When to bet against beta? Ask Google.

Pedro Piccoli[1*]

ABSTRACT

In this paper, I document that investor attention negatively predicts betting against beta returns. Using Google Search Volumes toward US market indices as my proxy to attention, I find that this relation holds after controlling for competitive factors and different search terminologies and in most of the other G7 countries. The results also indicate that investor attention presents a unique capacity to explain future BAB performance that is not shared by other famous variables, such as liquidity constraints, sentiment, lottery demand or volatility. On aggregate, the findings suggest that individual investors are a relevant barrier to arbitrage strategies such as BAB.

Keywords: Betting against beta, Investor attention, low beta anomaly, Google Trends

EFM Classification codes: 320, 350

[1]Corresponding author. Rua Imaculada Conceição 1155, Curitiba, PR, Brazil, 80215 901.Tel.: +55 41 99262-6661. *E-mail address:* pedro.guilherme@pucpr.br
*Assistant Professor of Finance, Pontifícia Universidade Católica do Paraná, Brazil

1. Introduction

The fact that low beta securities exhibit higher future returns vis-à-vis their high beta peers is one of the most striking puzzles in asset pricing and has led to a plethora of studies suggesting different motivations for this anomaly. Some studies claim that the puzzle is driven by macroeconomic elements such as liquidity constraints (Frazzini and Pedersen, 2014), idiosyncratic volatility (Liu et al., 2018), or additional risk factors (Fama and French, 2015). Others document that the anomaly can be partially explained by behavioral aspects such as heuristic biases (Baker et al., 2011), investor sentiment (Liu et al., 2018) or lottery demand from individual investors (Bali et al., 2017). Finally, a third group of papers advocates that the puzzle emerges from asset pricing model misspecification (Cederburg et al., 2016 and Schneider et al., 2020), disappearing after employing the consistent model approach. Relying on the robust evidence of the anomaly, Frazzini and Pedersen (2014) proposed a trading strategy that aims to profit from this mispricing by longing (shorting) low (high) beta assets. The pervasiveness of the strategy, allied to its arbitrageur nature, generated considerable attention from scholars and practitioners. Consequently, recent studies have investigated possible candidates that can forecast betting against beta returns (hereafter, BAB) since they can plausibly improve the profitability of the strategy as well as provide further explanations for the underlying puzzle. It is in this respect that the paper aims to make its main contribution.

Given that BAB is designed to take advantage of arbitrage opportunities, its performance should be jeopardized when limits to arbitrage increase, such as during funding constraints or extreme market circumstances that hinder shorting positions due to capital scarcity or high uncertainty regarding future prices (Shleifer and Vishny, 1997). Previous empirical studies have already presented evidence to this end (Frazzini and Pedersen, 2014; Asness et al., 2020; Blitz et al., 2019). There is another arbitrage barrier, however, whose role in BAB strategy has scarcely been scrutinized in the literature: mispricing by retail investors. According to the noise trader approach, the presence of naïve investors can make the price diverge from fundamentals, hampering the capacity of arbitrageurs to bet against the perceptions of individual traders (De Long et al., 1990). Consequently, the presence of retail investors should produce negative BAB returns. This paper provides empirical evidence of this underlooked relationship, making a novel contribution to the literature.

In the paper, the influence of individual investors on the stock market is proxied by investor attention which, in turn, is measured by online searches on stock market indices as assumed by a growing number of studies (e.g.: Da et al., 2011). In this respect, I download the weekly Google Search Volumes (hereafter, GSV) on different US stock market indices and use the log-difference as my proxy for attention. This is the variable of interest in a series of models in which the dependent variable is BAB returns.

In a preliminary investigation, I run a Vector Autoregressive analysis between Attention and BAB. The results indicate a negative impact of attention on future BAB performance that is observable in all the stock market indices employed. Even though my central hypothesis attests that investor attention negatively forecasts BAB performance, it is worth investigating the contrarian impact, since a negative relation between the variables could be driven by alternative factors like financial distress, which would provoke BAB decreases, and also grab the attention of uninformed investors. The results, however, do not support this conjecture. Moreover, the Granger causality test clearly indicates that the influence is only observable in the Attention-BAB direction, and significant at a level of 1% for all the indices.

In a second analysis, I investigate the influence of lagged attention on BAB in an OLS model controlling for several alternative factors that might influence the relationship under scrutiny as suggested in former studies. In concrete terms, these variables are: investor sentiment (Baker et al., 2011; Liu et al., 2018); the Fama-French 5 factor model (Fama and French, 2016); and Business Cycle variables (Rapach, 2005; Bali et al., 2009). The results indicate a negative relation between attention and future BAB returns for all the indices investigated. This influence is also economically important, since a one standard deviation of attention on the S&P-500 (DJIA) index generates an annual decrease in BAB returns of 8.2% (7.9%), which is greater than the average raw return of 6.9% obtained by the strategy in the research period. Moreover, the sensitivity analysis suggests that investor attention plays a unique role in explaining BAB variations, given that the t-statistics of the model move from -2.24 (-2.18) to -2.63 (-2.81) when the control variables are included in the regression regarding attention to the S&P-500 (DJIA).

These findings provide support for the hypothesis that individual investors represent an additional limit to arbitrage, as advocated by the Noise Trader Approach (De Long et al., 1990), undermining the performance of arbitrage investment strategies like BAB. Furthermore, given that BAB aims to take advantage of beta anomaly, the results suggest that the influence of individual investors plays an important role in this puzzle. The importance of retail traders to mispricing has also been suggested in previous studies, such as Yu and Yuan (2011), Stambaugh et al. (2012, 2015) and Yu (2013), which document that the anomalies are blunter when sentiment is high. Since individual investors are more abundant during bullish markets, their presence would increase mispricing given their lower sophistication, leading to more abundant anomalies. My results provide evidence in this respect, employing a more accurate measure of retail traders' influence on the market, since online searches are a more direct measure of individual investor attention and desire to trade (Da et al., 2011).

Since funding constraints are important drivers of negative BAB returns (Fazzini and Pedersen, 2014; Asness et al., 2020), I also investigate the role of liquidity in the BAB-attention relation. In a first analysis, I add the TED spread to the aforementioned OLS model to capture the influence of funding shortage on BAB performance. As expected, the results indicate a negative relation between TED spread and BAB returns that is significant at the 1% level and present in every index under analysis, meaning that an increase in TED spread leads to a negative performance of BAB. Notwithstanding, the negative influence of attention on BAB remains significant, albeit less pronounced, after including the TED spread. This suggests that both liquidity constraints and investor attention capture different dimensions of arbitrage limits that are important drivers of BAB falls. This dichotomy is consistent with the argument presented by Asness et al. (2020), who advocate that the BAB factor is influenced by leverage constraints, in addition to behavioral explanations. In an additional model, I add an interactive variable between attention and the TED spread to examine how the BAB-attention relation is moderated by funding constraints. The results indicate that liquidity plays an important role in this relationship, which suggests that funding shortage not only influences less constrained agents (affecting the short lag of the strategy) but also more constrained investors that are overweighting high-beta assets, as hypothesized by the BAB model (Frazzini and Pedersen, 2014). In this context, I run an intertemporal analysis on the influence of the TED spread on BAB-attention relation and find that this influence holds conditionally.

Based on the fact that the literature has suggested candidates that partially explain BAB performance, I employ some of them in a horse race analysis to gauge whether the influence of attention on BAB is already captured by these constructs. Hence, besides online searches, I also investigate the influence of Liquidity Constraints (Fazzini and Pedersen, 2014), Investor Sentiment (Liu et al., 2018), Lottery Demand (Bali et al., 2017) and Volatility (Liu et al., 2018) on BAB returns. In the aggregated OLS model containing all these variables, together with the aforementioned control variables, the coefficients of Attention, Liquidity Constraints and Sentiment were the only ones to remain significant. Once again, in the case of Attention, the t-statistic was almost untouched from the model containing only this variable and the aggregated model, moving from -2.39 to -2.41. In addition to attention, this pattern was only observable in the case of Liquidity Constraints, providing further support for the view that these variables capture singular aspects of BAB variations. Additionally, in the second part of the horse race investigation, I perform a VAR analysis between Attention and the alternative variables. The results indicate that attention significantly predicts every competitive construct, whereas the contrary direction prediction is only observable for Lottery Demand and Volatility. In those cases, however, the Granger causality tests demonstrate more significant explanatory power of Attention, clearly indicating the prominence of this variable for forecasting BAB behavior.

