ERASMUS UNIVERSITEIT ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS

The role of investor sentiment on the value premium

BACHELOR THESIS FINANCE

Ismael Junior Yerbanga

Student number: 409789

Supervisor: Dr. Maurizio Montone

Date: 10th July 2017

Abstract

This paper investigates whether investor sentiment has a predictive power on value premium in the United States. I found that Fama and French's (2015) five-factor model, augmented with an investor sentiment index as well as momentum and liquidity factors, is able to capture the value weighted value premium obtained from book-to-market ratios rankings well. Although it is argued that the value premium is related to a combination of multiple factors, the results suggest a fair positive interrelation between changes in investor sentiment and future value premia. Also, that correlation is more pronounced when the value premium is defined as the average value weighted return difference between the two extreme book-to-market ratios deciles. An analysis conducted on two distinct subgroups indicates that, among low limits-on-arbitrage stocks, shifts in investor sentiment are generally significantly related to future equally weighted value premia. After weighting portfolio stocks returns on value, the relation is more ambiguous. Lastly, the evidence supports the view that overpricing prevails more than underpricing.

Keywords: value premium, investor sentiment, limits on arbitrage

Table of Contents

I. Introduction

Decades ago the challenge was to identify whether investor sentiment had an effect on stocks returns. Investor sentiment can be defined as the aggregate attitude of investors towards price developments in a market. Now, the focus of researchers in the field has increasingly been directed on how to measure investor sentiment and quantify its effects. Existing literature made clear that a negative relationship exists between investor sentiment and stock returns. More specifically excessive optimism or pessimism can drive prices above or below intrinsic values. Accordingly, affected stocks will experience a low or high return following these period of high/low sentiment (Baker & Wurgler, 2007). This research primarily seeks to explore the relation between investor sentiment and the value premium, which corresponds to the greater absolute or risk-adjusted average returns on value stocks compared to growth stocks.

The value premium has been acknowledged since Fama and French (1992) in their three factor model. Furthermore, the existence and persistence of the value premium has been confirmed by numerous studies. For instance, Lakonishok, Sheifler and Vishny (1994) showed with a 1968 to 1989 sample that value stocks outperform growth stocks even in times of economic recession. The persistence of the value premium has caught the attention of many researchers, from the inception of Sharpe's (1964) single factor CAPM, to Fama and French's (1992) three-factor model unable to explain the value anomalies identified. Researchers have since then tried to explain this value premium and identified factors that could explain the premium.

Two alternative explanations prevail in explaining the value premium. One suggests that the premium is a compensation for risks such as costly reversibility and the countercyclical price of risk (Guo, Savickas, Wang, & Yang, 2009). Cooper (2006), Li, Brooks and Miffre (2009), and Gulen, Xing and Zhang (2011) also believe that the value premium is due to value firms being less flexible during adverse economic conditions relative to growth firms. Furthermore, Zhang (2005) argued that assets in place are riskier than growth options. Since it is costlier for firms to scale down capital than to expand, value firms are hit more negatively by economic turndowns. Similar to Zhang (2005), Carlson, Fisher and Giammarino (2004) and García-Feijóo and Jorgensen (2010) conjectured that the value premium relates to operating leverage.

The other explanation implies that investor sentiment has an effect on stock returns due to the irrational behaviour driving prices below or above the intrinsic value, with a lasting effect when binding arbitrage constraints exist. D'avolio (2002) analysed the market for borrowing stocks and concluded that investor optimism can impede arbitrage through the loan market. Du (2011) researched on the matter in a joint study by analysing the correlation between the value premium and the state of economy and found little evidence that the value premium is due to aggregate economic risk. Daniel and Titman (1997) directly tested whether the return premium relates to pervasive factors and conclude from their results that (1) there are no observable risk factors associated with the return premium as identified by Fama and French's (1992) three-factor model and (2) the book-to-market (B/M) characteristic ratio could not be linked to any detectable separate risk factor. The potential impact and influence of behavioural biases is therefore plausible.

Despite the various economic explanations of the value premium, including the heterogeneity in systematic risks, the standard economic model has increasingly been challenged. With the aim of improving the standard model, researchers in behavioural finance have proposed an alternative model built on two assumptions. The first one originates from De Long, Shleifer, Summers and Waldmann (1990) and points out that investors are subject to sentiments, a phenomenon that fosters noise trader risk. The second assumption as suggested by Shleifer and Vishny (1997) is that idiosyncratic volatility matters and deters arbitrage, when prices deviate far from fundamental levels. Arbitraging against sentimental investors is therefore costly and risky, creating limits to arbitrage. These limits to arbitrage and their effect on stock markets have been clarified by recent stock shocks such as the Internet bubble in the 90s and the following Nasdaq and telecom crashes (Baker & Wurgler, 2007). Brav, Heaton, and Li (2009) tested the limits of arbitrage claim and found support for the overvaluation anomaly of growth stocks in general.

If investor sentiment has an effect on the future returns of stocks, unless the value premium truly is solely a compensation for certain risks not captured by the CAPM, one can expect that on top of affecting returns, investor sentiment waves influence the value premium. Studying the sentiment of investors is important for two main reasons. Firstly, it teaches us about the biases affecting stock returns and potentially the value premium, and secondly, it may elucidate on extra returns that can be obtained by understanding the impact of these biases. Accordingly, this research paper will investigate how the investor sentiment is related to the value premium in the United States (U.S.) stock markets.

First, as past bubbles demonstrated a significant positive impact of investor sentiment on stock prices, a positive relation between shifts in investor sentiment and subsequent value premia is expected. This tendency would originate from a combination of excessive optimism towards growth stocks reversion and arbitrage forces. Second, the greater the limits-on-arbitrage, the stronger the relationship between investor sentiment and value premium should be, this follow directly from the key conditions for investor sentiment to significantly affect stock returns. Third, the link between investor sentiment and value premium is tested against a potential economic risk component in the value premium. The degree of relation observed between investor sentiment and value premium should be similar regardless of economic conditions if arbitrage forces and sentiment mechanisms maintain akin characteristics across these states.

In this paper, I analyse the relationship between an investor sentiment index and the value premium conditional on alternative partial explanations and across groups with different limits-on-arbitrage and during contrasting economic conditions. With the intend to empirically extend the value premium literature based on Gulen, Xing, and Zhang (2011), Baker and Wurgler (2006) and Du (2011), I find that investor sentiment has predictive power over the value premium in certain settings. Additionally, the value premium appears to also be correlated with other risk and premium factors than aggregate economic conditions. Economic state specific regressions suggest that shifts in investor sentiment have a significant positive association to subsequent value premium only outside of recession periods, when investor sentiment tends to be relatively high. This scenario is consistent with investor sentiment being positively correlated with aggregate economic conditions.

The remainder of the paper is built upon sections. Section II discusses the theoretical foundations supporting the research, Section III explores the data and variables employed, Section IV examines the methodology and research procedures, Section V presents and discusses the results. Finally, the conclusions reached, limitations of the research, and suggestions for future research are elaborated in Section VI.

II. Theoretical Framework

Theoretically, the effect of investor sentiment should be larger on securities whose valuations are highly subjective and difficult to arbitrage, and this has been shown by Baker and Wurgler (2006). Following Miller's (1977) argument that impediments to arbitrage drives overpricing of securities more than under-pricing in a market of well informed investors, the first anticipation of this research is that investor sentiment waves have a greater effect on growth stocks compared to value stocks. During increasing sentiment periods investors are bullish and more inclined to sentiment based demand of (speculative) growth stocks (Du, 2011). In contrast, during decreasing sentiment periods, bearish investors would foster a downwards pricing of growth stocks. According to theoretical expectations, there is a negative relation between prices and future returns due to both arbitrage forces and mean reversion of investor sentiment. Hence, high sentiment periods would inflate the prices of growth stocks and such periods would be followed by relatively low returns on growth stocks and low sentiment periods would be followed by relatively high returns on growth stocks by the same mechanism. The first hypothesis therefore expects the following:

H1: There is a positive relationship between investor sentiment and subsequent value premia

In a research using the Standard & Poor's (S&P) index options, Han (2007) determined that the investor sentiment has a stronger impact when impediments to arbitrage in the index options are greater. Additional support is brought by Wang and Yu (2014) who indicated that mispricing is more prominent on high limits-on-arbitrage firms than on low limits-on-arbitrage firms, such that the value premium is stronger among high to arbitrage groups than low to arbitrage groups. Past literature also suggests that momentum in the same direction help preserve larger mispricing, since the typical investors tend to behave sub optimally and extrapolate glamour stocks' past growth rates (Lakonishok, Sheifler, & Vishny, 1994). The second hypothesis thus investigates the following:

H2: Investor sentiment has a stronger link with the value premium of high limits-on-arbitrage stocks compared to that of low limits-on-arbitrage stocks

Previous literature mostly stipulates that the value premium varies conversely to the prior economic state. Gulen, Xing, and Zhang (2011) demonstrated a considerable countercyclical variation in the expected value premium by using the National Bureau of Economic Research's (NBER) economic index. Excess returns on value stocks are most affected during recession, raising the value premium, whereas during expansions, the return differential on value and growth stocks is mostly insignificant. As a result, the value premium quickly jumps during recession and gradually declines in the following period. Furthermore, Black and Fraser (2003) observed that the value premium varies over time in the U.S. and is negatively associated to past cumulative GDP growth. The test has been performed using logs of past cumulative real GDP growth over a period equal or greater than two years and indicates that long periods of shrinking GDP growth seem to be consistent with a rising value premium in the following period. Additionally, Kiku (2006) argued that differences in long-run consumptions risks in cash flows from assets are important determinants of the value premium. This long-run risks model implies that the value premium increases during high economic uncertainty and decreases during favourable economic conditions. However, recent evidence from Du (2011) indicates that while the value premium is indeed correlated to investor sentiment, it shows a weak correlation with the state of economy.

