lamont and Stein 2004). In the sample data set, the average put/call ratio is 0.63.

A heterogeneity measure is newly constructed in this work project[4]. The index is built by extracting information about the degree of heterogeneity in investor sentiment from the put/call ratio as well as the bullish percent index. The computation of the heterogeneity index is as follows: The difference between one and the share of sentiment neutral investors (extracted from the bullish percent index) is added to the difference of one minus the absolute difference between 0.7 and the put/call ratio. The further away this ratio is from 0.7, the higher is the degree of heterogeneity. The first part of the equation allows for investor sentiment to go in either direction, bullish or bearish. The second part of the equation additionally considers the degree of dispersion among investor sentiment while also allowing the dispersion to go in either bullish or bearish direction. *Equation 1* displays the construction of the heterogeneity index for time *t*,

$$heterogeneity_t = (1 - monthly\ average\ neutral_t) + (1 - (|0.7 - put\ call\ ratio_t|)) \quad (1)$$

The resulting value provides insights into the degree of uncertainty and heterogeneity of investor opinions about the markets. A higher value of the heterogeneity index indicates a higher degree of behavioral heterogeneity among investors.

Appendix 3 displays the evolution of the described variables over time. Monthly market excess return and volatility (VIX) present the same pattern as for the daily data with the biggest dips representing the 2008/2009 financial crisis and the Covid-19 crisis. The largest drops in the Baker and Wurgler sentiment index also coincide with these two financial crises. However, it has bounced back strongly in the end of 2021 and 2022. The bullish percent index exhibits a

[4] See Data monthly excel file for complete data and computation of the index

high level of volatility with the biggest drop at the beginning with 2022. This time coincides with the Russian invasion of Ukraine which explains the low portion of investors indicating that they have a bullish market outlook. The put/call ratio is also the highest during the 2008 crisis and has another peak at the beginning of the Covid-19 crisis. However, corresponding to the high levels of the sentiment index, the put/call ratio is dropping sharply in the middle of 2020, as the market recovered strongly. The heterogeneity index also has the biggest spikes in the 2008 financial crisis as well as at the beginning of 2020.

4. Methodology and Model

First, data was collected, sorted, and cleaned such that data quality is ensured and no missing values appear for any date included in the analysis. Dates for which there were any missing values were excluded for all variables. The excel files were prepared such that they can be loaded into Python for further analysis[5]. Subsequently, the models were built to test several hypotheses that will provide insights and answers to the presented research question. Then, several tests were run to check for prerequisites of linear regressions which are linearity, independence (or no perfect collinearity), homoscedasticity (or constant variance of errors) and normality.

4.1 Daily Model

The linearity assumption is checked by plotting scatter plots of each explanatory variable against market excess return. As can be seen in *Appendix 4*, all explanatory variables display a relatively linear relationship with market excess return. The normality of residuals assumption is checked with a Jarque-Bera test. With a statistically significant test-statistic of 25784, our dataset violates the normality assumption. The independence assumption can never be fully

[5] Please find the Python code attached to this work project.

fulfilled when working with time-series stock data because there is always autocorrelation present in this kind of data. Furthermore, residual plots show whether the variance of residuals is constant. In *Appendix 5* it becomes visible that especially for the sentiment variable, residuals are spreading out for small values of the sentiment index. Therefore, robust standard errors are used in the subsequent analysis. There is no large correlation between the variables, as can be seen in *Appendix 6*.

The first model which is constructed with daily data tests the null hypothesis that sentiment (as approximated by the Morgan Stanley sentiment indicator), expected volatility (VIX), tail risk (CBOE tail risk index), and the put/ call ratio have no significant systematic effect on excess market return. To adjust for the information lag, it is considered that at time t, investors act upon their information set of period t-1. Therefore, the explanatory variables were lagged one period. This lag is especially critical when working with daily or even intraday data. Furthermore, standard errors are heteroscedasticity robust (HC0).

Equation 2 below displays the regression equation for the model with daily data. The regression output will provide information about the degree to which market excess return on day t is influenced by sentiment, volatility, and tail risk on day t-1. The error term, ϵ_t, captures the volatility in market excess return that is not explained by the presented variables but might be driven by other factors.

$$(r_{mt} - r_{ft}) = \beta_0 + \beta_1 sent_{t-1} + \beta_2 vol_{t-1} + \beta_3 tailrisk_{t-1} + \beta_4 put\ call\ ratio_{t-1} + \epsilon_t. \tag{2}$$

4.2 Monthly Model

The tests of the regression assumptions revealed that the dataset with monthly data violates the linearity assumption. *Appendix 7* shows that the scatter plot of the VIX against the market excess return is exhibiting a nonlinear pattern as the relationship between the variables diminishes for increasing VIX values. Furthermore, the residuals exhibit heteroscedasticity. In *Appendix 8* it becomes visible that especially for larger values of the VIX, residuals are spreading out significantly. Therefore, the analysis is performed with robust standard errors that correct for the heteroscedasticity.

It is expected that the regression with the monthly data exhibits greater explanatory power than with daily data. One reason for this is that in daily stock data, there is lots of fluctuation which is not related to our variables of interest. By using monthly data, this noise can be excluded. The model aims at analyzing the effect of investor sentiment on stock prices which evolves over time and does normally follow trends that do not revert within hours but rather weeks.

With monthly data, the lag in the independent variables becomes less relevant, as investors do have the chance to update their information sets for period t with information from almost the whole period t-1. Therefore, a first regression is run without lagging explanatory variables, and to complete the analysis, a second regression is performed with the lagged covariates, as presented for the daily model. Thereby, the magnitude of the effect of lagging the independent variables also becomes apparent.

This model tests the null hypothesis that sentiment (as approximated by the index constructed by Baker & Wurgler, 2022), the bullish percent index, expected volatility, tail risk, the ratio of put volumes to call volumes, and the degree of heterogeneity of investor sentiment have no significant effect on excess market return. The resulting regression equation for the same-period model is presented by *Equation 3*: Market excess return in period t is given by an intercept (β_0), the effects of B&W sentiment index, the bullish percent index, the volatility index (VIX), the tail risk index, the put/call ratio, heterogeneity, as well as an error term, ϵ_t, that captures the variability which is not explained by the presented variables. The respective magnitude and direction of the relationships is given by the coefficients ($\beta_i, i = 1, \dots, 6$),

$$\left(r_{m,t} - r_{rf,t}\right) = \beta_0 + \beta_1\, sent_t + \beta_2 bullish\, percent_t + \beta_3 vol_t + \beta_4 tailrisk_t + \beta_5 put\, call\, ratio_t + \beta_6 heterogeneity_t + \epsilon_t. \tag{3}$$

An additional regression tests the effect of lagging: Like in the daily model, explanatory variables are shifted by one period to adjust for the information lag. The resulting regression equation is displayed by *Equation 4*. Here, it is analyzed, which effects sentiment, the share of bullish investors, volatility, tailrisk and put/call ratio in the previous month have on the current month's market excess return,

$$\left(r_{m,t} - r_{rf,t}\right) = \beta_0 + \beta_1\, sent_{t-1} + \beta_2 bullish\, percent_{t-1} + \beta_3 vol_{t-1} + \beta_4 tailrisk_{t-1} + \beta_5 put\, call\, ratio_{t-1} + \beta_6 heterogeneity_{t-1} + \epsilon_t. \tag{4}$$