The final analysis presented in the paper extends the investigation of the BAB-attention relation to the remaining G7 countries, controlling for the Fama-French 5 factors model and the business cycle variables described above. Among them, the influence is particularly pronounced in the UK, since a one standard deviation of attention on the FTSE index provokes an annual decrease of 11.7% on BAB in the following week, which is particularly remarkable, given that the BAB strategy in the UK earned a zero profit during the sampled period. Consistent with the pattern documented throughout the paper, there is a negative influence of investor attention on future local BAB returns in every G7 country, with the exception of Canada. Another common registered behavior is that the sensitivity analysis between the models with and without the control variables shows that the t-statistics of the significant lagged attention increases when the control variables are included, reinforcing the uniqueness of attention among the control variables employed. Analyzing the low-risk anomaly in the same G7 countries, Ang et al. (2009) document a persistent pattern in which high IVOL stocks earn lower future returns. The findings documented here add to this literature since they indicate that the low beta anomaly, a close friend of the low-risk puzzle, is influenced by investor attention in most of the remaining G7 countries.

On aggregate, the paper contributes to the literature in several ways. First, by providing a new construct that can be used to improve BAB performance adding to the practitioner-oriented literature on low-risk strategies (Alquist et al., 2020). Second, by suggesting a new behavioral variable that can help to explain the low beta puzzle (Baker et al., 2011; Bali et al., 2017; Asness et al., 2020). Third, by identifying that not only capital constraints are important limits to arbitrage, but also that retail investors play a significant role in this regard, providing empirical evidence of the relevance of noise traders to stock markets (De Long et al., 1990; Shleifer and Vishny, 1997; Barber et al., 2009). Finally, the findings also add to a growing body of studies reporting that market anomalies are plausibly driven by the influence of individual investors (Yu and Yuan, 2011; Stambaugh et al., 2012, 2015; Stambaugh and Yuan, 2017).

The paper is structured as follows. Section 2 describes the data. Section 3 presents the VAR analysis between BAB and attention. Section 4 investigates the relationship between these variables, employing OLS and quantile regression approaches. Section 5 examines the influence of funding constraints on the BAB-attention relation. Section 6 runs horse races between different variables related to BAB. Section 7 presents international evidence of the BAB-attention relation, while Section 8 concludes the study.

2. Data

The primary variables of this study are Google Searches, as a proxy for investor attention, and the returns of the Betting Against Beta strategy. The data for Google Searches were collected from the Google Trends platform using the expressions “S&P 500” and “Dow Jones Industry Average”. For robustness, I add the searches for a more generic term, namely "Stock Market", aiming to capture the attention of a broader spectrum of retail investors who plausibly would search for the market's news without typing the more technical terminology of the indexes' names. The search volumes are restricted to the US since I aim to capture the interest of retail investors in trading on the stock market. For the entire sample (i.e., 2004 to 2019), Google Trends makes available only monthly data where the period with the largest number of searches peaks at 100. To obtain weekly data, I use 4 overlapping windows of 5 years, which is the longest period for which Google search volume (GSV) information is displayed on a weekly basis. The windows begin in 2004, the first year with Google Trends data, and end in 2019. The overlaps of the windows have a length of one year and are used to estimate the corresponding search volumes in relative terms[1].

The data for the Betting Against Beta factor (hereafter, BAB) were obtained from the AQR Capital Management website (www.aqr.com), which continuously updates the portfolio construction based on Frazzini and Pedersen (2014). Since the dataset is available on a daily basis, to be consistent with the GSV data, I create an index by accumulating the daily returns of BAB, match the dates with the GSV dataset, and find the BAB weekly return using the log difference of each week. These data are available in the online appendix, together with the GSV. Table 1 shows the summary statistics of BAB returns and online searches, the latter measured by the log-difference of the GSV in the corresponding index[2].

(Table 1)

As expected, the BAB strategy provides positive average returns, despite the negative skewness. The online searches, in turn, exhibit both positive variation and skewness, indicating a growing interest of individual investors in the stock market. It is also worth mentioning that all distributions, apart from searches on the S&P, exhibit fat tails. This feature might raise the objection that some of the results can be biased by the presence of outliers. As will be discussed later, I do not find support for this hypothesis based on the results provided by a quantile analysis.

(Figure 1)

[1] The windows are: I) 01.2004 to 12.2008; II) 01.2008 to 12.2012; III) 01.2012 to 12.2016; IV) 01.2016 to 12.2019. The coefficients of the overlaps are all highly significant.
[2] These variables are highly stationary, since the Dickey-Fuller unit root test is of Z(t) = -34.01 for BAB, Z(t) = -39.54 for Attention to the S&P-500, Z(t) = -36.24 for Attention to the Dow Jones and Z(t) = -33.97 for attention on Stock Market terminology.

Figure 1 displays the cumulative weekly returns of BAB, starting with $100 (black line, left axis) together with the GSV for the S&P 500 (gray line, right axis) for the entire dataset. It can be seen that, even-though the time-series seem to comove, there are several spikes in attention that are accompanied by decreases in the BAB portfolio, as shown by the dashed lines. This suggests that increases in investors' attention are associated with negative returns in the BAB factor, which can be plausibly explained by the belief that retail investors increase the uncertainty of future prices, elevating the limits to arbitrage and, consequently, jeopardizing the profits of the BAB strategy. Another plausible explanation is that both the GSV and BAB are responding to the influence of an underlying force. In this respect, since some of the GSV spikes and BAB falls are clustered during recent financial crises (subprime and Eurozone), it is reasonable to conjecture that the negative BAB-GSV relation is explained by funding constraints, which provoke negative BAB returns (Frazzini and Pedersen, 2014) and concomitantly grab investor attention toward the stock market since they are related to economic downturns. This hypothesis is addressed in a specific section, and the results indicate that, although part of the BAB-GSV association is explained by liquidity tightness, there is still a significant relation between these constructs. A detailed investigation of this relation is presented in the following sections.

3. Predictability

Since noise traders can move prices away from their fundamentals, a growing influence of such investors on the market may shorten the arbitrage opportunities (Shleifer and Summers, 1990), limiting the possibility of rational investors betting against mispricing (De Long et al., 1990). Given that the BAB strategy is arbitrageur by its very nature, in the noise trader approach framework, one would expect that an increase in retail investor attention would negatively predict BAB returns. Nevertheless, since BAB decreases are reasonably associated with funding constraints, this hypothetical negative association could actually capture the increase in attention towards the market stress that is related to liquidity constraints rather than a direct influence of noise traders on BAB per se. Consequently, a natural way to investigate this possible bi-directional relation is running a VAR analysis of both variables, as addressed in the following equations:

$$BAB_t = \delta_0 + \delta_1 BAB_{t-1} + \cdots + \delta_n BAB_{t-n} + \gamma_1 S_{t-1,i} + \cdots + \gamma_n S_{t-n,i} + \varepsilon_t \qquad (1)$$

$$S_{t,i} = \delta_0 + \delta_1 BAB_{t-1} + \cdots + \delta_n BAB_{t-n} + \gamma_1 S_{t-1,i} + \cdots + \gamma_n S_{t-n,i} + \varepsilon_t \qquad (2)$$

Where BAB_t stands for the return of betting against a beta portfolio in week t and $S_{t,i}$ is the contemporaneous weekly log difference of GSV in the given index[3]. The results are shown in Table 2. For all the indexes, there is a negative influence of attention on BAB returns in the following week that is significant at the 1% level. This relation is also economically representative, given that an increase in one-standard deviation of online searches on the S&P-500 (Dow) provokes an annual negative return on BAB of 9.1% (8.9%) in the following week, which is substantial since the average annual return of this strategy in the sampled period is 6.9%. On the other hand, in the opposite direction, there is no significant influence of past BAB on attention for any of the indexes, supporting the hypothesis that an increase in attention induces a decrease in BAB performance. This explanation is also justified by the Granger causality test since the results clearly indicate that attention Granger causes BAB, while the opposite causation is not observable.

(Table 2)

To provide information on the timing and direction of the relationship, Figure 2 shows the impulse-response chart for each index, where the upper part refers to attention-BAB causation, while the bottom exhibits the results for the opposite direction. The orthogonalized impulse response functions are displayed for eight periods after the one standard deviation shock. The charts indicate that the influence of attention on BAB is short-lived, ceasing after two weeks. They also indicate that this relation is qualitatively the same, independent of the index used as the proxy for investor attention. Finally, as expected, there is no significant influence of BAB performance on attention.

(Figure 2)

So far, the results indicate that retail investor attention negatively forecasts BAB returns. To delve deeper into how these variables are related, and whether this association is driven by underlying variables, the next section investigates the BAB-Attention relation employing an OLS specification controlling for several alternative factors.