It is reasonable to assume that the investor sentiment is positively correlated to the state of economy, during good times investors are more exposed to favourable news and events and the opposite holds during bad times. Accordingly, the last hypothesis would effectively test whether a similar relation between investor sentiment and value premia exists both during and outside recessions.

H3: The strength of the relationship between investor sentiment and value premium is similar during different economy states (economic expansion and recession states)

This papers evolves in an environment in which market prices can deviate from intrinsic prices and examine whether value premia shift with sentiment. There are currently no infallible valuation models for stocks, hence it is difficult to accurately estimate deviations from true prices. A similar problem exists with the measurement of investor sentiment. Nevertheless, the relation between investor sentiment and asset valuation exists empirically, starting with Brown and Cliff (2005), who showed with a long horizon regression that investor sentiment affects the stocks prices. The authors perceive sentiment as a persistent variable and indeed find that sentiment is strongly related to long-run stock returns. Among others, Hong and Stein (1999) showed that prices underreact in the short-run and overreact in the long run. Further evidence indicates that stocks returns are more sensitive to optimism than pessimism (Ding, Charoenwong, & Seetoh, 2004); (Zhang & Semmler, 2009).

The most established and applied asset pricing model summarizing factors driving stocks excess returns so far is Fama and French's (1992) three factor model. The factors used in the model are frequently used as control variables in research attempting to explain cross-sectional returns or measuring the impact of an additional variable on a certain dependent variable (see (Brown & Cliff, 2004); (Baker & Wurgler, 2006)).

Fama and French's (1992) initial three factor model establishes that the expected excess return of a portfolio can be explained by two additional factors to the CAPM of Sharpe (1964) such that excess return E(Ri) – Rf is explained by factors loadings to: (i) the excess return on a well-diversified market portfolio, (ii) the small minus big factor, which is the difference in returns between a portfolio formed with small stocks and a portfolio formed with big stocks, and (iii) the high minus low factor, which reflect the return gap between portfolios of stocks of high-book-to market and those of stocks of low-book-to market. At inception, the model was a great improvement of Sharpe's (1964) single factor asset-pricing model linking average returns and risk. Fama and French (1992) found that both size and book-to-market increase the cross-sectional variation in average stock returns captured by the model. They also suggested that the value premium can be seen as a compensation due to greater risks carried by value stocks compared to growth stocks. An interesting feature of the three-factor model is its ability to capture the reversal of long-term returns as identified by Bondt and Thaler (1985). Stocks with a low past returns history lean towards positive SMB and HML slopes, whereas stocks with an history of high past returns tend to have negative slopes on HML (Fama & French, 1996). Yet, that three factor model was not bullet proof as it did not capture certain anomalies as we will discuss later.

Following Carhart's (1997) pioneer work, the model is often extended with a fourth factor, the 1-year momentum returns differential on past winners and past losers. Compared to the CAPM and three factor model, this four-factor model performs well in reducing the average pricing errors (Carhart, 1997).

Pástor and Stambaugh's (2003) liquidity risk factor is another common additional factor. Pástor and Stambaugh (2003) found that some securities are more sensitive to marketwide liquidity, and this risk is priced by the market as investors favour securities that are unlikely to require liquidity during low liquidity periods. The liquidity risk is measured via the magnitude of returns reversals following order flows (volume). Accordingly, securities with greater exposure to aggregate liquidity should expect higher returns. A sort on firm size indicated that smaller firms tend to have higher liquidity betas, which measure the sensitivity to aggregate liquidity (Pástor & Stambaugh, 2003). A higher liquidity beta indicates a stronger illiquidity.

Critics however proceeded to showcase anomalies unexplained by the three-factor model, namely the apparent link between expected earnings and investment. Accordingly, Fama and French (2015) decomposed the market capitalization and book-to-equity ratios and showed that these factors are in fact noisy proxies for expected earnings and investments. With advances in proxies for expected profitability and investment that are fairly related to average return, Fama and French (2015) enhanced their three-factor model with the profitability and investment factors, leading to a five-factor model. Profitability is measured as operating profitability minus interest expenses and investment is proxied by the growth in total assets divided by total assets (Fama & French, 2015). For different sets of portfolios return, the five-factor model performed better (yielded a lower absolute intercept) than the three factor model, reaching explanatory powers between 71% and 94%. Despite the improvements, the model is easily rejected by Gibbons, Ross, and Shanken's (1989) F-test, indicating that a significant portion of average return is still left unexplained by the five-factor model (Fama & French, 2015).

Fama and French (2007) found a convergence in B/Ms of value and growth firms, supposedly triggered by a mean reversion in profitability and expected returns. Namely, the B/Ms of value portfolios diminish as their profitability increase and the B/Ms of growth portfolios enlarge as growth firms do not meet the expected profitability level. While the significance of the sentiment coefficient in Baker and Wurgler's (2006) research remained, the coefficient decreased after controlling the regression for market excess return, SMB, HML, and the momentum factor. Hence, adding the profitability, momentum and liquidity factors as control variables may further alter the effect of investor sentiment on stock returns noted in the past literature.

III. Data and Variables

All the data are mutually available from January 1968 until September 2015, providing 573 monthly observations for each variable presented in the next subsections.

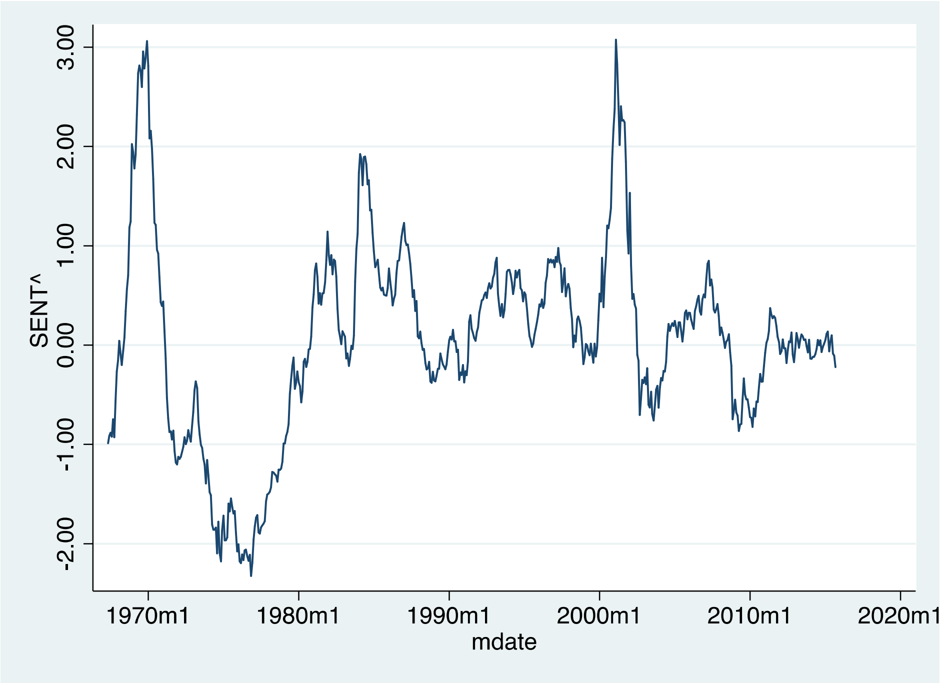
A. Investor sentiment

The research will employ the top down approach to investor sentiment, which focuses on tracing the effects of an aggregated form of sentiment on the markets. It considers the investor sentiment as an exogenous phenomenon and analyses its empirical effects. The composite index of sentiment used is that of Baker and Wurgler (2006) and was initially formed by combining six sentiment proxies: the share turnover from the New York Stock Exchange (NYSE), the close-end fund discount, the number and average first-day return on IPOS, the dividend premium and the share of equity in new issues. In the latest update for the composite index of sentiment, the NYSE share turnover has been dropped out since its link with investor sentiment significantly weakened. The index takes into account the relative timing of different variables, such that some proxies may signal a change in sentiment sooner. This can be the case of IPO volume lagging on high first-day returns for instance, or more generally, proxies based on investor supply reaction lagging behind proxies involving investor demand (Baker & Wurgler, 2006). Accordingly, SENTIMENT is defined by Baker and Wurgler (2016) as a combination of five variables, with each variable being the proxy's respective lead or lag that has the highest correlation with a pre-build first-stage index consisting of current and lagged estimates of each proxy.

Additionally, in order to remove any bias due to a common business cycle component Baker and Wurgler (2006) formed a second sentiment index, that excludes business cycle variation by orthogonalizing all five proxies in macroeconomics variables. Correlation tests between SENTIMENT and SENTIMENT⊥ show that the attempt to remove business cycle variation did not have a qualitative impact on neither any component nor the overall index. However, for future empirical test purposes the SENTIMENT⊥ constructed with cleaner proxies is preferred and will be used as investor sentiment index. Since all components of the index are standardized, the mean equals zero and one-unit increase in the index corresponds to a one standard deviation (SD) increase.

Graph 1 below plots the U.S. sentiment index level across time and suggests that the composite index is a fairly accurate indicator. Indeed, sentiment level were considerably higher than average before major crashes such as in the 1970s and the dot-com bubble crash beginning 2000s.