Ultimately, a sector analysis complements the study by a differentiation of the effect of sentiment and its heterogeneity on different sectors: The effect of the presented sentiment-variables on the financial, IT, healthcare, energy, and real estate sector is assessed with additional regressions. For the sector analysis, tail risk is excluded, and the independent variables are lagged by one period. The resulting regression models are presented by the following equations 5-9,

$$\left(r_{fin,t} - r_{rf,t}\right) = \beta_0 + \beta_1\, sent_{t-1} + \beta_2 bullish\, percent_{t-1} + \beta_3 vol_{t-1} + \beta_4 put\, call\, ratio_{t-1} + \beta_5 heterogeneity_{t-1} + \epsilon_t \tag{5}$$

$$\left(r_{IT,t} - r_{rf,t}\right) = \beta_0 + \beta_1\, sent_{t-1} + \beta_2 bullish\, percent_{t-1} + \beta_3 vol_{t-1} + \beta_4 put\, call\, ratio_{t-1} + \beta_5 heterogeneity_{t-1} + \epsilon_t; \tag{6}$$

$$\left(r_{HC,t} - r_{rf,t}\right) = \beta_0 + \beta_1\, sent_{t-1} + \beta_2 bullish\, percent_{t-1} + \beta_3 vol_{t-1} + \beta_4 put\, call\, ratio_{t-1} + \beta_5 heterogeneity_{t-1} + \epsilon_t; \tag{7}$$

$$\left(r_{RE,t} - r_{rf,t}\right) = \beta_0 + \beta_1\, sent_{t-1} + \beta_2 bullish\, percent_{t-1} + \beta_3 vol_{t-1} + \beta_4 put\, call\, ratio_{t-1} + \beta_5 heterogeneity_{t-1} + \epsilon_t; \tag{8}$$

$$(r_{en,t} - r_{rf,t}) = \beta_0 + \beta_1\, sent_{t-1} + \beta_2 bullish\, percent_{t-1} + \beta_3 vol_{t-1} + \beta_4 put\, call\, ratio_{t-1} + \beta_5 heterogeneity_{t-1} + \epsilon_t. \quad (9)$$

5. Analysis and Outcomes

5.1 Daily Model Regression Analysis

The output for the OLS regression of the lagged daily described variables on the daily excess market return can be seen in *Table 1*. The R-squared is only 0.082 and the coefficients for sentiment and volatility are not significant at a 10% level. The significantly negative coefficient of the put/call ratio is probably most interesting: This regression outcome hints at the fact that a bearish market sentiment has a negative effect on market returns. The coefficient for the tail risk Index is significantly positive; however, this result must be interpreted with caution as the analysis of the monthly data hints at the fact that there might be a multicollinearity bias between the tail risk Index returns and market excess returns. Therefore, the tail risk index will be excluded in later presented analyses.

Table 1: Regression output for the daily model with lagged explanatory variables

Dep. Variable:	S&P 500 Ex Ret	***R-squared:***	0.082
Model:	OLS	***Adj. R-squared:***	0.081
Method:	Least Squares	***F-statistic:***	45.96
No. Observations:	4309	***Covariance Type:***	HC0

	coef	***std err***	***z***	***P>|z|***	***[0.025***	***0.975]***
const	0.0104	0.002	6.681	0.000	0.007	0.013
Sentiment	0.0002	0.000	0.454	0.650	-0.001	0.001
VIX	5.58e-05	4.88e-05	1.142	0.253	-3.99e-05	0.000
CBOE S&P 500 tail risk index	0.3068	0.033	9.358	0.000	0.243	0.371
Put Call Ratio	-0.0178	0.002	-9.043	0.000	-0.022	-0.014

The relatively low predictive power of the daily model can be explained by the medium-term nature of behavioral market variables: investor sentiment and related behavioral investing patterns like herding behavior rather move over weeks/months than daily. Furthermore, there is a substantial amount of 'noise' within daily data that is uncorrelated to the variables of interest. Therefore, it is expected that a regression with monthly data will have more predictive power.

5.2 Monthly Model Regression Analysis

In a first regression, the values for sentiment, volatility (VIX), bullish percent index, tail risk, put/call ratio, and heterogeneity were regressed against market excess return (S&P 500). Firstly, here the explanatory variables were not lagged by one period. *Table 2* presents the regression outcome.

Table 2: *Regression output for the monthly model with same-period explanatory variables*

Dep. Variable:	Market Exc Ret	***R-squared:***	0.885
Model:	OLS	***Adj. R-squared:***	0.882
Method:	Least Squares	***F-statistic:***	305.9
No. Observations:	220	***Covariance Type:***	HC0

	coef	***std err***	***z***	***P>|z|***	***[0.025***	***0.975]***
const	0.0168	0.012	1.081	0.280	-0.010	0.035
Sentiment orth B&W	-0.0021	0.003	-0.815	0.415	-0.007	0.003
VIX monthly	-0.0004	0.000	-1.217	0.224	-0.001	0.000
Heterogeneity	-0.0042	0.014	-0.304	0.761	-0.031	0.023
Tail Risk Index monthly	1.1835	0.041	28.522	0.000	1.102	1.265
Bullish Percent	-0.0194	0.015	-1.278	0.201	-0.049	0.010
Put/Call Ratio	0.0051	0.013	0.398	0.690	-0.020	0.030

The R-squared is extraordinarily high with 0.885. This implies that 88.5% of variation in market excess return can be explained by the model. The adjusted R-squared is with 0.882 slightly lower, which indicates that the inclusion of some variables in the regression model may not be contributing significantly to explaining the variability in market excess return. Subsequent parts of the work project will focus on the significance of the single variables. For a stock-data regression this R-squared is relatively high and might be a sign for simultaneity bias or other biases. Simultaneity bias might occur because there exists a two-way causal relationship between the dependent variable and one or more independent variables. In other words, the relationship between the variables is bidirectional. This might indeed be the case in the model since there exists a feedback loop between market excess returns and investor sentiment: returns affect investor sentiment, and resulting investor actions, which consequently affect market excess return. It seems like even with monthly data, this bias might persist, and it is therefore more accurate, to lag the explanatory variables by one period.

As shown in *Table 2*, the coefficients of orthogonalized Sentiment, the VIX, heterogeneity, and the put/call ratio are not significant at a 10% level. This explains why the adjusted R-squared is lower than the R-squared. Since the tail risk index is highly significant and considering that the index tracks the return of a protective put strategy, it can be presumed that there is a very high correlation between market excess return and the tail risk index. A correlation matrix is presented by *Appendix 9*, and it becomes visible that the correlation between market excess return and the tail risk index is with 0.94 unacceptably high and indicates a multicollinearity bias. No other variables have a correlation above 0.6. Due to this detected multicollinearity problem, the other independent variables become statistically insignificant in the regression because the tail risk index is already explaining much of the variation in market returns, making the additional explanatory power of the other variables less apparent. An exclusion of the tail risk index will provide a higher quality regression and more information about the direction and magnitude of effects of the other variables on market excess return. *Table 3* presents the regression output resulting from an exclusion of the tail risk index.