4. BAB-Attention relation

[3]An alternative approach to calculate S is using the log-difference between the GSV for a given period and the median GSV of a past moving window. There is no consensus in the literature as to which option (if any) is preferable. To address this, I run the same analysis in accordance with Da et al. (2011), who employ this alternative approach and find that the results are virtually the same. Appendix A summarizes the main results using this operationalization.

In this section, I analyze the cross-sectional relation between past attention and BAB performance in the following OLS models:

$$BAB_t = a_0 + \sum_{n=1}^{4} b_{i,n} S_{i,t-n} + \sum_{n=1}^{4} c_n BAB_{t-n} + \varepsilon_t \quad (3)$$

$$BAB_t = a_0 + \sum_{n=1}^{4} b_{i,n} S_{i,t-n} + \emptyset X_t + \sum_{n=1}^{4} d_n BAB_{t-n} + \varepsilon_t \quad (4)$$

Where BAB_t and S_i were previously defined. The lagged BAB aims to control for autocorrelation, and X_t is a vector containing several control variables that could potentially influence the relation under scrutiny[4]. More precisely, these variables are:

a) Sentiment: The literature documents that market anomalies are stronger during high sentiment periods (Stambaugh et al., 2012), a pattern that is also observable in the case of a low-beta puzzle (Liu et al., 2018). Given that sentiment and investor attention are potentially connected (Bucher, 2017; Mbanga et al., 2019), to avoid this source of endogeneity, I employ the American Association of Individual Investors (AAII) weekly survey data and retain the percentage of bullish investors as my proxy for sentiment. My choice to use this rather than the more famous Baker-Wurgler sentiment index is justified by the fact that the latter is only available in monthly terms.
b) Fama-French 5-factor model. The extended FF-model has a higher capacity to explain anomalies in the cross-section, including the low-beta anomaly (Fama-French, 2016). The five factors were downloaded from Kenneth French's website. Since their data is not available on a weekly basis, I follow the same approach described in Section 2 to generate a weekly time-series for the Fama-French five factors.
c) Business Cycle. Given that business cycle fluctuations can be used to predict market performance (Rapach et al., 2005), I follow Bali et al., 2009 and Wu and Lee, 2015 and include four macro-economic variables on a weekly basis as follows: detrended riskless rate (RREL), dividend yield (DY), default spread (DEF) and term spread (TERM). The first business cycle variable (RREL) is set as the difference between the 1-month Treasury bill rate and its 12-month moving average. DEF is the difference between the yields in BAA- and

[4]The Spearman correlation between attention and the control variables are of small magnitude, indicating that online searches capture a distinct dimension of BAB variances that is not explained by these variables. For instance, the largest correlation exhibited by S is of -0.13 with Market Risk Premium. Appendix B includes the Spearman correlation table for all the variables.

AAA-rated corporate bonds. TERM is the difference between the yields in the ten-year Treasury bond and the three-month Treasury bill, and together with DEF and RREL data, was obtained from the Federal Reserve Bank of St. Louis website. Finally, DY is constructed using the twelve-month dividend data obtained from Robert Shiller's website[5].

The results are shown in Table 3. For brevity, I suppress the estimates for Business Cycle and autocorrelation controls. The table reports a negative influence of attention on BAB returns in the following week for all the indices, which remain significant at the 1% level when control variables are included in the model. This relation is also economically representative, since a one-standard deviation increase in attention corresponds to a decrease in BAB returns of 8.2% in annual terms the following week, using S&P searches as the proxy for attention. It should also be mentioned that the inclusion of the control variables better accommodates the BAB variation, given the substantial increase in the adjusted R^2 of Model (4). This increase is especially driven by Sentiment and the profitability (RMW) factor, which is consistent with previous studies. First, in the case of Sentiment, Liu et al. (2018) document that the beta anomaly is more prevalent during high sentiment regimes, while Baker et al., (2011) suggest that this puzzle can be partially driven by investor sentiment. The fact that I report a positive and significant contemporaneous association between Sentiment and BAB supports this notion. Second, in the case of the profitability (RMW) factor, Fama and French (2016) report that their extended five-factor model better accommodates well-known anomalies, including the flat beta, which is also documented by Asness et al. (2020). In this case, the positive relation indicates that BAB is longing (shorting) firms with robust (weak) profitability, which is consistent with the nature of the strategy of longing (shorting) firms with low (high) systemic risk.

The fact that investor attention remains significant after including these regressors suggests that this variable captures a different dimension of BAB performance that is not covered by these classical constructs. In this respect, it is interesting to note that the sensitivity analysis between Models (3) and (4) demonstrates that the significance of attention estimates increases when the control variables are included. Since attention coefficients decrease in Model (4), this indicates that their standard errors decrease more when control variables are included, lending support to the view that the influence of attention on BAB is not captured by the remaining variables.

(Table 3)

[5]The data for RREL, DEF and TERM is daily. I then match with the corresponding week of the GSV and BAB dataset to generate the weekly timeseries. In the case of DY, since the data is monthly, I repeat the given value throughout the weeks of the corresponding month. The dataset for all the variables used in the paper are available in online Appendix.

The results so far suggest that investor attention plays an important role in explaining future BAB performance. Nevertheless, given that the BAB strategy exhibits fat tails, as demonstrated in the descriptive statistics, a plausible question would be how prevailing the influence of attention is on BAB returns. The hypothesis that the results could be driven by outliers is especially reasonable for this investigation, since the BAB portfolio exhibits negative returns during liquidity constraint periods, which are commonly associated with market distress that, in turn, grabs investor attention. Furthermore, since Schneider et al. (2020) report that low volatility anomalies are largely explained by skewness, and given that individual investors exhibit a preference for skewness (Andrei and Hasler, 2015), the fat tails story seems to be a reliable alternative explanation for these preliminary findings.

To address this, I run a quantile regression using Model (4). For brevity, Table 4 shows the results only for searches on the S&P. The results for the other two indices are virtually the same and can be made available upon request. To avoid heteroskedastic and autocorrelation issues, the standard errors of the coefficients were obtained by block-bootstrap.

(Table 4)

The negative influence of attention on BAB returns one week ahead is observable in all quantiles, albeit more pronounced in the middle of BAB distribution, which is clearly opposed to the hypothesis that the relation is driven by tail events. The fact that the left tail quintiles (i.e., q05 and q10) are the ones for which past attention presents the lowest significance provides further support in this respect, since it indicates that this influence is less important during BAB's crashes (i.e., high skewness circumstances). Finally, it is also interesting to note that the impact of attention is more lasting when BAB exhibits more extreme positive performances (i.e., q90 and q95), since those are the quantiles for which the two-lag attention is more significant than the one-lag parameter. Given that the influence remains negative, this suggests that the positive returns of the strategy are partially driven by a lower level of investor attention. All in all, the estimates in Table 4 suggest a prevailing influence of attention on BAB performance that is driven neither by tail events nor by market states, lending further support to the view that retail investors are relevant players for future BAB returns.

Given that the BAB strategy aims to take advantage of arbitrage opportunities, its profits are jeopardized when limits to arbitrage increase. According to the Noise Trader Approach, the uncertainty in future stock prices brought about by the entrance of naïve investors is an important barrier to arbitrage. Consequently, in the phenomena investigated by the paper, one should expect that increases in online searches would lead to a negative performance of the BAB strategy. This is exactly the pattern I document. This relation also suggest that retail investors play an important role in the low-beta anomaly. Given that they exhibit a preference for risky stocks (Kumar, 2009), a massive entrance of individual traders in the market would raise the price of high beta assets in the short run, leading to negative BAB returns. In the medium term, however, when the market moves to equilibrium, these lottery stocks would exhibit future negative returns, leaving room for the appearance of the low-beta puzzle. Even though the investigation on how retail investors could be a driving force of this anomaly exceeds the scope of this paper, the findings here documented lend support to the view that uninformed investors are important agents of market inefficiency, as suggested by a growing number of recent studies (Bali et al., 2017; Shen et al., 2017; Stambaugh and Yu, 2017).

Another important limit to arbitrage is liquidity constraints. Since funding shortage is especially acute during financial crises and given that individual investors are attracted by skewness (Barberis and Huang, 2008; Barberis and Xiong, 2012), it is reasonable to investigate the role of funding constraints in the BAB-attention relation. The next section addresses this topic.