Graph 1: Time plot SENTIMENT⊥



B. Value premia

Academic research extensively researched different identification procedures of values and growth investments and evidence shows that the best selection criteria to analyze the value premia varies overtime (Pätäri & Leivo, 2015). Individual valuation ratios such as the earnings yield (E/P), the dividend yield (D/P), the book-to-market (B/M), the sales-to-price (S/P) and the cash flows-to price CF/P are quite popular value-based anomalies. Pätäri, Leivo, and Honkapuro (2010) recognized that value premium formed on composite value measures were more robust to changing market sentiment and pays off to the value investor. However, the evidence is relatively weak given that out of six studies, only that one study concluded that a composite value criteria yielded both higher returns and lower risks compared to the best individual valuation ratio alternative (Pätäri & Leivo, 2015). Currently, the B/M ratio is the most common measure to compute the value premium in literature (Pätäri & Leivo, 2015). Empirical research in the U.S. is well documented on the value premium as computed with the B/M ratio, with recent evidence provided by Israel and Moskowitz (2013). To allow for a smoother comparison with past research, the B/M valuation measure is employed in this paper as well.

A common concern in the value premium literature is the survivorship bias induced by the Compustat data. Kothari, Shanken, and Sloan (1995) claimed that the B/M return anomaly is largely exaggerated by a survivor bias. Yet, Lakonishok, Shleifer, and Vishny (1994) addressed this bias by adjusting their sample selection methodology to require at least five years of past data to arrange stocks returns into clusters and the value premium persisted. Their findings were further supported by Chan, Jegadeesh, and Lakonishok (1995), who stated that selection bias in Compustat is a minor issue for the value premium after using a sample free from survivorship bias. The survivorship bias claim will therefore be disregarded in this research.

Portfolio returns' data are obtained from Kenneth French's website. This data is convenient as it is clean with regards to companies with negative B/E, and B/E is computed according to Davis, Fama, and French (2000). Stocks are independently sorted on B/M of December of year t-1 among two sub groups Big (Market Capitalization(ME)>median) stocks and Small (ME< median) stocks. Three B/M groups (high, medium and low) are formed according to NYSE breakpoints.

Given the definition of value stocks and growths stocks, each month long–short portfolios are formed and the value premium is computed as follows:

$$Value\ premium = R_{Ht} - R_{Lt}$$

Where R_{Ht} corresponds to the average return on a portfolio composed of value (high book-to-market ratio) stocks

And R_{Lt} is the average return at time t on a portfolio composed of growth (low book-to-market ratio) stocks

The value premium is obtained in four ways to test the hypotheses at multiple levels. More specifically, the returns on portfolio are either equal-weighted or value weighted returns, and a high book-to-market ratio is defined as 70% or 90% percentile of NYSE stocks whereas low book-to-market corresponds to first 30% or 10% NYSE stocks.

Another key step in data modelling consist of classifying firms into subgroups such that the relation between investor sentiment and value premium can be analysed across different groups of firms. Motivated by the limits to arbitrage literature, two groups of stocks are formed: high limits-on-arbitrage (HLA) and low limits-on-arbitrage (LLA). Limits to arbitrage arise due to transaction and holding costs (Pontiff, 2006). As smaller stocks tend to be more difficult to evaluate and trade due to opaqueness and illiquidity, to study the differential impact of sentiment index on the value premium with respect to difficulty to arbitrage, the Fama and French' portfolios sorted two-ways on size and B/M are used. These portfolios permit the value premium to be assessed within two subsamples: high limits-on-arbitrage (small stocks) and low limits-on-arbitrage (large stocks). The portfolios are formed every June of year t such that B is book equity at the end of the previous fiscal year M is market cap at the end of December of the previous calendar year (2015). B/Ms are matched to monthly returns from July t to June t+1. For brevity and simplicity, the value premium for these subgroups corresponds to either equal-weighted or value weighted returns but uniquely for the High30 minus Low30 version.

C. Control variables

To identify sentiment-driven changes on value premium, the five empirically motivated factors of Fama and French (2015) are considered to reduce the risk of an omitted variable bias. The five-factor model was designed to explain the relation between average return and excess return to the market portfolio (MktRf), size, book-to-market (B/M), profitability (OP) and investment (Inv). Size, B/M, profitability, and investment factors are respectively labelled as

SMB (small minus big), HML (high minus low B/M), RMW (robust minus weak OP), and CMA (conservative minus aggressive Inv) (Fama & French, 2015).

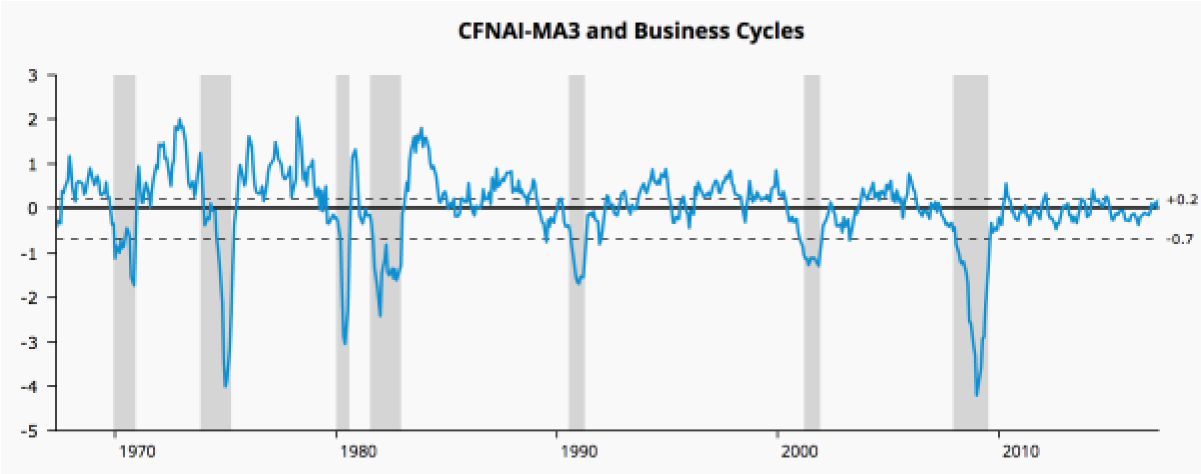
These factors were obtained through 2x3 sorts on stocks available on both CRSP and Compustat. The sorts are based on Size and B/M, or Size and OP, or Size and Inv and the intersection of these groups are used to construct the factors, making these portfolios roughly neutral in terms of size effect. Four sorts 2x2x2x2 portfolios to jointly control for size, B/M, OP and Inv were also considered to neutralize correlation effects and better isolate the premiums. Yet, as multivariate regression slopes focus on measuring the factors' marginal effects, the 2x3 sorts produce more diversified portfolios to isolate premiums and perform as well as 2x2x2x2 factors (Fama & French, 2015).

Moreover, the momentum factor (Mom) and liquidity factor (LIQ) are added to capture Jegadeesh and Titman's (1993) one-year momentum anomaly and Pástor and Stambaugh's (2003) liquidity risk premium. In Fama and French's (2015) models, the momentum and liquidity factor resulted to insignificant changes to the model's performance for portfolios examined, with regression slopes close to zero. Nevertheless, for robustness the momentum and liquidity factors will be included in the list of controlling parameters.

Fama and French (2017) computed Mom in similar ways as previous factors, through the intersection of a 2x3 sorts based on size and 1-year prior return for all NYSE, AMEX, and NASDAQ stocks with sufficient information. Although the portfolios contain all stocks jointly listed on CRSP and Compustat, to avoid a bias due to an overpopulation of small caps stocks listed on NASDAQ and AMEX stock exchanges, the breakpoints used are the NYSE median for size and 30th and 70th NYSE percentiles for the second order sorts. The resulting factors are public and obtained directly from French's online data library. The liquidity risk premium corresponds to the spread in returns between extreme decile portfolios formed on predicted liquidity betas. The resulting liquidity factor is obtained from Pástor's online data library.

Similar to Du (2011), the Chicago Fed National Activity Index (CFNAI-MA3) will be used as a measure of the economy's state, and proxy periods of falling GDP. The CFNAI-MA3 comprises four broad categories: (1) Production & Income, (2) Employment, Unemployment & Hours, (3) Personal Consumption & Housing and (4) Sales, Orders & Inventories (Federal Reserve Bank of Chicago, 2017) and totalizes 85 macroeconomic variables. The resulting composite is considered more complete and more effective in capturing economic activity than the NBER business cycle dates (Du, 2011). Also, since the CFNAI-MA3 is realized monthly, it represents a more real-time measure of aggregate economic conditions compared to the lagged NBER indicator that usually officialise turning points only after a few months' delay.

Graph 2: National Activity Index over time and NBER business cycles



Source: Federal Reserve Bank of Chicago, 2017

The above plot shows that the CFNAI (blue line) effectively captures recession periods as identified by the NBER recessions indicated by the shading. The index is constructed to have an average value of zero and a standard deviation of one, such that positive values correspond to a growth above trend and negative values to a growth below trend.

D. Summary descriptive statistics

To reveal patterns and responses to investor sentiment in value premium during different regimes, the SENTIMENT⊥ and CFNAI time series have been split into two states using a Markov switch model (see Appendix B. Graphs B1 and B2). SENTIMENT⊥ and CFNAI are assumed to follow a first-order Markov process able to switch between states multiple times throughout the sample. According to the model, there exist a low sentiment state which corresponds to a period in which the sentiment index has a mean of -1,1 and a recession state in which the CFNAI has a mean of -1,65; within states means are denoted by Constant (see Appendix B. Tables B1 and B3). These two states characteristics have been approximated for both SENTIMENT⊥ and CFNAI with two dummies, respectively Low and Recession, by replicating similar mean in both states (see Appendix B. Tables B2 and B4). The investor sentiment is considered low when SENTIMENT⊥ < -0,3 and the economy is in recession when CFNAI<-0.7. A CFNAI below -0,70 also corresponds to a recession period as historically associated by the Federal Reserve Bank of Chicago (FRBC). Tables below summarize the dependent variable, value premia, under all forms and across regimes. “EW” represents equally weighted, and “VW” value weighted.