Table 3: Regression output for monthly model with same-period explanatory variables and excluding tail risk index

Dep. Variable:	Market Exc Ret	***R-squared:***	0.326
Model:	OLS	***Adj. R-squared:***	0.310
Method:	Least Squares	***F-statistic:***	18.85
No. Observations:	220	***Covariance Type:***	HC0

	coef	***std err***	***z***	***P>|z|***	***[0.025***	***0.975]***
const	0.1150	0.043	2.693	0.007	0.031	0.199
Sentiment orth B&W	-0.0158	0.006	-2.825	0.005	-0.027	-0.005
VIX monthly	-0.0017	0.000	-3.900	0.000	-0.003	-0.001
Heterogeneity	-0.0087	0.037	-0.232	0.817	-0.082	0.065
Bullish Percent	0.0855	0.035	2.413	0.016	0.016	0.155
Put/Call Ratio	-0.1513	0.044	-3.437	0.001	-0.238	-0.065

The R-squared dropped to 0.326, which is a more realistic value for a stock-returns regression. Furthermore, all explanatory variables have become significant at the 1% level which underlines the relevance of including investor sentiment and, in particular, its degree of dispersion into asset pricing models. Sentiment, as constructed by the Baker & Wurgler Index, has a negative effect on market return. The same applies for the VIX: A higher market volatility makes investors more hesitant and market excess return drops. The bullish percent index has a positive effect on market excess return, whereas the put/call ratio has a negative coefficient. This is plausible because a higher put/call ratio implies a higher demand for puts than for calls. Since put options are commonly used as insurances against price declines this indicates that investors are relatively cautious and bearish. The negative coefficient of the heterogeneity variable shows that not only a bearish sentiment itself among investors has a negative impact on the stock market. It provides evidence that a high level of dispersion in sentiment among investors, independent of the (bullish/bearish) direction, also has a negative effect on market returns.

The following paragraph addresses the previously discussed simultaneously bias. To mitigate this bias, the explanatory variables are lagged by one period such that the feedback loop between the explanatory variables and market return is broken. Moreover, the lag makes the analysis more realistic and conservative because it incorporates the time-delayed updating of investors' information sets. *Table 4* presents the regression output of the associated model.

Table 4: *Regression output for the monthly model with lagged explanatory variables*

Dep. Variable:	Market Exc Ret	***R-squared:***	0.165
Model:	OLS	***Adj. R-squared:***	0.146
Method:	Least Squares	***F-statistic:***	6.091
No. Observations:	219	***Covariance Type:***	HC0

	coef	***std err***	***z***	***P>|z|***	***[0.025***	***0.975]***
const	0.1521	0.043	3.580	0.000	0.069	0.235
Sentiment orth B&W	-0.0081	0.007	-1.186	0.236	-0.021	0.005
VIX monthly	-0.0010	0.000	-2.283	0.022	-0.002	-0.000
Heterogeneity	-0.0866	0.039	-2.219	0.026	-0.163	-0.010
Bullish Percent	0.0952	0.043	2.235	0.025	0.012	0.179
Put/Call Ratio	-0.0356	0.048	-0.740	0.459	-0.130	0.059

With the incorporation of the lag, the number of observations drops from 220 to 219 because the explanatory variables are shifted back by one period. As expected, the R-squared of the new model falls considerably to 0.165. The model can be considered relatively conservative after variables that violate regression assumptions have been excluded and simultaneously bias has been eliminated by incorporating a lag. Considering this conservativeness and the fact that the analysis is restricted to sentiment-related variables, it is still satisfactory that 16.5% of variation in market excess return can be explained by the model. The coefficients of the sentiment index and the put/call ratio become insignificant due to the lag. All other independent variables are still significant at a 5% level. Particularly noteworthy is that the heterogeneity coefficient is significantly negative. This further strengthens the hypothesis that not only a generally bearish market sentiment has a negative effect on market returns but that a high level of heterogeneity in investor sentiment also negatively impacts market returns.

5.3 Sentiment and the Fama French Factor Model (FFM)

The Fama & French three-factor model aims at capturing additional risk factors that investors require a return for. Next to market excess return, as captured by the traditional CAPM model, the additional risk factors are size and book-to-market ratio. A regression of the factors (retrieved from the website of Kenneth R. French[6]) on the S&P 500 excess return and incorporating a one-period lag leads to an R-squared of 0.009 and insignificant coefficients. The adjusted R-squared is even only -0.005 (see *Appendix 10*). Removing the lag has an enormous effect: the R-squared becomes 0.997, the adjusted R-squared increases to 0.996, and all 3 factors are significant (see *Appendix 11*). This shows that the FFM3 is highly sensitive to simultaneity which is consistent with the results in Allen and McAleer (2018). The authors conclude that there are significant time varying relationships among the variables and incorporate (among others) a lag to reduce the bias of the model (Allen and McAleer 2018). Since most of the variations in the S&P 500 is most likely explained by the FFM market risk factor and there is a very high correlation of 0.99 between these two variables, the market risk factor will be excluded for further analysis. By this means, the FFM factors become more comparable to the factors of this work project's sentiment model (in which market risk is also excluded as a coefficient). *Appendix 12* presents the output of a regression of only the size

[6] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Breakpoints

(SMB) and value (HML) factor on the S&P 500 excess return. The adjusted R-squared drops to 0.094, and while the SMB coefficient is significantly positive, the HML is positive but with a p-value of 6%.

In the following, it is analyzed whether this 2-factor model can be improved by incorporating sentiment and heterogeneity. Since the R-squared always increases by adding more explanatory variables, the adjusted R-squares, will be compared here. *Appendix 13* presents the output of a regression of the SMB and HML factor as well as this work project's sentiment variables on the S&P 500 excess return. The addition of the sentiment and heterogeneity factors to the 2-factor model significantly improves the model: The adjusted R-squared increases from 0.094 to 0.34. This implies that even without the market-risk factor, 34% of variation in the S&P 500 can be explained by the model. Except for the heterogeneity coefficient, all coefficients are significant at the 5% level. This regression has strengthened the hypothesis that traditional financial models that assume fully rational investors and efficient markets cannot explain a substantial part of variation in real-world stock data. Such models can be improved by considering behavioral aspects and adding investor sentiment and its degree of heterogeneity.

5.4 Sector Analysis

The following paragraph extends the study by a sector analysis and breaks down the effect of sentiment and behavioral heterogeneity on asset prices of different sectors. Additional regression analyses are performed for the financial sector, the IT sector, the healthcare sector, energy sector, and the real estate sector. Indices that track the respective sectors serve as dependent variables. To create a conservative and more realistic outcome, the independent variables are again lagged by one period.

The R-squared for the financial sector regression is with 0.181 higher than the one presented above for the general market regression (see *Appendix 14* for regression output). This indicates that investor sentiment and behavioral heterogeneity have a stronger effect on the financial sector because more variation of sector returns can be explained by the model's variables. The coefficient of the bullish percent index has increased, implying that the magnitude of its effect on the financial-sector returns has risen. The same holds true for the absolute values of the negative coefficients for volatility and the put/call ratio. The B&W sentiment coefficient is still negative, however not significant (after introducing the lag).

Contrary to the observation for the financial sector, the R-squared for the IT-sector regression is with 0.095 smaller than the one for the general market (*Appendix 15*). This suggests that the IT sector is less affected by investor sentiment and behavioral heterogeneity. Furthermore, all priorly significant variables decrease in their significance, and except for the bullish percent index, all other variables become insignificant at a 10% level. It is striking that for the IT-sector the lag has such a big effect on the significance of the variables: *Appendix 16* shows that without a lag, all explanatory variables, except for the heterogeneity index, are highly significant at a 5% level. This might indicate that returns in the IT-sector move especially dynamically and fast such that a lag of one month reduces predictive power of sentiment variables enormously.

For the healthcare sector (*Appendix 17*), the R-squared is with 0.099 only slightly higher than for the IT sector (with lag). However, except for volatility, all explanatory variables become insignificant at an alpha of 5% level. It turns out that even without the lag, only volatility and put/call ratio are significant indicators for returns in the healthcare sector. This implies, that investor sentiment might not have as a significant effect on the healthcare sector as on other sectors.