5. Liquidity Constraints

Following the standard approach in the literature, I employ the TED spread as my proxy for funding limitation. The data are from the Federal Reserve Bank of St. Louis website. To analyze its influence on the BAB-attention relation, I include the TED spread in Models (3) and (4), resulting in Models (5) and (6). Moreover, since the results in the previous section demonstrate that the influence of attention on BAB is limited to two lags, for conciseness, I suppress lags three and four of both attention and BAB in the new models' specifications:

$$BAB_t = a_0 + \sum_{n=1}^{2} b_{i,n} S_{i,t-n} + \sum_{n=1}^{2} c_n TED_{t-n} + \sum_{n=1}^{2} d_n BAB_{t-n} + \varepsilon_t \tag{5}$$

$$BAB_t = a_0 + \sum_{n=1}^{2} b_{i,n} S_{i,t-n} + \emptyset X_t + \sum_{n=1}^{2} c_n TED_{t-n} + \sum_{n=1}^{2} d_n BAB_{t-n} + \varepsilon_t \tag{6}$$

Furthermore, to gauge whether investor attention influence on BAB is conditional to changes in liquidity, I add an interaction term $S_{i,t-n} \times TED_{t-n}$ to the above equations, resulting in Models (7) and (8).

$$BAB_t = a_0 + \sum_{n=1}^{2} b_{i,n} S_{i,t-n} + \sum_{n=1}^{2} c_n TED_{t-n} + \sum_{n=1}^{2} d_{i,n} S_{i,t-n} \times TED_{t-n} + \sum_{n=1}^{2} d_n BAB_{t-n} + \varepsilon_t \tag{7}$$

$$BAB_t = a_0 + \sum_{n=1}^{2} b_{i,n} S_{i,t-n} + \emptyset X_t + \sum_{n=1}^{2} c_n TED_{t-n} + \sum_{n=1}^{2} d_{i,n} S_{i,t-n} \times TED_{t-n} + \sum_{n=1}^{2} e_n BAB_{t-n} + \varepsilon_t \tag{8}$$

The estimates are summarized in Table 5. For clarity, the table displays only the results for the variables of interest in this subsection, i.e., Attention, TED spread and the interaction term. The results for the remaining variables can be made available upon request. As expected, the TED spread exhibits a negative influence on BAB returns one week ahead that is significant at the 1% level for every model and in all the indices, indicating that liquidity constraints (high TED spread) deteriorate future BAB performance. Despite this, in Models (5) and (6), the one-lag attention remains significant, albeit in a less pronounced way, to influence BAB returns. Taken together, these results suggest that, although funding restrictions play an important role in BAB returns, as assumed in Frazzini and Pedersen's (2014) third proposition, this relation does not overlap the influence of attention on betting against beta. More recently, Asness et al. (2020) disentangle the BAB factor into two components: betting against correlation (BAC), and betting against volatility (BAV). They advocate that, while the first component is more exposed to liquidity shortage, the second is more dependent on investor sentiment. In some way, the results of Model (5) and (6) provide support in this respect by documenting that both the TED spread and investor attention coexist as explanatory variables for BAB's overall performance. Based on Asness et al. (2020), one could plausibly conjecture that my proxy for attention may actually capture investor sentiment. The estimates do not find support for this, since Model (6) controls for sentiment, and the coefficient for attention remains significant. Furthermore, as will be addressed later, in a horse races analysis among behavioral explanations for BAB performance, I find evidence of causation in the opposite direction, i.e., attention forecasting sentiment.

(Table 5)

When I include the interaction term, the significance of one-lag attention ceases, while the interaction between attention and the TED spread in this lag exert a significant influence on future

BAB returns. Moreover, whereas the significance of one-lag attention vanishes, the influence of one-lag TED spread remains almost untouched. This behavior indicates that the impact of attention on future BAB performance depends on liquidity conditions. Since retail investors are more constrained, when funding conditions tighten, their overweight on high-beta stocks becomes limited, exhausting one of the sources of BAB profit: the opportunity to bet against the preference of retail investors for risky stocks (Kumar, 2009; Bali et al., 2017).

Given that the TED spread peaks during financial crises, it is worth investigating whether the influence of liquidity shortage on the BAB-attention relation is restricted to these turbulent circumstances. To investigate how prevailing the influence of liquidity constraints is on this relation, I generate a conditional estimate of the relation using a moving window regression of BAB on one-lag Attention (S_{t-1}) together with BAB_{t-1} to control for autocorrelation. The length of the moving window regression is 30 weeks. I then use this time-series of the coefficient of Attention as the dependent variable of Models (9) and (10):

$$bS_{i,t} = a_0 + a_1 TED_{t-1} + \varepsilon_t \tag{9}$$

$$bS_{i,t} = a_0 + a_1 TED_{t-1} + \emptyset X_{t-1} + \varepsilon_t \tag{10}$$

The estimates are shown in Table 6. They demonstrate that that TED spread exerts a significant intertemporal influence on the BAB-attention relation, even after controlling for competitive behavioral and macroeconomic variables. These results indicate that the prevailing negative influence of attention on BAB performance is moderated by funding constraints, becoming more important when liquidity tightens (TED spread increases). It should also be noted that, even though liquidity plays a relevant role in the relation between BAB and investor attention, most of the variance of this association is not explained by the TED spread. This fact suggests that investor attention has a peculiar influence on BAB returns. To explore this further, the next section presents horse race analyses of behavioral and macroeconomic variables that are believed to help explain the low beta anomaly.

(Table 6)

6. Horse races

Given the substantial interest in the BAB strategy, there is a plethora of studies that attribute the performance of the strategy to different economic or behavioral variables. Consequently, it is plausible to question whether the influence of retail investor attention is indeed unique, or if it is already captured by the available studies. To shed light in this direction, I develop horse races among some of these potential explanatory variables, together with online searches. The choice of these variables is based on the reasonable connection with investor attention together with the facility for data access, and are listed below:

a) Liquidity Constraints: As discussed in the previous section, it is well stablished that liquidity constraints play an important role in BAB performance due to arbitrage limits that affect short selling from arbitrageurs (Frazzini and Pedersen, 2014; Asness et al., 2020). Once more, this variable is proxied by the TED spread.
b) Investor Sentiment: There is growing literature reporting that mispricing is more acute during high sentiment circumstances (Stambaugh and Yuan, 2017), leaving room for the appearance of market anomalies (Stambaugh et al., 2012), including the low-beta puzzle (Baker et al., 2011, Liu et al., 2018). As previously addressed, in this analysis this variable is proxied by the proportion of bullish respondents of the American Association of Individual Investors (AAII) weekly survey.
c) Lottery Demand: Recently, Bali et al. (2017) document that the low-beta anomaly can be partially explained by the preference of retail investors for lottery-like stocks. Using a factor that aims to capture this gambling propensity (FMAX), they find that BAB profit ceases after controlling for FMAX. Given that this factor is intrinsically linked to the influence of individual investors, it is reasonable to believe that investor attention is endogenous to lottery demand, which makes this investigation particularly interesting. The drawback is that the FMAX factor available on Turan Bali's website is on a monthly basis. To have weekly data, I employ the following process. The FMAX in a given week is the weighted average of the FMAX between the current and the previous month. The weight for the current month is the ratio between the day of the given week and the total days of the given month, while the weight of the previous month is given by one minus this ratio. For example, the FMAX for 2019.09 (2019.10) is -3.66 (0.18). Hence, the FMAX for the week 2019.10.07 is $FMAX^{week} = \frac{7}{31} \times 0.18 + \left(1 - \frac{7}{31}\right) \times (-3.66) = -2.796$.
d) Volatility: The literature documents that the beta anomaly is more important when volatility is high. For example, Liu et al. (2018) report a positive correlation between the anomaly and idiosyncratic volatility (IVOL) and that the anomaly is observable only when the correlation between beta and IVOL is high. Moreover, Schneider et al. (2020) find that BAB displays insignificant alphas when controlling for skewness. Based on these findings, I employ a fourth alternative variable in the horse races that is related to volatility. Since my proxy for

investor attention is given by online searches on the market index, for consistency I employ a measure of volatility based on the corresponding index. Specifically, I run a GARCH (1,1) model on the index weekly return to generate a conditional measure of volatility[6] (VOL_i).

In the first part of the horse races, I employ an OLS regression of BAB on each of the above lagged variables together with the Fama-French 5 factors, Business Cycle variables and four-lag BAB, the latter aiming to control for autocorrelation. Table 7 presents the estimates of the models for each variable separately (columns 1 to 5), together with an additional specification that aggregates all these variables (column 6). For brevity, the table shows only the results for the Dow Jones.

(Table 7)

Apart from VOL, all the variables described above significantly predict future BAB performance in the expected direction: while Attention and the TED Spread are negatively related to BAB, for SENT and FMAX the association is positive. Nevertheless, when these variables are brought together (right column), the significance of FMAX to forecast BAB ceases. Furthermore, it should be noted that in the aggregated model, Attention and the TED Spread are the only variables for which the t-statistics remain virtually untouched, which suggests that these variables do indeed capture a unique variation of BAB. This pattern is in line with Asness et al. (2020), who advocate that BAB variations can be driven by funding constraints (e.g., TED spread) and behavioral factors (e.g., Attention).