Table 1: Average monthly equal and value weighted value premia per investor sentiment level

	EW High30-Low30		EW High10-Low10		VW High30-Low30		VW High10-Low10		
Low	Mean	Std. Dev	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Freq.
0	0,784	3,074	1,130	4,090	0,214	2,822	0,356	4,216	416
1	0,669	3,072	1,130	4,520	0,514	3,603	0,789	6,001	157
Total	0,752	3,071	1,130	4,210	0,296	3,056	0,475	4,772	573

Table 2: Subgroups average monthly equal and value weighted value premia per investor sentiment level

	EW HLA		EW LLA		VW HLA		VW LLA		
Low	Mean	Std. Dev	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Freq.
0	0,902	3,551	0,165	3,120	0,577	3,434	0,128	2,981	416
1	0,519	2,984	0,564	3,097	0,431	2,943	0,387	3,601	157
Total	0,797	3,407	0,274	3,116	0,537	3,305	0,199	3,162	573

At first sight, the equally weighted returns are constantly higher than value weighted returns regardless of group. Given the dichotomy of the dummy Low, periods of high sentiment level are expected to be followed by periods of low sentiment level and vice versa. Table 1 above shows that the value premium tends to be lower in the periods following a high sentiment. Accordingly, the value premium appears higher following low sentiment periods and that holds for the HLA subgroup. The LLA subgroup follows a reverse trend, namely the value premium is higher during low sentiment periods compared to the second state (See Table 2 above).

The tables below provide an overview of value premia trends during different economic states.

Table 3: Average monthly equal and value weighted value premia per economic state

	EW High30-Low30		EW High10-Low10		VW High30-Low30		VW High10-Low10		
Recession	Mean	Std. Dev	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Freq.
0	0,760	2,839	1,120	3,810	0,293	2,691	0,452	4,341	492
1	0,707	4,238	1,180	6,120	0,317	4,727	0,611	6,864	81
Total	0,752	3,071	1,130	4,210	0,296	3,056	0,475	4,772	573

Table 4: Subgroups average monthly equal and value weighted value premia per economic state

Recession	EW HLA		EW LLA		VW HLA		VW LLA		Freq.
	Mean	Std. Dev	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	
0	0,818	3,256	0,257	2,840	0,499	3,163	0,208	2,779	492
1	0,667	4,228	0,376	4,462	0,767	4,076	0,143	4,909	81
Total	0,797	4,407	0,274	3,116	0,537	3,305	0,199	3,162	573

Overall, the value premium appears larger following a recession period. However, across subgroups it is unclear what the trend is, as Table 4 indicates a mixture of slightly larger following a recession period for some categories and higher during the computed recession period for others.

Table 5: Summary statistics regressors

Variable	Observations	Mean	Std. Dev.	Min	Max
SENT	573	0,068	0,974	-2,325	3,076
$\Delta SENT$	572	-0,000	0,171	-0,730	0,778
CFNAI	573	0,000	0,889	-4,207	2,036
MktRF	573	0,473	4,568	-23,240	16,100
SMB	573	0,189	3,072	-15,280	18,730
Mom	573	0,680	4,358	-34,580	18,380
RMW	573	0,262	2,311	-19,110	13,520
CMA	573	0,352	2,028	-6,880	9,550
LIQ	573	0,422	3,382	-12,489	11,078

Globally, $SENTIMENT^{\perp}$ (SENT) and CFNAI both display a mean close to zero as well as a standard deviation close to 1 as expected and all factors defined previously are positive on average. Next, the regression approach requires not only stationary time series data, but also exogenous explanatory variables. These conditions are analysed through tests and data analyses presented below. The sentiment index seems auto-correlated as shown on its correlogram (see Appendix A.D. Graph 1). Accordingly, it has been detrended by the first difference method such that:

$$\Delta SENTIMENT^{\perp} = SENTIMENT^{\perp}{}_{t} - SENTIMENT^{\perp}{}_{t-1}$$

Subsequently, all the relevant regressors are stationary according to the Augmented Dickey–Fuller unit root test at 1% critical value. Another assumption of the Ordinary Least Square

(OLS) regression method is no correlation in residuals, and the Durbin-Watson Statistics on Stata hint on no evident serial correlation on regressors used.

IV. Methodology

The aim of this research is to estimate a controlled relationship between investor sentiment and the value premium. Accordingly, formal tests will be performed using the following multivariate regression:

$$R_{Ht} - R_{Lt} = \alpha_i + b_i(MktRF_t) + s_iSMB_t + r_iRMW_t + c_iCMA_t + m_iMOM_t + \ell_iL_t + \gamma_i\Delta SENTIMENT^{\perp} + e_{it} \quad (3)$$

Where $MktRF_t$ is the excess returns of a value weighted market index over a U.S 1 month treasury bill, SMB_t is the return gap between small and large firms portfolios formed on market capitalization, RMW_t is the return gap between portfolios of stocks with robust and weak profitability, CMA_t is the return gap between diversified portfolios of stocks of low investment firms and stocks of high investments firms and $\Delta SENTIMENT^{\perp}$ is the first difference in composite sentiment index. During the tests, factors will be progressively added as to assess the marginal power brought by factors and their effect on adjacent regressors. As HML are the portfolios of interest, the HML factor is excluded from the equations' right side. It is assumed that e_{it} approximately follows the normal distribution.

In equation (3), the intercept α_i is zero if factor loadings $b_i, s_i, r_i, c_i, m_i, \ell_i$ and γ_i capture all the variation in the average value premium. This zero-intercept intercept hypothesis is principally supported by Huberman and Kandel's (1987) proposition that the mean-variance-efficient tangency portfolio, which is the foundation to all asset pricing, combines the market portfolio, the risk-free asset and all four other factors included in the model (Fama & French, 2015). Adding these factors permits to have a relation between investor sentiment and value premium that is distinct from other common factors' influence.

Given that asset valuation implies that expected returns be determined by the combination of current prices and expectations of future dividends, the expected returns should be the same for all horizons (Fama & French, 2015). Most asset pricing researches are therefore conducted on short-horizon returns. Medium and long term predictability of investor sentiment has been confirmed by Brown and Cliff (2005). The effect of sentiment on stock returns and subsequently value premium is observed on a monthly basis. Standard errors are robust to avoid heteroskedasticity and prevent a possible bias due to autocorrelated sentiment index and portfolio returns innovations as documented by Stambaugh (1999) (Baker & Wurgler, 2006).

Another problem with the regression approach is that causality can run two ways, and sentiment levels are unlikely to be distinct from recent movements in stock prices. However, since the value premia rather individual stock returns are considered as dependent variable, the potential bias is tampered down. A causal relationship between value premium and sentiment is expected to be weak.

Lastly, equation (3) can take the general form:

$$R_{Ht} - R_{Lt} = \begin{cases} \alpha_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + r_i RMW_t + c_i CMA_t + m_i MOM_t + \gamma_i \Delta SENTIMENT^{\perp} + \ell_i L_t + e_{it} & \text{if Recession} = 0 \\ \alpha'_i + b'_i(R_{Mt} - R_{Ft}) + s'_i SMB_t + r'_i RMW_t + c'_i CMA_t + m'_i MOM_t + \gamma'_i \Delta SENTIMENT^{\perp} + \ell_i L'_t + e'_{it} & \text{if Recession} = 1 \end{cases} \quad (4)$$

such that state-specific parameters of sentiment can be predicted by using a threshold model. With equation (4) it is possible to observe whether there is a differential response of value premium to investor sentiment during adverse aggregate shocks. That is, whether coefficients γ_i and γ'_i differ significantly.

V. Results and Discussion

The regressions performed and presented next have a common goal to formally test the hypotheses. As the value premium is the difference in returns between value and growth stocks portfolios, any factor presenting a significant coefficient will be interpreted to exhibit a differential association to value and growth stocks, such that the level of value premium fluctuates with changes in that factor. More generally, the coefficients can be seen as ceteris paribus tendencies of the factor on the value premium. Significant coefficients (at 5% significance level on a two-sided test) are in bold.

Tables 6a & 6b: Models summary on equal-weighted value premia

Table 6a summarizes five OLS regressions on High30-Low30 equal-weighted value premia and Table 6b summarizes five OLS regressions on High10-Low10 equal-weighted value premia. MktRF represents the excess returns of a value weighted market index over a U.S 1-month treasury bill, SMB the return gap between small and large firms' portfolios formed on market capitalization, RMW the return gap between portfolios of stocks with robust and weak profitability, CMA the return gap between diversified portfolios of stocks of low investment firms and stocks of high investments firms and ΔSENT, the first difference in composite sentiment index. The factors are progressively added in order to assess their marginal power and their effect on adjacent regressors. In Basemodel, the value premium is regressed solely on MktRF. In CAPM, the value premium is regressed on MktRF and $\Delta SENT$. In ThreeFactor, the value premium is regressed on MktRF, SMB, and $\Delta SENT$. In FiveFactor, the value premium is regressed on MktRF, SMB, RMW, CMA and $\Delta SENT$. In FiveFactorMomLIQ, the value premium is regressed on MktRF, SMB, CMA, Mom, LIQ, and $\Delta SENT$. Bolded coefficients are significant at 5% level.