The regression output for the energy-sector regression exhibits similar characteristics as the R-squared is with 0.091 much lower than for the general market and only the coefficient for the VIX is significant (*Appendix 18*). Without the lag, only VIX and bullish percent index are significant. The analysis indicates that the energy and healthcare sectors are less affected by investor sentiment as other sectors. This might be because for these sectors, demand is determined by more exogeneous factors such as seasonal temperature fluctuations or diseases and therefore, less prone to sentiment as e.g., the financial sector.

The real-estate sector seems more affected by investor sentiment and heterogeneity. As can be seen in *Appendix 19*, the R-squared for the regression on the real estate index is 0.161 which is close to the one for the general market. The bullish percent index has a significantly positive coefficient and the VIX a significantly negative coefficient. Orthogonalized sentiment by Baker & Wurgler, heterogeneity, and put/call ratio have negative coefficients which are, however, not significant.

The sector analysis has shown that asset pricing models do not work equally well for each sector and a sector-specific differentiation can be beneficial.

6. Conclusion

This work project contributes to the academic behavioral finance literature by providing a detailed and structured analysis of the effect of investor sentiment and its degree of heterogeneity on market returns. Only a detailed understanding of the origins of asset price movements, allows analysts and federal institutions to generate precise forecasts and risk assessments.

The findings of the work project can be summarized as follows: Generally, investor sentiment plays an important role for asset prices and can explain up to almost 20% of variation in market excess returns. This finding contrasts with the efficient markets hypothesis which assumes that investors are rational and security prices correctly and fully reflect all public information. Furthermore, the presented analysis provides evidence for the hypothesis that heterogeneity in investor sentiment, regardless of the bullish/bearish direction of the sentiment itself, has a negative effect on asset prices. An addition of the sentiment and heterogeneity factors to the Fama French factor model has revealed that the inclusion of such behavioral factors increases the explanatory power of asset pricing factor models. A sector analysis has shown that sentiment and its degree of dispersion among investors have a bigger effect on the financial sector than on other sectors which were included in the analysis.

6.1 Limitations and Further Research

The most relevant limitation for all existing models aiming at analyzing investor behavior is the measurement of such behavior. For this work project, this applies likewise to the measurement of sentiment. It is probably almost impossible to completely capture all relevant behavioral investor characteristics and certainly more research should be dedicated to improving and refining the measurement of investor sentiment. Furthermore, the data which is analyzed in this work project is restricted to the US market. More research shall be dedicated to a wider set of stock market data to detect potential regional differences in the effect of investor sentiment on market excess returns.

References

abc News. 2022. *Why the stock market has fallen in December, and what it means going forward.* December 15. Accessed November 2023. https://abcnews.go.com/Business/stock-market-fallen-december-means-forward/story?id=95066510.

Allen, David, and Michael McAleer. 2018. "'Choosing Factors' by Fama and French (2018): A Comment." *SSRN.*

Babalos, V, A Kostakis, N Philippas , and F Economou. 2013. "Herding behavior in REITs: Novel tests and the role of financial crisis." *International Review of Financial Analysis* 166-174.

Baker, Malcolm, and Jeffrey Wurgler. 2006. "Investor Sentiment and the Cross-Section of Stock Returns." *The Journal of Finance* 1645-1680.

Barone-Adesi, Giovanni, Loriano Mancini, and Shefrin Hersh. 2011. "Systemic Risk and Sentiment." *National Centre of Competence in Research Financial Valuation and Risk Management* Working Paper No. 739.

Barras, Laurent. 2019. "A large-scale approach for evaluating asset pricing models." *Journal of Financial Economics* 549-569.

Bloomberg. 2023. *CBOE Volatility Index.* New York, October.

Carlin, Bruce, Francis Lingstaff, and Kyle Matoba. 2014. "Disagreement and asset prices." *Journal of Financial Economics* 226-238.

CBOE Global Indices, LLC. 2021. "CBOE Global Indices." *EXTERNAL Cboe S&P 500 Tail Risk Index.* Accessed October 29, 2023. https://cdn.cboe.com/api/global/us_indices/governance/PPUT3M_Methodology.pdf.

Cornea-Madeira, Adriana, and Joao Madeira. 2020. "Intraday dynamics of behavioral heterogeneity in stock prices." *University of York.*

Doukas, John A, Chansog Kim, and Christos Pantzalis. 2004. "Divergent Opinions and the Performance of Value Stocks." *Financial Analysts Journal* 55-64.

Fama, Eugene F., and Kenneth R. French. 2012. “Size, value, and momentum in international stock returns.” *Journal of Financial Economics* 457-472.

French, Eugene F, and Kenneth R French. 1993. "Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics* 3-56.

Habibah, Ume, M Bhayo, and M Shahid Iqbal. 2021. "Investor Sentiments and Fama–French Five-Factor Premia." *SAGE Open.*

Huang, T, and K Wang. 2017. "Investors’ Fear and Herding Behavior: Evidence from the Taiwan Stock Market." *Emerging Markets Finance and Trade* 2259 - 2278.

Lamont, A Owen, and C Jeremy Stein. 2004. "Aggregate Short Interest and Market Valuations." *AMERICAN ECONOMIC REVIEW* 29-32.

Nguyen, T. 2022. "A Study of Herding Behavior on Vietnam Stock Market." *Journal of Economics, Finance and Accounting Studies.*

Ren, R, and D Wu. 2020. "An Innovative Sentiment Analysis to Measure Herd Behavior." *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 3841-3851.

Reuters . 2022. *Wall Street ends 2022 with biggest annual drop since 2008.* Decmber 30. Accessed November 2023. https://www.reuters.com/markets/us/futures-slip-last-trading-day-torrid-year-2022-12-30/.

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

Shefrin, Hersh. 2008. *A Behavioral Approach to Asset Pricing.* Santa Clara University: Elsevier.

Welch, Ivo, and Amit Goyal. 2008. "Empirical Performance of Equity Premium Prediction." *The Review of Financial Studies* 21 (4).

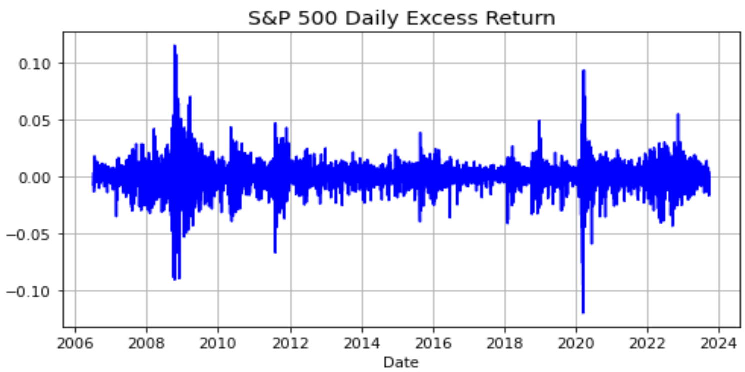
Wurgler, Jeffrey. 2022. *NYU Department of Finance.* Accessed September 2023. https://pages.stern.nyu.edu/~jwurgler/.