To better explore the relation between investor attention and the five alternative variables described above, the second part of the horse races consists of a VAR analysis between online searches and each of these competitor constructs. The results are shown in Table 8, the bottom part of which also shows the Granger causality test for the corresponding relation under study.

(Table 8)

The results indicate a remarkable influence of investor attention on every alternative variable. This is especially observable in the case of Liquidity Constraints and Volatility, since all lagged attention is highly significant to explain current changes in these constructs. The fact that this relation is positive suggests that retail investors are a source of instability in the stock market, given that both variables are associated with market distress. The relation in the opposite direction, however, is observable only for the Lottery Demand factors and Volatility, and yet it is less significant than the

[6]For the "Stock Market" terminology, I calculate the conditional volatility using the returns of the CRSP index, which was obtained from Kenneth French's website. For the other indices, the dataset is from Yahoo Finance.

influence of attention on each variable. This pattern is also observable in the results of the Granger causality test, which indicate that investor attention can forecast future changes in these variables, whereas the contrary causation is only observable for Volatility. Taken together, the results in Tables 7 and 8 indicate not only that investor attention captures a singular dimension of BAB variations but also that it presents a unique capacity to forecast BAB performance.

7. International Evidence

In this last part of the empirical analysis, I investigate whether the influence of investor attention on BAB performance can be extended to other G7 countries, given the evidence that this strategy is also profitable in other markets (Frazzini and Pedersen, 2014). Consequently, I downloaded the weekly GSV on the local indices[7] from 2004 to 2019, replicating the process described in Section 2 to generate the proxy for attention in the given stock market. I employ Models (3) and (4) to run this analysis, but without controlling for sentiment, due to data constraints. On the other hand, I include the TED spread, given the importance of liquidity constraints in this relation, as previously demonstrated. The data for BAB returns were also downloaded from the AQR Capital Management website[8]. Finally, the factors of Fama-French's model are the factors for developed countries excluding the US, which are available on a daily basis on Kenneth French's website. To generate weekly data for BAB and the FF factors, I employ the same approach described in Section 2. The results are shown in Table 9.

(Table 9)

Apart from Canada, in all the remaining countries there is a negative influence of attention on future BAB performance in at least one of the models employed. This impact is especially important in the UK and Germany, the two most representative markets in Europe. In the case of the UK, for example, an increase of one standard deviation in attention provokes an annualized decrease in BAB of -11.7% in the following week. One possible explanation for the exception in the case of Canada relies in the peculiar performance of the BAB in this country, which provided an astonishing geometric average return of 31% in annual terms, limiting the capacity of investor attention to explain

[7] The terminologies used were: "FTSE 100" for the UK, "DAX" for Germany, "Nikkei" for Japan, "CAC 40" for France, and "S&P/TSX" for Canada. In the case of Italy, employing the name of the local index (i.e., "FTSE MIB") resulted in many missing data. To illustrate, from 2004 to 2008, the data from Google Trends employing this terminology generated 72% of the GSV = 0.0. To overcome this, I alternatively employed the term "Stock Market", which reduced the number of missing data to 20% in the same five-year window. Furthermore, in the case of Japan, Italy and France, there were some weeks in the first window of the dataset (i.e., 01.2004 to 12.2008) where the GSV reported by Google Trends was zero. Since this precludes the calculation of the log-differences, I replaced this data with one (1.0).

[8] The data for BAB and the GSV of every country are available in the online Appendix.

BAB's variations. It is also interesting to note that all the significant coefficients of attention are negative, providing strong support for the hypothesis that retail investors do indeed increase limits to arbitrage, jeopardizing the performance of strategies that aim to take advantage of mispricing, such as betting against beta. Furthermore, a sensitivity analysis comparing the coefficients of attention between Models (3) and (4) reveals that the vast majority increased their significance when adding control variables, indicating once more that online searches exert a singular influence on BAB. Investigating the same G7 countries, Ang et al (2009) document that idiosyncratic volatility and future average returns are negatively correlated in the cross-section. Since BAB and idiosyncratic risk are related (Asness et al., 2020), the results in Table 9 can shed light on the low-risk puzzle by suggesting that this anomaly is related to the presence of individual investors in stock markets.

8. Conclusion

According to the Noise Trader Approach, retail investors are an important force behind limits of arbitrage, undermining the performance of strategies that aim to profit from mispricing, as the betting against beta portfolio does. Based on this argument, an increase in the attention of individual investors should lead to a decrease in future BAB returns. Employing the volume of online searches on Google toward different US stock market indexes, this paper presents robust evidence in this respect, since I document that attention negatively affects future BAB performance. The results also indicate that investor attention has a singular capacity to explain future BAB returns that is not captured by other explanatory variables such as liquidity constraints, investor sentiment, lottery demand or volatility. On the contrary, the horse race analysis demonstrates that investor attention affects every one of these variables. Finally, the international analysis on the remaining G7 countries also attests to the central hypothesis of the paper: that increases in investor attention provoke decreases in future BAB performance.

The findings reported here can also shed more light on one of the most striking anomalies in finance: the low beta puzzle. Since retail investors exhibit a preference for risky stocks (Kumar, 2009), a massive entrance of individual traders would increase the demand for high beta securities, elevating their prices in the short run, leading to a decrease in BAB performance. In the mid-term, however, when the market moves to equilibrium, this behavior would produce negative future returns on high beta assets. There is a growing literature on asset pricing that advocates the prominence of individual investors for market inefficiency. The results presented in this paper make a contribution in this respect. Even though investigating the mechanism by which individual investors can lead to low beta anomaly exceeds the scope of this paper, this process seems to be an interesting agenda for future research. For the time being, the most important message presented in this paper is that one should consult Google before betting against beta.

References

Alquist., R., Frazzini, A., Ilmanen, A., Pedersen, L. H., 2020. Fact and Fiction about Low-Risk Investing. *Journal of Portfolio Management*, 46(6), 1-21

Andrei, D., Hasler, M., 2015. Investor attention and stock market volatility. *Review of Financial Studies*, 28(1), 33-71.

Ang., A., Hodrick, R. J., Xing, Y., Zhang, X., 2009. High idiosyncratic volatility and low returns: International and further U.S. evidence. *Journal of Financial Economics*, 91, 1-23

Asness, C., Frazzini, A., Gormsen, N. J., Pedersen, L. H., 2020. Betting against correlation: Testing theories of the low-risk effect. *Journal of Financial Economics*, 135(3), 629-652

Baker, M., Bradley, B., Wurgler, J., 2011. Benchmarks as Limits to Arbitrage: Understanding the Low-Volatility Anomaly. *Financial Analyst Journal*, 67, 40-54

Bali, T., Brown, S. J., Murray, S., Tang, Y., 2017. A Lottery-Demand-Based Explanation of the Beta Anomaly. *Journal of Financial and Quantitative Analysis*, 52(6), 2369-237

Bali, T., Demirtas, K., Levy, H., 2009. Is there an intertemporal relation between downside risk and expected returns? *Journal of Financial and Quantitative Analysis*, 44, 883–909.

Barber, B. M., Odean, T., Zhu, N., 2009. Do retail traders move the market? *Review of Financial Studies*, 22(1), 151-186

Barberis, N., Huang, M., 2008. Stock as lotteries: the implication of probability weighting for security prices. *American Economic Review*, 98 (5), 2066-2100

Barberis, N., Xiong, W., 2012. Realization utility. *Journal of Financial Economics*, 104, 251-271

Blitz, D., Vliet, P., Baltussen, G., 2019. The volatility effect revisited. *SSRN Working Paper.*

Bucher, M. C., 2017. Investor Attention and Sentiment: Risk or Anomaly? *SSRN Working Paper.*

Cederburg, S., O'Doherty, M. S., 2016. Does It Pay to Bet Against Beta? On the Conditional Performance of the Beta Anomaly. *Journal of Finance*, 71(2), 737-774

Da, Z., Engelberg, J., Pengjie, G., 2011. In Search of Attention. *Journal of Finance*, 66(5), 1461-1498

De Long, J. B., Shleifer, A., Summers, L. H., Waldmann, R. J., 1990. Noise trader risk in financial markets. *Journal of Political Economy*, 98(4), 703-738

Fama, E. F., French, K. R., 2015. A five-factor asset pricing model. *Journal of Financial Economics*, 116, 1-22

Fama, E. F., French, K. R., 2016. Dissecting anomalies with a five-factor model. *Review of Financial Studies*, 29(1), 69-103