Table 6a: Models summary on equal-weighted High30-Low30 value premia

Variable	Basemodel	CAPM	ThreeFactor	FiveFactor	FiveFactorMomLIQ
MktRF	**-0,302**	**-0,296**	**-0,299**	**-0,118**	**-0,122**
ΔSENT		**2,042**	**2,014**	0,910	0,884
SMB			0,022	0,086	0,088
RMW				**0,287**	**0,291**
CMA				**0,911**	**0,909**
Mom					-0,031
LIQ					0,019
Constant	**0,895**	**0,883**	**0,881**	**0,393**	**0,407**
N	573	572	572	572	572
R-squared	0,202	0,213	0,214	0,526	0,528
Adj. R-squared	0,200	0,211	0,210	0,522	0,523

Table 6b: Models summary on equal-weighted High10-Low10 value premia

Variable	Basemodel	CAPM	ThreeFactor	FiveFactor	FiveFactorMomLIQ
MktRF	**-0,349**	**-0,340**	**-0,359**	**-0,110**	**-0,124**
ΔSENT		**3,193**	**3,056**	1,547	1,449
SMB			0,106	**0,187**	**0,192**
RMW				**0,362**	**0,378**
CMA				**1,258**	**1,251**
Mom					**-0,106**
LIQ					0,0722
Constant	**1,298**	**1,282**	**1,272**	**0,610**	**0,656**
N	573	572	572	572	572
R-squared	0,143	0,159	0,164	0,477	0,493
Adj. R-squared	0,142	0,156	0,159	0,473	0,487

Tables 7a & 7b: Models summary on value weighted value premia

Table 7a summarizes five OLS regressions on High30-Low30 value weighted value premia and Table 7b summarizes five OLS regressions on High10-Low10 value weighted value premia. MktRF represents the excess returns of a value weighted market index over a U.S 1-month treasury bill, SMB the return gap between small and large firms' portfolios formed on market capitalization, RMW the return gap between portfolios of stocks with robust and weak profitability, CMA the return gap between diversified portfolios of stocks of low investment firms and stocks of high investments firms and ΔSENT, the first difference in composite sentiment index. The factors are progressively added in order to assess their marginal power and their effect on adjacent regressors. In Basemodel, the value premium is regressed solely on MktRF. In CAPM, the value premium is regressed on MktRF and $\Delta SENT$. In ThreeFactor, the value premium is regressed on MktRF, SMB, and $\Delta SENT$. In FiveFactor, the value premium is regressed on MktRF, SMB, RMW, CMA and $\Delta SENT$. In FiveFactorMomLIQ, the value premium is regressed on MktRF, SMB, CMA, Mom, LIQ, and $\Delta SENT$. Bolded coefficients are significant at 5% level.

Table 7a: Models summary on value weighted High30-Low30 value premia

Variable	Basemodel	CAPM	ThreeFactor	FiveFactor	FiveFactorMomLIQ
MktRF	-0,067	-0,594	**-0,108**	0,054	0,035
ΔSENT		**2,701**	**2,357**	1,384	1,324
SMB			**0,265**	**0,237**	**0,242**
RMW				-0,086	-0,064
CMA				**0,961**	**0,951**
Mom					**-0,144**
LIQ					0,019
Constant	**0,328**	**0,315**	**0,290**	-0,088	0,009
N	573	572	572	572	572
R-squared	0,010	0,032	0,098	0,444	0,486
Adj. R-squared	0,008	0,029	0,093	0,439	0,480

Table 7b: Models summary on value weighted High10-Low10 value premia

Variable	Basemodel	CAPM	ThreeFactor	FiveFactor	FiveFactorMomLIQ
MktRF	0,525	0,064	-0,029	**0,205**	**0,174**
ΔSENT		**4,433**	**3,770**	**2,358**	**2,265**
SMB			**0,511**	**0,460**	**0,469**
RMW				-0,165	-0,127
CMA				**1,412**	**1,395**
Mom					**-0,241**
LIQ					0,026
Constant	**0,450**	**0,431**	**0,383**	-0,159	0,006
N	573	572	572	572	572
R-squared	0,003	0,028	0,128	0,439	0,487
Adj. R-squared	0,001	0,025	0,124	0,434	0,481

Table 8a & 8b: Models summary comparison across HLA and LLA for equal-weighted High30-Low30 value premia

Table 8a compares three OLS regressions on High30-Low30 equal-weighted value premia across high limits-on-arbitrage (HLA) and low limits-on-arbitrage (LLA) group and Table 8b compares two OLS regressions on High30-Low30 equal-weighted value premia across HLA and LLA groups. MktRF represents the excess returns of a value weighted market index over a U.S 1-month treasury bill, SMB the return gap between small and large firms' portfolios formed on market capitalization, RMW the return gap between portfolios of stocks with robust and weak profitability, CMA the return gap between diversified portfolios of stocks of low investment firms and stocks of high investments firms and ΔSENT, the first difference in composite sentiment index. All independent variables are as of time t and are progressively added in order to assess their marginal power and their effect on adjacent regressors. In Basemodel, the value premium is regressed solely on MktRF. In CAPM, the value premium is regressed on MktRF and $\Delta SENT$. In ThreeFactor, the value premium is regressed on MktRF, SMB, and $\Delta SENT$. In FiveFactor, the value premium is regressed on MktRF, SMB, RMW, CMA and $\Delta SENT$. In FiveFactorMomLIQ, the value premium is regressed on MktRF, SMB, CMA, Mom, LIQ, and $\Delta SENT$. Bolded coefficients are significant at 5% level

Table 8a: Models summary comparison Basemodel, CAPM and ThreeFactor regressed on HLA and LLA equal-weighted High30-Low30 value premia

Variable	HLABasemodel	LLABasemodel	HLACAPM	LLACAPM	HLAThreeFactor	LLAThreeFactor
MktRF	**-0,355**	**-0,183**	**-0,351**	**-0,175**	**-0,316**	**-0,169**
ΔSENT			2,077	**2,643**	1,585	**3,164**
SMB					**-0,194**	-0,029
Constant	**0,965**	**0,368**	**0,959**	**0,349**	**0,978**	**0,352**
N	573	572	572	572	572	572
R-squared	0,226	0,072	0,230	0,099	0,258	0,101
Adj. R-squared	0,225	0,070	0,227	0,097	0,254	0,096

Table 8b: Models summary comparison FiveFactor and FiveFactorMOMLIQ on HLA and LLA equal-weighted High30-Low30 value premia

Variable	HLAFiveFactor	LLAFiveFactor	HLAFiveFactorMomLIQ	LLAFiveFactorMomLIQ
MktRF	-0,116	0,023	-0,117	0,007
ΔSENT	0,374	**2,002**	0,367	**1,971**
SMB	**-0,092**	-0,003	**-0,092**	0,000
RMW	**0,438**	0,133	**0,439**	**0,152**
CMA	**0,941**	**1,036**	**0,941**	**1,028**
Mom			-0,001	**-0,119**
LIQ			0,008	-0,003
Constant	**0,424**	-0,135	**0,421**	-0,048
N	573	572	572	572
R-squared	0,557	0,470	0,557	0,497
Adj. R-squared	0,553	0,466	0,551	0,491

Tables 9a & 9b: Models summary comparison across HLA and LLA for value weighted High30-Low30 value premia

Table 9a compares three OLS regressions on High30-Low30 value weighted value premia across high limits-on-arbitrage (HLA) and low limits-on-arbitrage (LLA) group and Table 9b compares two OLS regressions on High30-Low30 value weighted value premia across HLA and LLA groups. MktRF represents the excess returns of a value weighted market index over a U.S 1-month treasury bill, SMB the return gap between small and large firms' portfolios formed on market capitalization, RMW the return gap between portfolios of stocks with robust and weak profitability, CMA the return gap between diversified portfolios of stocks of low investment firms and stocks of high investments firms and ΔSENT, the first difference in composite sentiment index. All independent variables are as of time t and are progressively added in order to assess their marginal power and their effect on adjacent regressors. In Basemodel, the value premium is regressed solely on MktRF. In CAPM, the value premium is regressed on MktRF and $\Delta SENT$. In ThreeFactor, the value premium is regressed on MktRF, SMB, and $\Delta SENT$. In FiveFactor, the value premium is regressed on MktRF, SMB, RMW, CMA and $\Delta SENT$. In FiveFactorMomLIQ, the value premium is regressed on MktRF, SMB, CMA, Mom, LIQ, and $\Delta SENT$. Bolded coefficients are significant at 5% level.