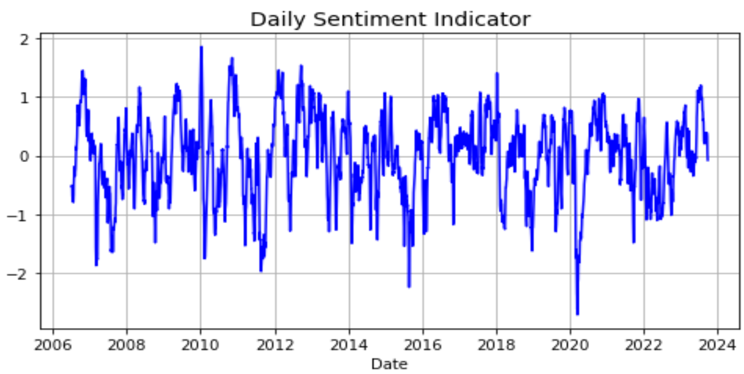
Appendix

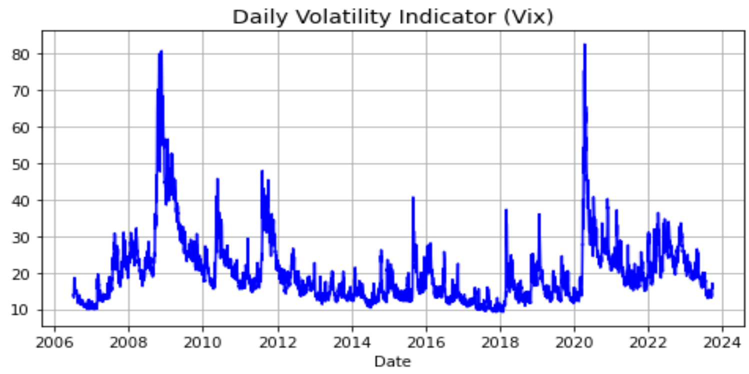
Appendix 1: Description of the daily data

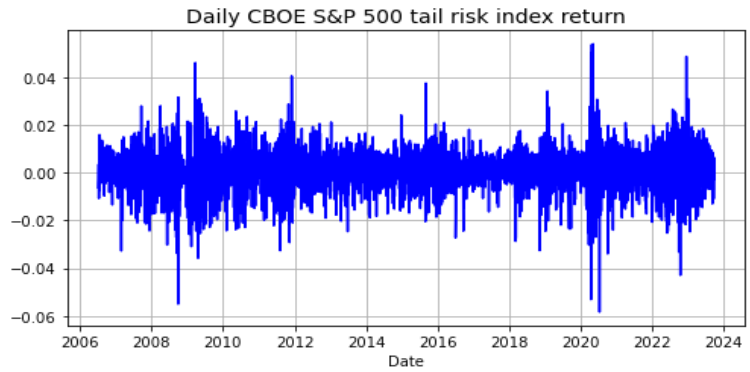
	S&P 500 Ex Ret	*Sentiment*	*VIX*	*CBOE S&P 500 tail risk index*	*Put Call Ratio*
count	4310	4310	4310	4310	4310
mean	0.000202	0.000664	19.999724	0.000324	0.638480
std	0.012706	0.698948	9.158996	0.008697	0.129075
min	-0.119902	-2.720000	9.140000	-0.058424	0.320000
25%	-0.004392	-0.430000	13.790000	-0.003658	0.560000
s50%	0.000544	0.060000	17.625000	0.000431	0.630000
75%	0.005719	0.500000	23.250000	0.004889	0.700000
max	0.115586	1.860000	82.690000	0.054013	2.400000

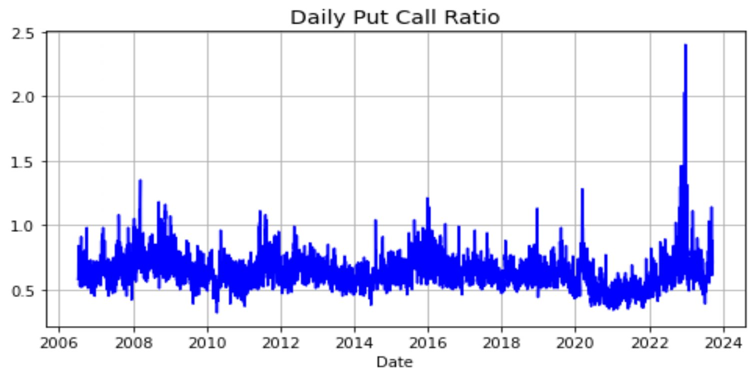
Appendix 2: Time – Series Plots for the Daily Model Variables



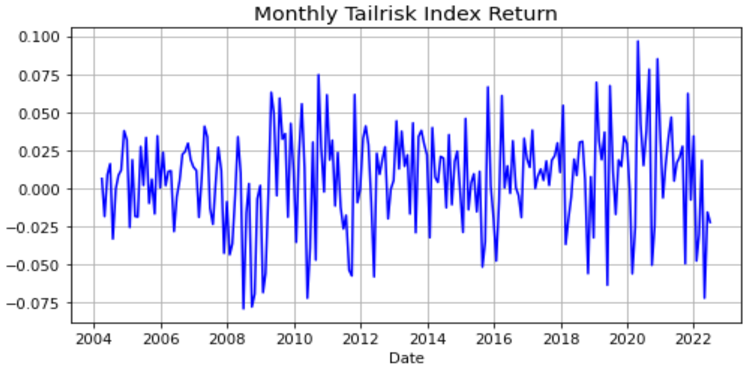




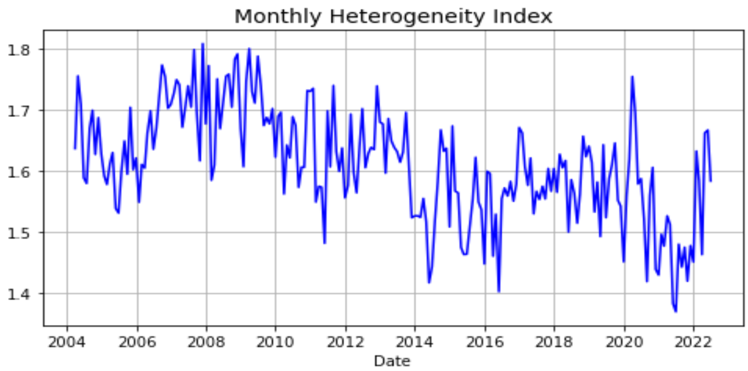


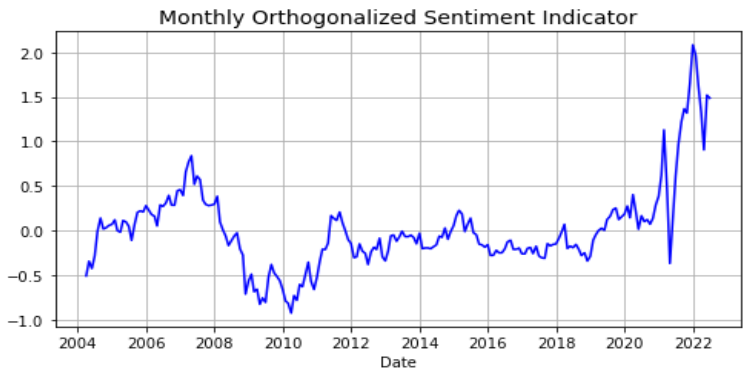


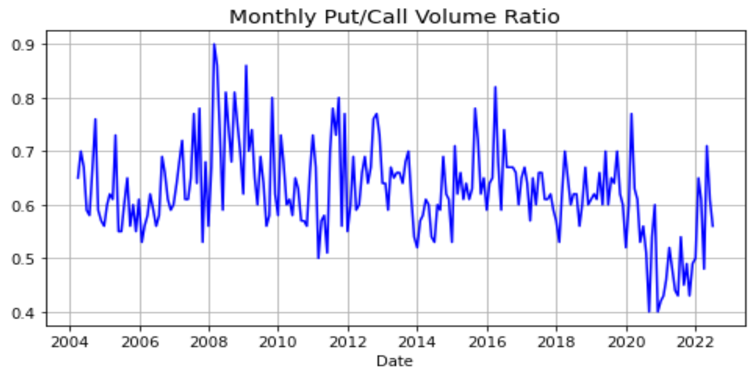
Appendix 3: Time-series plots for the monthly model variables