Frazzini, A., Pedersen, L. H., 2014. Betting Against Beta. *Journal of Financial Economics*, 111, 1-25

Kumar, A., 2009. Who gambles in stock market? *Journal of Finance*, 64, 1889-1933

Liu, J., Stambaugh, R. F., Yuan, Y., 2018. Absolving beta of volatility's effects. *Journal of Financial Economics*, 128, 1-15

Mbanga, C., Darrat, A. F., Park, J. C., 2019. Investor sentiment and aggregate stock returns: the role of investor attention. *Review of Quantitative Finance and Accounting*, 53, 397-428

Rapach, D., Wohar, M., Rangvid, J., 2005. Macro variables and international stock return predictability. *International Journal of Forecasting*, 21, 137–166

Schneider, P., Wagner, C., Zechner, J., 2020. Low-Risk Anomalies? *Journal of Finance*, 75(5), 2673-2718

Shen, J., Yu, J., Zhao, S., 2017. Investor sentiment and economic forces. *Journal of Monetary Economics*, 86, 1-21

Shleifer, A., Summers, L. H., 1990. The Noise Trader Approach to Finance. *Journal of Economic Perspectives*, 4(2), 19-33

Shleifer, A., Vishny, R., 1997. The limits of arbitrage. *Journal of Finance*, 52(1), 35-55

Stambaugh, R. F., Yuan, Y., 2017. Mispricing factors. *Review of Financial Studies*, 30(4), 1270-1315

Stambaugh, R., Yu, J., Yuan, Y., 2012.The short of it: investor sentiment and anomalies. *Journal of Financial Economics*, 104, 288–302

Stambaugh, R., Yu, J., Yuan, Y., 2015. Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle. *Journal of Finance*, 70(5), 1903-1948

Wu, S., Lee, W., 2015. Intertemporal risk–return relationships in bull and bear markets. International *Review of Economics and Finance*, 38, 308-325.

Yu, J., 2013. A sentiment-based explanation of the forward premium puzzle. *Journal of Monetary Economics*, 60, 474-491

Yu, J., Yuan, Y., 2011. Investor sentiment and the mean-variance relation. *Journal of Financial Economics*, 100(2), 367–381

Table 1: Descriptive Statistics

The table exhibits the descriptive statistics of the weekly BAB returns together with the log-difference of the weekly Google Search Volumes on the corresponding index. The data for BAB returns is from the AQR Capital Management website (www.aqr.com), while the data for online searches is from Google Trends, using the name of the index (e.g. Dow Jones Industry Average) and restricted to US searches. The data ranges from 01.01.2004 to 12.31.2019.

	BAB	Online Searches		
		S&P	Dow Jones	Stock Market
Mean	0.0013	0.0010	0.0019	0.0001
Median	0.0019	0.0000	0.0000	0.0000
Std Dev	0.0142	0.1795	0.2406	0.2148
Skewness	-0.7248	0.3962	1.4717	0.9454
Kurtosis	8.2008	3.0186	9.6783	15.7300

Table 2: VAR analysis between investor attention and BAB

In the upper section, the table shows the estimates of the VAR analysis between investor attention (S) on three market indexes and BAB returns. Investor attention is measured by the log-difference of the GSV on the corresponding index (e.g. Dow Jones Industry Average). The number of lags was based on the AIC. The standard errors are reported below the coefficients. The data is weekly from 01.01.2004 to 12.31.2019. In the lower section, the table reports the p-values of the Granger Causality tests beside the respective chi-squared between brackets. N = 830

	S_{t-1}	S_{t-2}	S_{t-3}	S_{t-4}	BAB_{t-1}	BAB_{t-2}	BAB_{t-3}	BAB_{t-4}	Intercept	Adj. R^2
S&P 500										
BAB_t	-0.010***	-0.004	-0.001	-0.002	-0.162***	0.032	0.096***	0.035	0.001***	5.2%
	(0.003)	(0.003)	(0.003)	(0.003)	(0.035)	(0.035)	(0.035)	(0.034)	(0.000)	
S_t	-0.389***	-0.235***	-0.149***	-0.150***	0.385	0.348	-0.297	0.337	0.001***	14.9%
	(0.034)	(0.037)	(0.037)	(0.035)	(0.416)	(0.419)	(0.419)	(0.413)	(0.006)	
Dow Jones										
BAB_t	-0.007***	-0.005**	-0.004*	-0.004**	-0.168***	0.026	0.101***	0.049	0.001***	5.8%
	(0.002)	(0.002)	(0.002)	(0.002)	(0.035)	(0.035)	(0.035)	(0.034)	(0.000)	
S_t	-0.258***	-0.133***	-0.043	-0.102***	0.512	0.970	0.263	-0.041	0.001	7.7%
	(0.035)	(0.036)	(0.036)	(0.035)	(0.582)	(0.586)	(0.586)	(0.578)	(0.863)	
Stock Market										
BAB_t	-0.008***	-0.003	-0.001	-0.004*	-0.167***	0.031	0.092	0.041	0.001***	5.2%
	(0.002)	(0.002)	(0.002)	(0.002)	(0.035)	(0.035)	(0.035)	(0.034)	(0.000)	
S_t	-0.208***	-0.191***	-0.081	-0.073	0.798	0.110	0.228	-0.216	-0.001	6.7%
	(0.035)	(0.035)	(0.035)	(0.035)	(0.521)	(0.526)	(0.526)	(0.518)	(0.007)	
Granger causality test										
			S&P 500			Dow Jones			Stock Market	
S does not cause BAB			0.01	[13.09]		0.00	[18.60]		0.01	[13.38]
BAB does not cause S			0.54	[3.08]		0.53	[3.18]		0.60	[2.73]

***Significant at 1%, **5% and *10% levels

Table 3: Influence of investor attention on BAB returns

The table shows the results for the regression between BAB returns and investor attention on the given market index (S_i) according to the following models:

$$BAB_t = a_0 + \sum_{n=1}^{4} b_{i,n} S_{i,t-n} + \sum_{n=1}^{4} c_n BAB_{t-n} + \varepsilon_t \quad (3)$$

$$BAB_t = a_0 + \sum_{n=1}^{4} b_{i,n} S_{i,t-n} + \emptyset X_t + \sum_{n=1}^{4} d_n BAB_{t-n} + \varepsilon_t \quad (4)$$

Where X_t is a vector containing the following control variables: Investor Sentiment, captured by the proportion of optimistic investor of the American Association of Individual Investors (AAII) weekly survey, Fama-French 5 factors model and four Business Cycle variables. The latter group is composed of detrended riskless rate (RREL), dividend yield (DY), default spread (DEF) and term spread (TERM). RREL is the difference between the 1-month Treasury bill rate and its 12-month moving average. DEF is the difference between the yields on BAA- and AAA-rated corporate bonds. TERM is the difference between the yields on the ten-year Treasury bond and the three-month Treasury bill, and DY is constructed using the twelve-month dividend data. The sample period is from 01.01.2004 to 12.31.2019. The numbers in parentheses are t-statistics from the Newey-West standard error estimator. The number of lags for S_i and BAB was based on the AIC. The coefficients in bold are significant at the 5% level. N=830.

	S&P		Dow Jones		Stock Market	
	(3)	(4)	(3)	(4)	(3)	(4)
S_{t-1}	**-0.0102**	**-0.0092**	**-0.0074**	**-0.0066**	**-0.0076**	**-0.0066**
	(-2.24)	(-2.63)	(-2.18)	(-2.81)	(-2.11)	(-2.58)
S_{t-2}	-0.0038	-0.0045	**-0.0046**	**-0.0047**	-0.0026	-0.0030
	(-1.14)	(-1.66)	(-1.99)	(-2.48)	(-1.07)	(-1.55)
S_{t-3}	-0.0009	-0.0026	-0.0036	-0.0039	-0.0015	-0.0024
	(-0.28)	(-0.92)	(-1.5)	(0.06)	(-0.64)	(-1.17)
S_{t-4}	-0.0021	-0.0026	-0.0043	-0.0042	-0.0040	-0.0037
	(-0.57)	(-0.81)	(-1.63)	(-1.71)	(-1.20)	(-1.23)
SENT		**0.027**		**0.027**		**0.027**
		(4.10)		(4.08)		(4.09)
MRP		-0.019		-0.034		-0.029
		(-0.43)		(-0.80)		(-0.64)
SMB		**-0.201**		**-0.193**		**-0.190**
		(-3.36)		(-3.26)		(-3.19)
HML		**-0.155**		**-0.145**		**-0.145**
		(-2.13)		(-2.00)		(-2.01)
RMW		**0.276**		**0.269**		**0.276**
		(3.84)		(3.86)		(3.98)
CMA		0.118		0.110		0.110
		(1.29)		(1.20)		(1.22)
Intercept	**0.001**	-0.012	**0.001**	-0.014	**0.001**	-0.013
	(2.57)	(-1.37)	(2.63)	(-1.53)	(2.55)	(-1.41)
Business Cycle Control	No	Yes	No	Yes	No	Yes
Lagged BAB Control	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	4.3%	19.1%	4.9%	19.6%	4.3%	19.0%