Table 9a: Models summary comparison Basemodel, CAPM and ThreeFactor regressed on HLA and LLA value weighted High30-Low30 value premia

Variable	HLABasemodel	LLABasemodel	HLACAPM	LLACAPM	HLAThreeFactor	LLAThreeFactor
MktRF	**-0,297**	-0,061	**-0,292**	-0,054	**-0,255**	-0,078
ΔSENT			**2,077**	**2,643**	**2,343**	**2,472**
SMB					**-0,206**	**0,132**
Constant	**0,678**	0,227	**0,670**	0,217	**0,690**	0,205
N	573	572	572	572	572	572
R-squared	0,169	0,008	0,179	0,028	0,213	0,043
Adj. R-squared	0,168	0,006	0,177	0,024	0,209	0,037

Table 9b: Models summary comparison FiveFactor and FiveFactorMOMLIQ on HLA and LLA value weighted High30-Low30 value premia

Variable	HLAFiveFactor	LLAFiveFactor	HLAFiveFactorMomLIQ	LLAFiveFactorMomLIQ
MktRF	-0,039	**0,084**	-0,051	0,063
ΔSENT	1,037	1,492	0,977	1,442
SMB	**-0,126**	0,107	**-0,122**	**0,113**
RMW	**0,354**	-0,073	**0,369**	-0,047
CMA	**1,071**	**0,096**	**1,065**	**0,949**
Mom			**-0,097**	**-0,169**
LIQ			0,035	0,001
Constant	0,110	-0,177	0,166	-0,055
N	573	572	572	572
R-squared	0,588	0,363	0,605	0,417
Adj. R-squared	0,584	0,358	0,600	0,409

Tables 10a & 10b: Models summary equal-weighted value premia by economy state

Table 10a summarizes three OLS regressions on High30-Low30 equal-weighted value premia and Table 10b summarizes three OLS regressions on High10-Low10 equal-weighted value premia. MktRF represents the excess returns of a value weighted market index over a U.S 1-month treasury bill, SMB the return gap between small and large firms' portfolios formed on market capitalization, RMW the return gap between portfolios of stocks with robust and weak profitability, CMA the return gap between diversified portfolios of stocks of low investment firms and stocks of high investments firms, ΔSENT the first difference in composite sentiment index and Recession, a dummy variable capturing when the economy is in recession as determined with the CFNAI. In Recession0, the regression uses the sample outside recession periods. In Recession1, the regression uses the sample during recession periods. In Recession Control, the value premium is regressed on MktRF, $\Delta SENT$, SMB, RMW, CMA, Mom, LIQ and Recession to control for a joint hypothesis bias. Bolded coefficients are significant at 5% level.\

Table 10a: Models summary High30-Low30 equal-weighted value premia by economy state

Variable	Recession0	Recession1	Recession Control
MktRF	**-0,159**	-0,030	**-0,121**
ΔSENT	0,236	1,688	0,759
SMB	0,447	0,185	**0,092**
RMW	**0,274**	**0,395**	**0,299**
CMA	**0,880**	**1,232**	**0,917**
Mom	0,033	**-0,167**	-0,033
LIQ	0,005	0,045	0,017
Recession			-0,509
Constant	**0,477**	-0,301	**0,476**
N	491	81	572
R-squared	0,534	0,653	0,532
Adj. R-squared	0,527	0,619	0,525

Table 10b: Models summary High10-Low10 equal-weighted value premia by economy state

Variable	Recession0	Recession1	Recession Control
MktRF	**-0,171**	-0,012	**-0,123**
ΔSENT	0,392	2,780	1,306
SMB	0,114	0,339	**0,197**
RMW	**0,337**	**1,692**	**1,260**
CMA	**1,211**	**1,692**	**1,260**
Mom	-0,016	**-0,315**	**-0,108**
LIQ	0,022	0,183	0,070
Recession			-0,583
Constant	**0,754**	0,244	**0,735**
N	491	81	572
R-squared	0,501	0,620	0,495
Adj. R-squared	0,494	0,584	0,488

Tables 11a & 11b: Models summary value weighted value premia by economy state

Table 11a summarizes three OLS regressions on High30-Low30 value weighted value premia and Table 11b summarizes three OLS regressions on High10-Low10 value weighted value premia. MktRF represents the excess returns of a value weighted market index over a U.S 1-month treasury bill, SMB the return gap between small and large firms' portfolios formed on market capitalization, RMW the return gap between portfolios of stocks with robust and weak profitability, CMA the return gap between diversified portfolios of stocks of low investment firms and stocks of high investments firms, ΔSENT the first difference in composite sentiment index and Recession, a dummy variable capturing when the economy is in recession as determined with the CFNAI. In Recession0, the regression uses the sample outside recession periods. In Recession1, the regression uses the sample during recession periods. In Recession Control, the value premium is regressed on MktRF, $\Delta SENT$, SMB, RMW, CMA, Mom, LIQ and Recession to control for a joint hypothesis bias. Bolded coefficients are significant at 5% level.

Table 11a: Models summary High30-Low30 value weighted value premia by economy state

Variable	Recession0	Recession1	Recession Control
MktRF	0,007	0,106	0,035
ΔSENT	1,030	1,089	1,228
SMB	**0,187**	**0,442**	**0,246**
RMW	**-0,073**	-0,039	-0,058
CMA	**0,898**	**1,321**	**0,957**
Mom	**-0,088**	**-0,246**	**-0,146**
LIQ	0,015	0,008	0,018
Recession			-0,395
Constant	0,063	-0,686	0,062
N	491	81	572
R-squared	0,477	0,613	0,488
Adj. R-squared	0,469	0,576	0,481

Table 11b: Models summary High10-Low10 value weighted value premia by economy state

Variable	Recession0	Recession1	Recession Control
MktRF	**0,143**	**0,242**	**0,174**
ΔSENT	**2,358**	0,262	**2,151**
SMB	**0,370**	**0,826**	**0,473**
RMW	-0,158	0,014	-0,121
CMA	**1,325**	**1,910**	**1,403**
Mom	**-0,154**	**-0,412**	**-0,243**
LIQ	0,004	0,046	0,024
Recession			-0,462
Constant	0,068	-1,047	0,068
N	491	81	572
R-squared	0,440	0,718	0,488
Adj. R-squared	0,432	0,691	0,481

Table 6a indicates that adding a sentiment factor to the MktRF regressor does not particularly improve neither the intercept, which represents the mean value premium if all the factor premia were to be zero, nor the R-squared, which represents the variance in value premium explained by the regressors. The first difference sentiment index coefficient (2,042) nevertheless appears significant and suggests that the equally weighted value premium raises in the month following an increase in investor sentiment. Yet, the investor sentiment's coefficient significantly drops when expanding the right hand side of the equation (See Table 6a and 6b). More specifically, controlling for additional factors such as profitability (RMW) and investment (CMA) renders the coefficient insignificant.

Tables 7a and 7b display the same results as Table 6a and 6b but with value weighted instead of equal-weighted returns value premia. The main difference is that after value weighting returns in the portfolios, the SMB factor is more often positive and significant, which is consistent with a value portfolio containing relatively smaller stocks (in terms of market capitalization). Additionally, the significance of SMB's coefficient arising at the expense of the RMW factor may be an indicator that firms in the lower spectrum of market capitalization presumably tend to have a greater operating profitability measure. Also, the negative coefficient for momentum becomes significant when regressing on a price weighted value premium, evoking that growth portfolios contain more past winners on average than value portfolios.

Now to summarize and clarify the findings, on one hand, when the value premium is defined as equal-weighted returns, even if hypothetically all the known risk factors in the regression were not priced (hence all factors=0), the alpha (mean value premium) remains positive and significant for the most extended regression. This would imply that the value premium is at least partly derived from additional factors not reflected in my most complete model. On the other hand, once the value premium is taken as a value weighted return spread, momentum matters and the alpha is nearly zero and insignificant after including all the factors. In both cases (i.e. equally and value weighted), the investment factor always displays a significant positive coefficient. Perhaps as there is a higher proportion of firms with high investment rates in value stocks portfolios, a greater compensation for the investment factor will likely push the value premium upwards. The positive and significant coefficients for the investment factor (CMA) are also consistent with value firms following more conservative investment procedures with riskier assets in place, and due to the costly reversibility, these firms suffer more during economic down turns. Accordingly, higher returns on value stocks portfolios would be justified by this additional factor risk, as suggested by Zhang (2005). This effect may have gone unnoticed in previous researches controlling the right hand side only for the three-factor model. Furthermore, the liquidity factor remains insignificant for all specifications of value premium and apparently plays no role in explaining the value premium given its close to zero and insignificant regression slope. These results imply that value and glamour stocks have a similar exposure to liquidity risks.

The investor sentiment has a fully controlled significant positive relationship with subsequent value premia only on the High10-Low10 value weighted value premium (See Table 7b). In that case, the value premium is higher, more concentrated and there is thus more room for observable patterns. This outcome is consistent with Phalippou's (2008) evidence stipulating that most of the value premium is concentrated on 7% of the total stocks with relatively lower institutional ownership and that the spread in B/M declines with institutional ownership. It is thus not possible to readily reject the first hypothesis. The outputs depict a positive relationship between investor sentiment and value premium, and this relationship can be significant.

Comparing the relationship between investor sentiment and value premium on equally weighted returns across the HLA and LLA subgroups with Tables 8a and 8b above show that contrary to the overall value premium results, changes in investor sentiment have a strong positive and significant association with the future value premium in the LLA group throughout all the models with equally weighted High30-Low30 premium. By contrast, a change in sentiment remains positively related to the value premium but is mostly insignificant in the HLA group across all the regressions specifications. Unsurprisingly, controlling for SMB often matters only for the HLA group (small stocks).

Interestingly, the differential effect observed in Tables 8a and 8b disappears when the value premium is computed with value weighted stock returns portfolios instead of equally weighted ones (See Tables 9a and 9b). Indeed, all investor sentiment 's coefficients become insignificant at 5%. The second hypothesis is therefore rejected, there is not sufficient evidence supporting that investor sentiment has a stronger relationship with the value premium amongst high limits-on-arbitrage groups as defined by size. The coefficients of investor sentiment across subgroups are insignificant on value weighted value premia, and it is unclear whether sentiment has a significant relationship in the HLA group when the value premium is based on equal-weighted returns.

The third hypothesis predicted a similar relationship of investor sentiment with value premium during different states and tests conducted are presented in Tables 10a, 10b 11a and 11b. The full model (including all regressors in equation (3)) is regressed in different regimes and the recession factor is tested for significance to address the joint hypothesis problem that the results and association of sentiment might be driven by a state of economy risk compensation instead (Du, 2011).