Monthly Tailrisk Index Return
0.100
0.075
0.050
0.025
0.000
−0.025
−0.050
−0.075
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022
Date

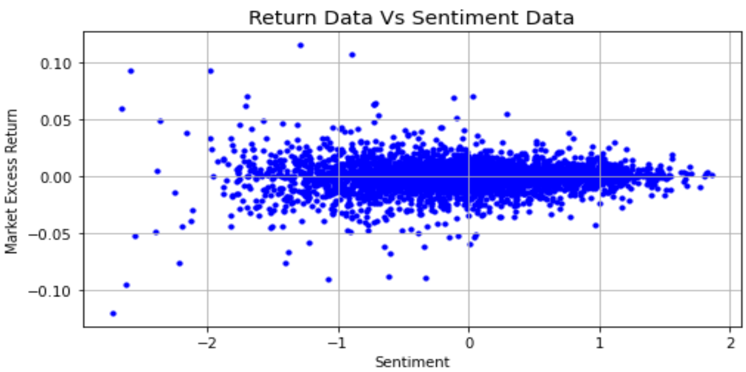

Monthly Volatility Index (VIX)
60
50
40
30
20
10
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022
Date

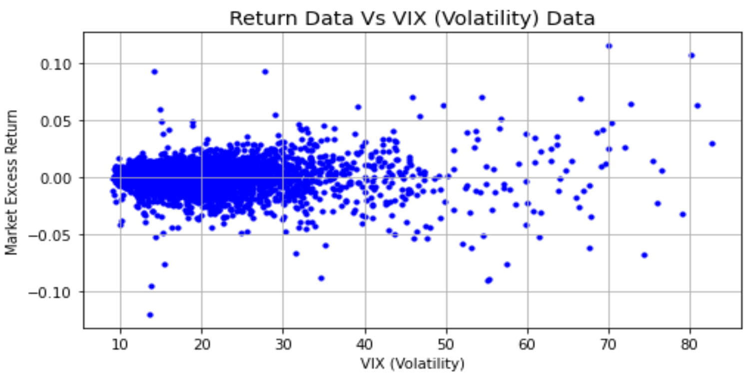

Monthly Heterogeneity Index
1.8
1.7
1.6
1.5
1.4
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022
Date

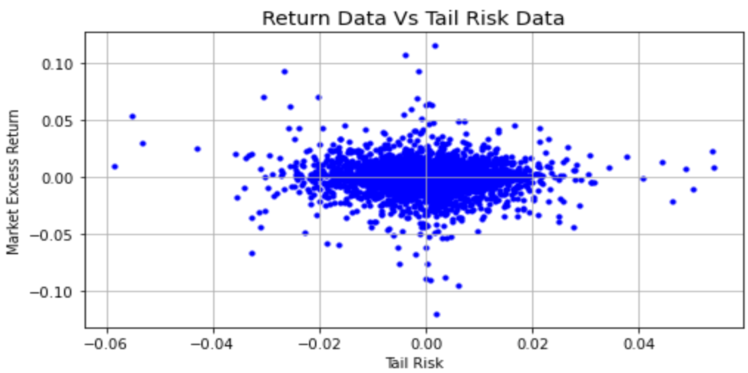

Monthly Orthogonalized Sentiment Indicator
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022
Date



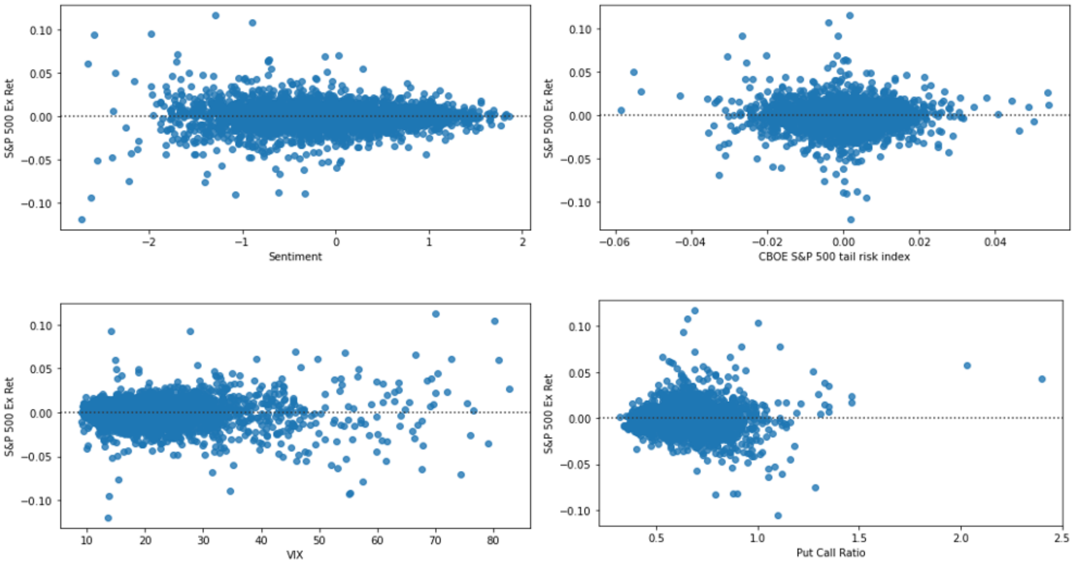
Appendix 4: Scatter plots for daily variables (checking linearity assumption)







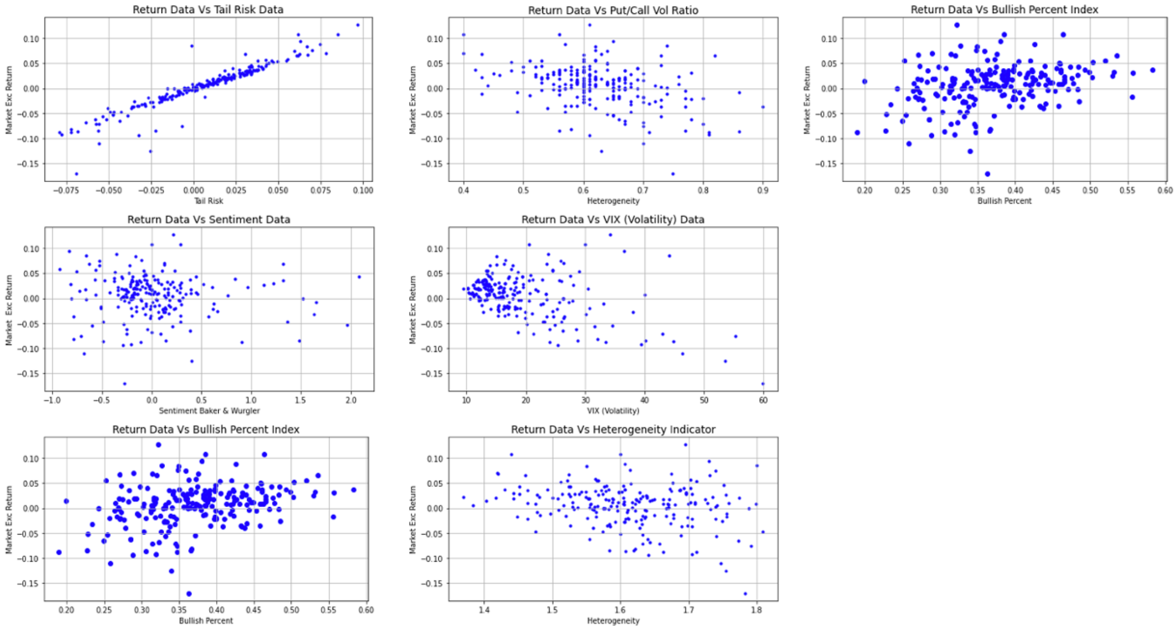
Appendix 5: Residual Plots of daily regression variables (checking for homoskedasticity)