Table 4: Quantile regression of BAB on Attention to the S&P-500

The table presents the results for the estimators from the quantile regression of BAB on Attention (S) to the S&P-500, employing model (4). The corresponding quantile is identified in the first row. The sample period is from 01.01.2004 to 12.31.2019. The numbers in parentheses are the block-bootstrap t-statistics. The coefficients in bold are significant at 5%. N = 830.

	q05	q10	q25	q50	q75	q90	q95
S_{t-1}	-0.0094	-0.0070	**-0.0098**	**-0.0086**	**-0.0089**	**-0.0067**	**-0.0086**
	(-1.57)	(-1.75)	(-3.81)	(-3.26)	(-2.26)	(-2.38)	(-2.42)
S_{t-2}	0.0011	-0.0011	0.0004	-0.0044	-0.0067	**-0.0066**	**-0.0138**
	(0.2)	(-0.25)	(0.16)	(-1.57)	(-1.83)	(-2.6)	(-4)
S_{t-3}	-0.0008	-0.0025	-0.0006	0.0009	-0.0009	-0.0033	-0.0066
	(-0.12)	(-0.71)	(-0.22)	(0.34)	(-0.26)	(-0.89)	(-1.56)
S_{t-4}	0.0032	0.0011	0.0023	-0.0039	-0.0034	-0.0053	-0.0062
	(0.64)	(0.26)	(0.92)	(-1.56)	(-1.3)	(-1.94)	(-1.37)
SENT	**0.022**	**0.021**	**0.020**	**0.023**	**0.024**	**0.019**	0.019
	(2.15)	(2.88)	(3.41)	(4.31)	(3.37)	(3.32)	(1.89)
MRP	0.004	0.014	-0.025	-0.061	0.328	**-0.125**	**-0.147**
	(0.14)	(0.31)	(-1.12)	(-1.85)	(1.08)	(-2.61)	(-3.24)
SMB	-0.046	-0.078	**-0.105**	**-0.114**	-0.119	**-0.162**	**-0.207**
	(-0.86)	(-1.02)	(-2.41)	(-2.77)	(-1.78)	(-3.64)	(-2.43)
HML	-0.150	**-0.174**	**-0.219**	**-0.134**	-0.107	0.010	0.031
	(-1.36)	(-2.42)	(-5.25)	(-2.25)	(-1.42)	(0.1)	(0.25)
RMW	**0.399**	**0.328**	**0.283**	**0.200**	**0.248**	**0.229**	0.194
	(4.05)	(3.73)	(5.47)	(3.17)	(3.48)	(3.29)	(1.85)
CMA	-0.123	-0.037	0.018	0.019	0.048	-0.020	-0.030
	(-1.19)	(-0.37)	(0.35)	(0.2)	(0.46)	(-0.21)	(-0.19)
Intercept	-0.020	-0.006	0.000	-0.006	-0.013	**-0.020**	-0.017
	(0.02)	(0)	(0)	(-0.84)	(-1.37)	(-2.19)	(-1.11)
Business Cycle Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lagged BAB Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	7.4%	7.5%	12.8%	17.0%	11.8%	2.4%	1.5%

Table 5: Influence of investor attention on BAB returns controlled by liquidity constraints

The table shows the results for the regression between BAB returns and investor attention to the given market index (S_i) according to the following models:

$$BAB_t = a_0 + \sum_{n=1}^{2} b_{i,n} S_{i,t-n} + \sum_{n=1}^{2} c_n TED_{t-n} + \sum_{n=1}^{2} d_n BAB_{t-n} + \varepsilon_t \tag{5}$$

$$BAB_t = a_0 + \sum_{n=1}^{2} b_{i,n} S_{i,t-n} + \emptyset X_t + \sum_{n=1}^{2} c_n TED_{t-n} + \sum_{n=1}^{2} d_n BAB_{t-n} + \varepsilon_t \tag{6}$$

$$BAB_t = a_0 + \sum_{n=1}^{2} b_{i,n} S_{i,t-n} + \sum_{n=1}^{2} c_n TED_{t-n} + \sum_{n=1}^{2} d_{i,n} S_{i,t-n} \times TED_{t-n} + \sum_{n=1}^{2} d_n BAB_{t-n} + \varepsilon_t \tag{7}$$

$$BAB_t = a_0 + \sum_{n=1}^{2} b_{i,n} S_{i,t-n} + \emptyset X_t + \sum_{n=1}^{2} c_n TED_{t-n} + \sum_{n=1}^{2} d_{i,n} S_{i,t-n} \times TED_{t-n} + \sum_{n=1}^{2} e_n BAB_{t-n} + \varepsilon_t \tag{8}$$

Where TED_t is the TED spread in week t. The remaining variables were defined in Section 4. The interaction variables in Models (7) and (8) aim to capture the effect of attention conditional to the TED spread. The sample period is from 01.01.2004 to 12.31.2019. The numbers in parentheses are the t-statistics from the Newey-West standard error estimator. The number of lags was determined using the AIC. The coefficients in bold are significant at 10%. N = 830.

	S&P				Dow Jones				Stock Market			
	(5)	(6)	(7)	(8)	(5)	(6)	(7)	(8)	(5)	(6)	(7)	(8)
S_{t-1}	**-0.0080**	**-0.0070**	0.0009	0.0009	**-0.0051**	**-0.0045**	0.0022	0.0011	**-0.0058**	**-0.0050**	0.0031	0.0006
	(-2.02)	(-2.01)	(0.28)	(0.28)	(-1.78)	(-1.95)	(0.79)	(0.40)	(-2.07)	(-2.11)	(1.06)	(0.19)
S_{t-2}	-0.0011	-0.0016	**-0.0079**	**-0.0082**	-0.0018	-0.0022	-0.0038	-0.0037	-0.0009	-0.0013	-0.0035	-0.0034
	(-0.4)	(-0.64)	(-2.32)	(-2.68)	(-0.91)	(-1.19)	(-1.59)	(-1.60)	(-0.45)	(-0.71)	(-1.26)	(-1.18)
TED_{t-1}	**-2.12**	**-1.76**	**-2.30**	**-1.97**	**-2.10**	**-1.75**	**-1.90**	**-1.59**	**-2.12**	**-1.78**	**-1.91**	**-1.70**
	(-3.93)	(-3.38)	(-4.11)	(-3.65)	(-3.85)	(-3.31)	(-3.23)	(-2.72)	(-3.95)	(-3.40)	(-3.15)	(-2.83)
TED_{t-2}	**1.43**	**1.33**	**1.59**	**1.43**	**1.41**	**1.31**	**1.19**	**1.08**	**1.43**	**1.33**	**1.18**	**1.16**
	(3.25)	(3.28)	(3.25)	(2.98)	(3.18)	(3.22)	(2.47)	(2.20)	(3.26)	(3.27)	(2.24)	(2.14)
$S_{t-1}.TED_{t-1}$			**-1.93**	**-1.84**			**-1.26**	**-1.06**			**-1.72**	**-1.14**
			(-2.45)	(-2.07)			(-2.74)	(-2.22)			(-3.58)	(-2.36)
$S_{t-2}.TED_{t-2}$			**1.64**	**1.49**			0.40	0.26			0.53	0.40
			(2.17)	(1.96)			(0.88)	(0.53)			(0.94)	(0.62)

Intercept	**0.005**	-0.010	**0.005**	-0.011	**0.005**	-0.010	**0.005**	-0.012	**0.005**	-0.010	**0.005**	-0.011
	(4.82)	(-1.15)	(5.40)	(-1.27)	(4.81)	(-1.17)	(5.15)	(-1.30)	(4.86)	(-1.16)	(5.56)	(-1.23)
Sentiment Control	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
FF-5 Factors Control	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Business Cycle Control	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Lagged BAB Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	10.5%	22.2%	14.0%	25.1%	10.2%	22.1%	12.4%	23.4%	10.2%	22.0%	12.8%	23.4%

Table 6: Intertemporal influence of liquidity constraints on the BAB-Attention relation

The table presents the estimates of the following models:

$$bS_{i,t} = a_0 + a_1 TED_{t-1} + \varepsilon_t \quad (9)$$

$$bS_{i,t} = a_0 + a_1 TED_{t-1} + \emptyset X_{t-1} + \varepsilon_t \quad (10)$$

Where $bS_{i,t}$ is the time-varying coefficient of the BAB-Attention relation, generated employing a moving window regression of BAB_t on one-lag Attention ($S_{i,t-1}$) together with BAB_{t-1} to control for autocorrelation. The length of the moving window regression is 30 weeks. The remaining variables of Models (9) and (10) were defined in Sections 4 and 5. The sample period is from 01.01.2004 to 12.31.2019. The numbers in parentheses are the t-statistics from the Newey-West standard error estimator. The coefficients in bold are significant at 5%. N = 830.