Firstly, when defining value premium as High30 minus Low30 B/M portfolios returns (equal and value weighted) spread, the coefficients of investor sentiment's first difference in both states are insignificant at 5% significance level (See Tables 10a and 11a). Secondly, when the value premium is defined as value weighted return spread between High10-Low10, the first difference on monthly investor sentiment has a significant positive relationship with subsequent value premia only outside the recession state (see Table 11b). Similar tests have been performed on the HLA and LLA groups and it seems that the sentiment and value premium correlation previously observed among the LLA group for equal-weighted value premia remains significant only outside recession periods (See Appendix C, Table C1). Given the persistence of investor sentiment, the value premium seems quite unaffected by shifts in investor sentiment during worsening economic conditions, when sentiment tends to be low.

One possible inference is that investor sentiment is indeed positively correlated with economic aggregate conditions such that high sentiment generally prevails outside of recessions and monthly sentiment consequences are significant only during extended high sentiment periods. Hence, following low sentiment (and thus following months in recession) the relation is insignificant. This is consistent with Stambaugh, Yu and Yuan (2012) reporting that the value premium anomaly is stronger following high sentiment than following low sentiment. On the grounds of these results, the third hypothesis, stating that investor sentiment has a similar relationship with value premia across economic states, is rejected. Furthermore, it appears that changes in economic aggregate condition have no extra influence on the value premium. Results indicate that the coefficient for recession is not only insignificant, but also has no impact on others regressors. Hence, there is, although limited, statistical evidence for a positive relationship between the change in investor sentiment and next month's value premium.

A potential explanation to the observed positive coefficients would be that as the regressions are working with the monthly difference in investor sentiment, its coefficient is in fact reflecting a mispricing correction resulting from a shift in investor sentiment between the current and the preceding month. For instance, Chung, Hung, and Yeh (2012) regressed portfolios sorted on B/M ratios and found that the lagged monthly sentiment indicator has a positive association with the long-short portfolio returns. In this sentiment–based value premium world, as modelled by the regression, a plausible scenario would be that rising sentiment index levels elevate prices (and returns) of growth stocks, and this overpricing is corrected by the market in the following month, such that the value premium rises. The tendency of this mispricing correction to occur more frequently among LLA stocks (perhaps because arbitrage is more accessible) would explain the enhanced statistical significance of the sentiment index factor being specific to that group as displayed in Tables 8a and 8b above. Perhaps it takes longer for mispricing within the HLA group to adjust following increases in sentiment. Hence, the effect of investor sentiment is not fully captured in that subgroup for a monthly horizon.

VI. Conclusion

Previous analyses exhibit a positively correlated monthly changes of investor sentiment and future monthly value premia. However, the pattern between the two is complex; it varies according to how portfolios characteristics are defined to perform the tests and seems more evident among stocks with low limits-on-arbitrage. The results also suggest that besides the investor sentiment index, profitability, momentum, as well as investment's counter cyclicality risk factors have predictive power on the value premium. Moreover, the evidence signals a greater role for investor sentiment during economic prosperity.

This research attempted to analyze how the value premium relate to investor sentiment in a viable way using a regression approach. Identifying stock mispricing trends directly is a difficult task and I therefore settled on estimating the possible market correction in the subsequent period. This method is noisy as the trends measured are subject to biases due to other relevant factors' prices and value premia, hence results must be interpreted cautiously. Looking for systematic patterns in price corrections conditioned on beginning-of-period sentiment may be a viable alternative to the regression approach to assess the relationship between investor sentiment and value premium. This method would also give more insight into whether the direction of sentiment change matters for the predictive power. Additionally, using a longer horizon (e.g. annual time frame) might allow for these corrections to fully reflect in the model for the HLA group.

An extended research using another proxy for difficulty of arbitrage may shed light on additional patterns and solidify the current evidence. Indeed, while it is true smaller stocks are more opaque and illiquid and thus present higher limits-on-arbitrage, future work may explore whether stocks with greater illiquidity or idiosyncratic volatility respond better to investor sentiment shocks due to additional impediments to arbitrage.

Bibliography

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance*, 1645-1680.

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *The Journal of Economic Perspectives*, 129-151.

Black, A. J., & Fraser, P. (2003). The value premium: rational, irrational or random. *Managerial Finance*, 57-75.

Bondt, W. F., & Thaler, R. (1985). Does the stock market overreact? *The Journal of Finance*, 793-805.

Brav, A., Heaton, J. B., & Li, S. (2009). The limits of the limits of arbitrage. *Review of Finance*, 157-187.

Brown, G. W., & Cliff, M. T. (2005). Investor sentiment and asset valuation. *The Journal of Business*, 405-440.

Brown, G., & Cliff, M. T. (2004). Investor sentiment and the near-term stock market. *Journal of Empirical Finance*, 1-27.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 57-82.

Carlson, M., Fisher, A., & Giammarino, R. (2004). Corporate investment and asset price dynamics: implications for the cross-section of returns. *The Journal of Finance*, 2577-2603.

Chan, L. K., Jegadeesh, N., & Lakonishok, J. (1995). Evaluating the performance of value versus glamour stocks The impact of selection bias. 269-296.

Chung, S. L., Hung, H., C., & Yeh, C. Y. (2012). When does the investor sentiment predict stock returns? *Journal of Empirical Finance*, 217-240.

Cooper, I. (2006). Asset pricing implications of nonconvex adjustment costs and irreversibility of investment. *The Journal of Finance*, 139-170.

Daniel, K., & Titman, S. (1997). Evidence on the characteristics of cross sectional variation in stock returns. *The Journal of Finance*, 1-33.

Davis, J. L., Fama, E. F., & French, K. R. (2000). Characteristics, covariances and average returns: 1929 to 1997. *The Journal of Finance*, 389-406.

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance*, 379-395.

D'avolio, G. (2002). The market for borrowing stock. *Journal of Financial Economics*, 271-306.

Ding, D. K., Charoenwong, C., & Seetoh, R. (2004). Prospect theory, analyst forecasts, and stock returns. *Journal of Multinational Financial Management*, 425-442.

Du, D. (2011). Value premium and investor sentiment. *The Journal of the Academy of Behavioral Finance*, 87-101.

Fama, E. F., & French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, 55-84.

Fama, E. F., & French, K. R. (2007). Disagreement, tastes, and asset prices. *Journal of Financial Economics*, 667-689.

Fama, E., & French, K. (1992). The Cross-Section of Expected Stock Returns. *Journal of Finance*, 427-465.

Fama, E., & French, K. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 1-22.

García-Feijóo, L., & Jorgensen, R. D. (2010). Can operating leverage be the cause of the value premium? *Financial Management*, 1127-1154.

Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica*, 1121-1152.

Gulen, H., Xing, Y., & Zhang, L. (2011). Value versus Growth: Time-Varying Expected Stock Returns. *Financial Management*, 381-407.

Guo, H., Savickas, R., Wang, Z., & & Yang, J. (2009). Is the value premium a proxy for time-varying investment opportunities? Some time-series evidence. *Journal of Financial and Quantitative Analysis*, 133-154.

Han, B. (2007). Investor sentiment and option prices. *The Review of Financial Studies*, 387-414.

Hong, H., & Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance*, 2143-2184.

Huberman, G., & Kandel, S. (1987). Mean-variance spanning. *The Journal of Finance*, 873-888.

Israel, R., & Moskowitz, T. J. (2013). The role of shorting, firm size, and time on market anomalies. *Journal of Financial Economics*, 275-301.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, 65-91.

Kiku, D. (2006). *Is the value premium a puzzle?* Wharton Business School, Philadelphia.

Kothari, S. P., Shanken, J., & loan, R. G. (1995). Another look at the cross-section of expected returns. *The Journal of Finance*, 185-224.

Lakonishok, J., Sheifler, A., & Vishny, R. (1994). Contrarian investment, extrapolation, and risk. *The Journal of Finance*, 1541-1578.

Li, X., Brooks, C., & Miffre, J. (2009). The Value Premium and Time-Varying Volatility. *Journal of Business Finance & Accounting*, 1252-1272.

Miller, E. (1977). Risk, Uncertainty, and Divergence of Opinion. *The Journal of Finance*, 1151-1168.

Pástor, L., & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, 642-685.

Pätäri, E., & Leivo, T. (2015). A CLOSER LOOK AT VALUE PREMIUM: LITERATURE REVIEW AND SYNTHESIS. *Journal of economic Surveys*.

Pätäri, E., Leivo, T., & Honkapuro, J. (2010). Enhancement of value portfolio performance using data envelopment analysis. *Studies in Economics and Finance*, 223-246.

Phalippou, L. (2008). Where is the value premium? *Financial Analysts Journal*, 41-48.

Pontiff, J. (2006). Costly arbitrage and the myth of idiosyncratic risk. *Journal of Accounting and Economics*, 35-52.

Sharpe, W. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 425-442.

Shleifer, A., & Vishny, R. W. (1997). The limits of arbitrage. *The Journal of Finance*, 35-55.

Stambaugh, R. F. (1999). Predictive regressions. *Journal of Financial Economics*, 375-421.

Stambaugh, R. F., Yu, J., & Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics*, 288-302.

Wang, H., & Yu, J. (2014). An empirical assement of Models of the Value premium. *AFA 2013 San Diego Meetings Paper*.