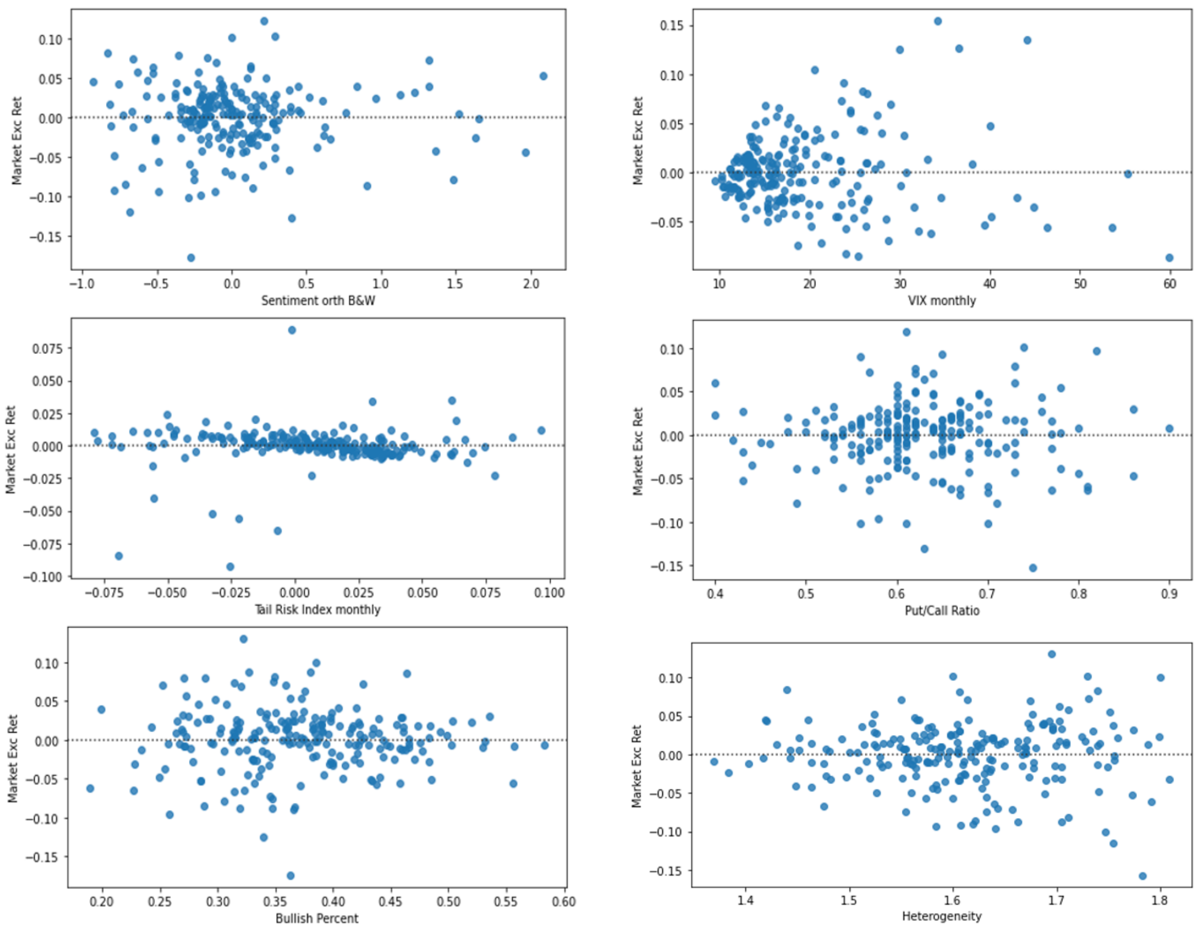
Appendix 6: Correlation Matrix for the daily regression variables

	S&P 500 Ex Ret	***Sentiment***	***VIX***	***tail risk index***	***Put Call Ratio***
S&P 500 Ex Ret	1.000000	0.026336	0.042475	-0.041125	-0.317885
Sentiment	0.026336	1.000000	-0.28868	0.097691	-0.270568
VIX	0.042475	-0.288689	1.000000	-0.110014	0.162668
tail risk index	-0.041125	0.097691	-0.11001	1.000000	-0.088102
Put Call Ratio	-0.317885	-0.270568	0.162668	-0.088102	1.000000

Appendix 7: Scatter plots for monthly variables (checking for linearity)



Appendix 8: Residual Plots of monthly regression variables (checking for homoskedasticity)



Appendix 9: Correlation Matrix for the monthly regression variables

	Market Exc Ret	***Sentiment orth B&W***	***Bullish Percent***	***VIX monthly***	***Tail Risk Index monthly***	***Heterogeneity***	***Put/Call Ratio***
Market Exc Ret	1	-0.082442	0.304038	-0.44171	0.937178	-0.227305	-0.37259
Sentiment orth B&W	-0.08244	1	-0.11260	-0.05926	-0.072172	-0.229341	-0.29149
Bullish Percent	0.304038	-0.112600	1	-0.23852	0.347501	0.181688	-0.19751
VIX monthly	-0.44171	-0.059269	-0.23852	1	-0.400174	0.302977	0.241147
Tail Risk Index monthly	0.937178	-0.072172	0.347501	-0.40017	1	-0.216418	-0.40527
Heterogeneity	-0.22730	-0.229341	0.181688	0.302977	-0.216418	1	0.572567
Put/Call Ratio	-0.37259	-0.291498	-0.19751	0.241147	-0.405274	0.572567	1

Appendix 10: Regression output for the FFM3 Model with lagged factors

Dep. Variable:	Market Exret	***R-squared:***	0.009
Model:	OLS	***Adj. R-squared:***	-0.005
Method:	Least Squares	***F-statistic:***	0.3492
No. Observations:	219	***Covariance Type:***	HC0

	coef	***std err***	***z***	***P>|z|***	***[0.025***	***0.975]***
const	0.0054	0.003	1.927	0.054	-9.29e-05	0.011
Mkt-RF	0.0004	0.001	0.394	0.693	-0.002	0.002
SMB	0.0004	0.001	0.368	0.713	-0.002	0.003
HML	0.0009	0.001	0.838	0.402	-0.001	0.003