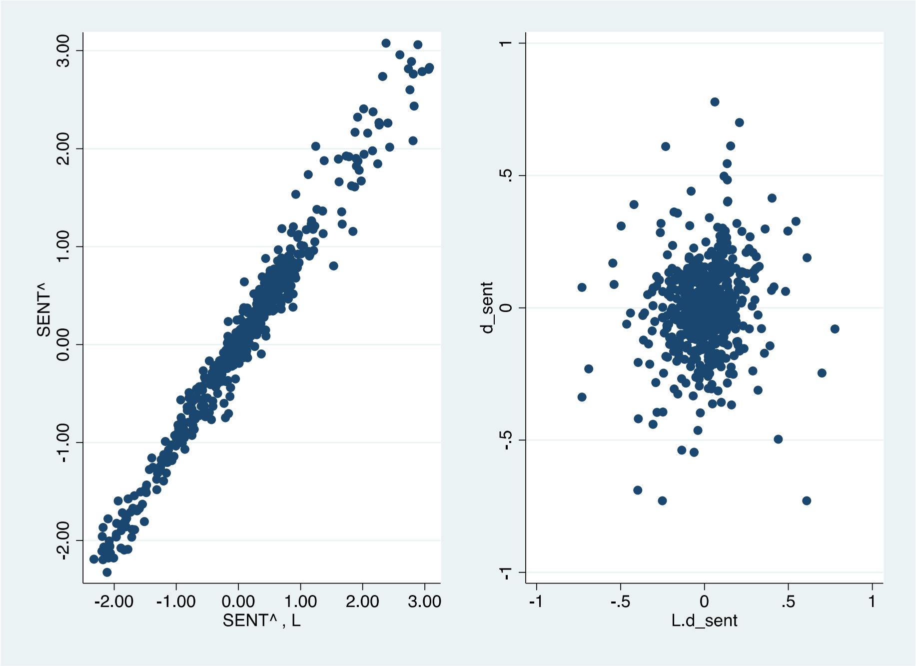
Zhang, L. (2005). The Value Premium. *The Journal of Finance*, 67-103.

Zhang, W., & Semmler, W. (2009). Prospect theory for stock markets: Empirical evidence with time-series data. *Journal of Economic Behavior & Organization*, 835-849.

Appendices

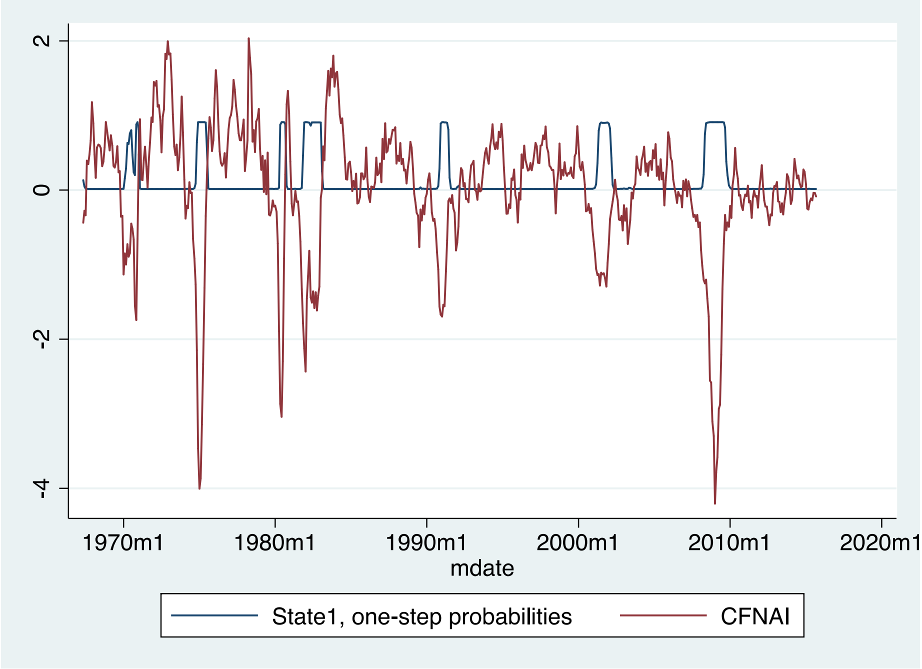
Appendix A

Graph A1: Correlation of SENT& ΔSENT (d_sent) between current and lagged value



Appendix B

Graph B1: Plot CFNAI with Markov-Switching regimes



Graph B2: Plot SENT with Markov-Switching regimes

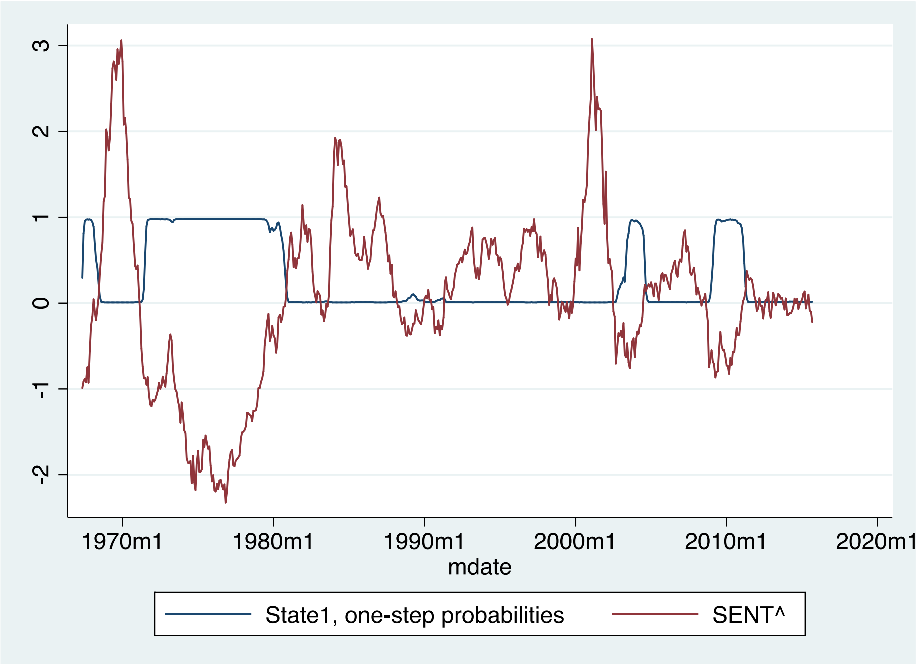


Table B1: Markov-Switching regression model CFNAI two states parameters summary

CFNAI	Coefficient	Standard Error	z	P> |z|
State 1				
ΔSENT	0,071	0,383	0,18	0,854
Constant	-1,762	0,092	-19,18	0,000
State 2				
ΔSENT	0,563	0,183	3,08	0,002
Constant	0,247	0,027	9,00	0,000
sigma	0,584	0,178		
p11	0,899	0,036		
P21	0,014	0,005		

Table B2: Mean CFNAI by dummy variable recession

Low	Mean	Std. Dev.	Freq.
0	0,27	0,51	492
1	-1,66	0,90	81
Total	0,00	0,89	573

Table B3: Markov-Switching regression model SENT two states parameters summary

SENT	Coefficient	Standard Error	z	P> |z|
State 1				
Constant	-1,129	0,072	-15,58	0,000
State 2				
Constant	0,484	0,037	13,05	0,000
sigma	0,671	0,020		
p11	0,978	0,119		
P21	0,007	0,004		

Table B4: Mean SENT by dummy variable low

Low	Mean	Std. Dev.	Freq.
0	0,51	0,68	416
1	-1,09	0,60	157
Total	0,07	0,97	573

Appendix C

Table C1: Models summary equal-weighted HLA value premia by economy state

Variable	Recession0	Recession1	Recession Control
MktRF	**-0,152**	-0,042	**-0,116**
ΔSENT	-0,367	1,650	0,212
SMB	**-0,125**	-0,015	-0,086
RMW	**0,444**	**0,452**	**0,447**
CMA	**0,938**	**1,177**	**0,951**
Mom	0,069	**-0,135**	-0,003
LIQ	-0,004	0,030	0,006
Recession			-0,629
Constant	**0,487**	-0,243	**0,505**
N	491	81	572
R-squared	0,574	0,640	0,561
R-squared adjusted	0,568	0,605	0,555

Table C2: Models summary equal-weighted LLA value premia by economy state

Variable	Recession0	Recession1	Recession Control
MktRF	-0,032	0,104	0,008
ΔSENT	**1,866**	1,202	**1,903**
SMB	-0,064	**0,224**	0,003
RMW	0,129	0,258	**0,155**
CMA	**0,971**	**1,458**	**1,033**
Mom	**-0,044**	**-0,257**	**-0,119**
LIQ	-0,004	-0,020	-0,004
Recession			-0,278
Constant	-0,015	**-0,765**	-0,010
N	491	81	572
R-squared	0,504	0,648	0,498
R-squared adjusted	0,497	0,614	0,491

Table C3: Models summary value weighted HLA value premia by economy state

Variable	Recession0	Recession1	Recession Control
MktRF	**-0,086**	0,041	-0,051
ΔSENT	0,915	1,023	0,926
SMB	**-0,142**	-0,011	**-0,120**
RMW	**0,391**	0,254	**0,371**
CMA	**1,024**	**1,358**	**1,068**
Mom	-0,048	-0,150	**-0,098**
LIQ	0,044	0,003	0,034
Recession			-0,207
Constant	0,179	-0,205	0,193
N	491	81	572
R-squared	0,611	0,680	0,605
R-squared adjusted	0,605	0,649	0,599

Table C4: Models summary value weighted LLA value premia by economy state

Variable	Recession0	Recession1	Recession Control
MktRF	0,042	0,123	0,063
ΔSENT	1,027	1,218	1,330
SMB	0,048	**0,358**	**0,118**
RMW	-0,061	-0,044	-0,040
CMA	**0,881**	**1,350**	**0,956**
Mom	**-0,124**	**-0,251**	**-0,171**
LIQ	-0,004	-0,018	-0,000
Recession			-0,455
Constant	0,020	-0,830	0,006
N	491	81	572
R-squared	0,405	0,546	0,419
R-squared adjusted	0,397	0,502	0,411