Appendix 11: Regression output for the FFM3 Model with same-period factors

Dep. Variable:	Market Exret	***R-squared:***	0.997
Model:	OLS	***Adj. R-squared:***	0.996
Method:	Least Squares	***F-statistic:***	1.827e+04
No. Observations:	220	***Covariance Type:***	HC0

	coef	std err	z	P>|z|	[0.025	0.975]
const	-0.0018	0.000	-10.499	0.000	-0.002	-0.001
Mkt-RF	0.0099	4.53e-05	218.349	0.000	0.010	0.010
SMB	-0.0015	8.58e-05	-17.924	0.000	-0.002	-0.001
HML	0.0002	6.67e-05	2.526	0.012	3.78e-05	0.000

Appendix 12: Regression output for the FFM Model without Mkt-RF and same-period factors

Dep. Variable:	Market Exret	**R-squared:**	0.102
Model:	OLS	**Adj. R-squared:**	0.094
Method:	Least Squares	**F-statistic:**	10.06
No. Observations:	220	**Covariance Type:**	HC0

	coef	std err	z	P>|z|	[0.025	0.975]
const	0.0052	0.003	1.922	0.055	-0.000	0.010
SMB	0.0047	0.001	3.979	0.000	0.002	0.007
HML	0.0021	0.001	1.879	0.060	-9.15e-05	0.004

Appendix 13: Regression output for the full model with FFM factors and sentiment variables

Dep. Variable:	Market Exret	**R-squared:**	0.369
Model:	OLS	**Adj. R-squared:**	0.348
Method:	Least Squares	**F-statistic:**	17.37
No. Observations:	220	**Covariance Type:**	HC0

	coef	std err	z	P>|z|	[0.025	0.975]
const	0.1098	0.041	2.655	0.008	0.029	0.191
Sentiment orth B&W	-0.0148	0.006	-2.583	0.010	-0.026	-0.004
Bullish Percent	0.0720	0.035	2.037	0.042	0.003	0.141
VIX monthly	-0.0016	0.000	-4.043	0.000	-0.002	-0.001
Heterogeneity	-0.0092	0.036	-0.258	0.796	-0.079	0.061
Put/Call Ratio	-0.1357	0.042	-3.225	0.001	-0.218	-0.053

SMB	0.0029	0.001	2.883	0.004	0.001	0.005
HML	0.0016	0.001	2.089	0.037	9.85e-05	0.003

Appendix 14: Regression output for the financial-sector regression

Dep. Variable:	Financials sector	***R-squared:***	0.181
Model:	OLS	***Adj. R-squared:***	0.162
Method:	Least Squares	***F-statistic:***	7.488
No. Observations:	219	***Covariance Type:***	HC0

	coef	***std err***	***z***	***P>|z|***	***[0.025***	***0.975]***
const	0.1425	0.062	2.281	0.023	0.020	0.265
Sentiment orth B&W	-0.0072	0.010	-0.742	0.458	-0.026	0.012
Bullish Percent	0.1492	0.055	2.698	0.007	0.041	0.258
VIX monthly	-0.0018	0.001	-2.682	0.007	-0.003	-0.000
Heterogeneity	-0.0702	0.059	-1.197	0.231	-0.185	0.045
Put/Call Ratio	-0.0767	0.065	-1.188	0.235	-0.203	0.050

Appendix 15: Regression output for the IT-sector regression

Dep. Variable:	IT sector	***R-squared:***	0.095
Model:	OLS	***Adj. R-squared:***	0.074
Method:	Least Squares	***F-statistic:***	3.818
No. Observations:	219	***Covariance Type:***	HC0

	coef	***std err***	***z***	***P>|z|***	***[0.025***	***0.975]***
const	0.1401	0.060	2.325	0.020	0.022	0.258
Sentiment orth B&W	-0.0079	0.008	-0.956	0.339	-0.024	0.008
Bullish Percent	0.1028	0.057	1.804	0.071	-0.009	0.215
VIX monthly	-0.0008	0.001	-1.444	0.149	-0.002	0.000
Heterogeneity	-0.0749	0.056	-1.348	0.178	-0.184	0.034
Put/Call Ratio	-0.0491	0.068	-0.725	0.469	-0.182	0.084

Appendix 16: Regression output for the IT-sector regression with same-period explanatory variables

Dep. Variable:	IT sector	***R-squared:***	0.242
Model:	OLS	***Adj. R-squared:***	0.224
Method:	Least Squares	***F-statistic:***	13.44
No. Observations	220	***Covariance Type***	HC0

	coef	***std err***	***z***	***P>|z|***	***[0.025***	***0.975]***
const	0.1731	0.061	2.860	0.004	0.054	0.292
Sentiment orth B&W	-0.0182	0.007	-2.731	0.006	-0.031	-0.005
Bullish Percent	0.1325	0.047	2.816	0.005	0.040	0.225
VIX monthly	-0.0014	0.001	-2.729	0.006	-0.002	-0.000
Heterogeneity	-0.0600	0.050	-1.189	0.234	-0.159	0.039
Put/Call Ratio	-0.1398	0.055	-2.532	0.011	-0.248	-0.032

Appendix 17: Regression output for the healthcare-sector regression

Dep. Variable:	Healthcare sector	***R-squared:***	0.099
Model:	OLS	***Adj. R-squared:***	0.078
Method:	Least Squares	***F-statistic:***	3.511
No. Observations:	219	***Covariance Type***	HC0

	coef	***std err***	***z***	***P>|z|***	***[0.025***	***0.975]***
const	0.1376	0.043	3.170	0.002	0.053	0.223
Sentiment orth B&W	-0.0046	0.007	-0.643	0.520	-0.019	0.009
Bullish Percent	0.0048	0.042	0.116	0.907	-0.077	0.086
VIX monthly	-0.0010	0.000	-2.440	0.015	-0.002	-0.000
Heterogeneity	-0.0558	0.039	-1.445	0.148	-0.131	0.020
Put/Call Ratio	-0.0377	0.044	-0.859	0.390	-0.124	0.048

Appendix 18: Regression output for the energy-sector regression

Dep. Variable:	Energy sector	***R-squared:***	0.091
Model:	OLS	***Adj. R-squared:***	0.070
Method:	Least Squares	***F-statistic:***	3.333
No. Observations:	219	***Covariance Type***	HC0

	coef	***std err***	***z***	***P>|z|***	***[0.025***	***0.975]***
const	0.1095	0.081	1.345	0.179	-0.050	0.269
Sentiment orth B&W	0.0102	0.011	0.946	0.344	-0.011	0.031
Bullish Percent	0.0779	0.076	1.019	0.308	-0.072	0.228
VIX monthly	-0.0015	0.001	-2.322	0.020	-0.003	-0.000
Heterogeneity	-0.0411	0.072	-0.571	0.568	-0.182	0.100
Put/Call Ratio	-0.0592	0.077	-0.765	0.444	-0.211	0.093

Appendix 19: Regression output for the real estate-sector regression

Dep. Variable:	RE sector	***R-squared:***	0.161
Model:	OLS	***Adj. R-squared:***	0.141
Method:	Least Squares	***F-statistic:***	4.991
No. Observations:	219	***Covariance Type***	HC0

	coef	***std err***	***z***	***P>|z|***	***[0.025***	***0.975]***
const	0.1011	0.057	1.763	0.078	-0.011	0.213
Sentiment orth B&W	-0.0046	0.009	-0.509	0.611	-0.022	0.013
Bullish Percent	0.1447	0.050	2.886	0.004	0.046	0.243
VIX monthly	-0.0016	0.001	-2.575	0.010	-0.003	-0.000
Heterogeneity	-0.0725	0.056	-1.304	0.192	-0.181	0.036
Put/Call Ratio	-0.0027	0.063	-0.042	0.966	-0.127	0.122