	S&P		Dow Jones		Stock Market	
	(9)	(10)	(9)	(10)	(9)	(10)
TED_{t-1}	**-2.113**	**-2.374**	**-1.416**	**-1.765**	**-2.184**	**-2.309**
	(-4.98)	(-4.47)	(-6.60)	(-4.76)	(-5.83)	(-3.91)
$SENT_{t-1}$		-0.001		0.007		0.002
		(-0.04)		(0.64)		(0.87)
MRP_{t-1}		0.016		0.000		0.036
		(0.52)		(0.01)		(1.14)
SMB_{t-1}		-0.004		0.000		0.016
		(-0.07)		(0.00)		(0.28)
HML_{t-1}		0.029		0.037		-0.068
		(0.41)		(0.78)		(-1.14)
RMW_{t-1}		-0.017		0.025		-0.105
		(-0.23)		(0.46)		(-1.45)
CMA_{t-1}		-0.039		0.008		0.128
		(-0.37)		(0.10)		(1.27)
Intercept	0.002	0.001	**0.004**	-0.010	0.003	**-0.042**
	(0.86)	(0.03)	(2.66)	(-0.55)	(1.67)	(-2.03)
Business Cycle Control	No	Yes	No	Yes	No	Yes
Adj. R^2	18.0%	21.3%	13.8%	16.4%	18.6%	23.8%

Table 7: BAB returns and concurrent explanatory variables

The table presents the estimates of the BAB regression on different explanatory variables, controlling for the Fama-French 5-factor model. The variables are Investor Attention, (measured by the log-difference of the weekly GSV on the Dow Jones index), Liquidity Constraints (measured by the TED spread), Investor Sentiment (measured by the percentage of bullish respondents of the AAII weekly survey), Lottery Demand (measured by the FMAX factor of Bali et al., 2017) and Volatility (measured by the conditional variance using a GARCH (1,1) model on the weekly returns of the Dow Jones). The sample period is from 01.01.2004 to 12.31.2019. The numbers in parentheses are the t-statistics from the Newey-West standard error estimator. The number of lags was determined using the AIC. The coefficients in bold are significant at 5%. N = 830.

	(1)	(2)	(3)	(4)	(5)	(6)
S_{t-1}	**-0.0060**					**-0.0054**
	(-2.39)					(-2.41)
TED_{t-1}		**-0.8759**				**-0.8398**
		(-3.63)				(-3.92)
$SENT_{t-1}$			**0.0278**			**0.0215**
			(4.44)			(3.48)
$FMAX_{t-1}$				**0.0007**		0.0002
				(3.66)		(1.15)
VOL_{t-1}					-2.8338	0.7653
					(-1.34)	(0.32)
MRP	-0.009	-0.0100	0.0080	0.0033	0.0127	-0.0292
	(-0.19)	(-0.22)	(0.15)	(0.07)	(0.23)	(-0.72)
SMB	**-0.178**	**-0.1906**	**-0.1889**	**-0.2021**	**-0.1952**	**-0.1917**
	(-2.87)	(-2.88)	(-2.95)	(-3.11)	(-2.85)	(-3.15)
HML	-0.134	**-0.155**	**-0.147**	**-0.158**	**-0.158**	**-0.143**
	(-1.84)	(-2.16)	(-1.97)	(-2.10)	(-2.26)	(-2.02)
RMW	**0.258**	**0.279**	**0.275**	**0.296**	**0.287**	**0.268**
	(3.29)	(4.05)	(3.75)	(3.97)	(3.94)	(3.78)
CMA	0.120	0.128	0.122	0.159	0.146	0.119
	(1.22)	(1.40)	(1.26)	(1.60)	(1.56)	(1.34)
Intercept	**0.001**	**0.005**	**-0.009**	**0.001**	**0.003**	-0.003
	(2.28)	(5.23)	(-3.59)	(2.81)	(2.35)	(-1.27)
Lagged BAB Control	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	11.6%	16.2%	13.4%	12.1%	12.3%	18.9%

Table 8: VAR between investor attention and concurrent BAB's explanatory variables

The table presents the estimates of the VAR analysis between Investor Attention to the Dow Jones index (S) and four concurrent variables that are plausibly related to BAB performance. These variables are Liquidity Constraints (measured by the TED spread), Investor Sentiment (measured by the percentage of bullish respondents of the AAII weekly survey), Lottery Demand (measured by the FMAX factor from Bali et al., 2017) and Volatility (measured by the conditional variance using a GARCH (1,1) model on the weekly returns of the Dow Jones). In the VAR model, these variables are named by X. The sample period is from 01.01.2004 to 12.31.2019. The numbers in parentheses are the standard errors. The number of lags was determined using the AIC. In the lower section, the table reports the p-values of the Granger Causality tests beside the respective chi-squared in brackets. N = 830.

	Liquidity Constraints		**Investor Sentiment**		**Lottery Demand**		**Volatility**	
	S_t	TED_t	S_t	$SENT_t$	S_t	$FMAX_t$	S_t	VOL_t
S_{t-1}	-0.256***	0.001***	-0.262***	-0.028***	-0.256***	-0.173*	-0.405***	$3.6\text{x}10^{-4}$***
	(0.035)	($2.1\text{x}10^{-4}$)	(0.035)	(0.009)	(0.035)	(0.094)	(0.034)	($0.4\text{x}10^{-4}$)
S_{t-2}	-0.135***	0.001***	-0.141***	-0.009	-0.140***	-0.233**	-0.253***	$2.9\text{x}10^{-4}$***
	(0.036)	($2.2\text{x}10^{-4}$)	(0.036)	(0.010)	(0.036)	(0.097)	(0.038)	($0.5\text{x}10^{-4}$)
S_{t-3}	-0.050	$4.3\text{x}10^{-4}$**	-0.053	0.016	-0.057	-0.025	-0.154	$1.3\text{x}10^{-4}$***
	(0.037)	($2.2\text{x}10^{-4}$)	(0.036)	(0.010)	(0.036)	(0.097)	(0.038)	($0.5\text{x}10^{-4}$)
S_{t-4}	-0.105***	0.001***	-0.103***	0.015	-0.107***	-0.125	-0.129***	$3.3\text{x}10^{-4}$***
	(0.035)	($2.1\text{x}10^{-4}$)	(0.035)	(0.009)	(0.035)	(0.094)	(0.035)	($0.4\text{x}10^{-4}$)
X_{t-1}	0.592	0.800***	-0.156	0.552***	-0.032**	1.785***	-54.351**	0.810***
	(5.767)	(0.035)	(0.128)	(0.035)	(0.013)	(0.035)	(26.96)	(0.033)
X_{t-2}	1.006	0.151***	0.033	0.103***	0.058**	-0.948***	19.52	-0.059
	(7.399)	(0.045)	(0.146)	(0.039)	(0.026)	(0.071)	(34.557)	(0.043)
X_{t-3}	-4.316	-0.002	0.070	0.006	-0.036	-0.020	-52.974	0.203***
	(7.397)	(0.045)	(0.146)	(0.039)	(0.026)	(0.071)	(34.281)	(0.043)
X_{t-4}	0.843	-0.005	0.186	0.097***	0.010	0.094***	36.916	-0.057*
	(5.738)	(0.035)	(0.126)	(0.034)	(0.013)	(0.035)	(26.881)	(0.033)
Intercept	0.012	$2.3\text{x}10^{-4}$***	-0.047	0.091***	0.003	-0.032***	0.028	$0.5\text{x}10^{-4}$***
	(0.012)	($0.7\text{x}10^{-4}$)	(0.043)	(0.012)	(0.420)	(0.022)	(0.010)	($0.1\text{ x}10^{-4}$)
Adj. R^2	7.5%	88.8%	8.0%	47.7%	8.2%	95.8%	9.5%	79.4%
Granger causality test								
X does not cause S	0.87	[1.25]	0.22	[5.74]	0.12	[7.42]	0.01	[19.27]
S does not cause X	0.00	[23.41]	0.01	[14.86]	0.06	[9.15]	0.00	[135.32]

***Significant at 1%, **5% and *10% levels

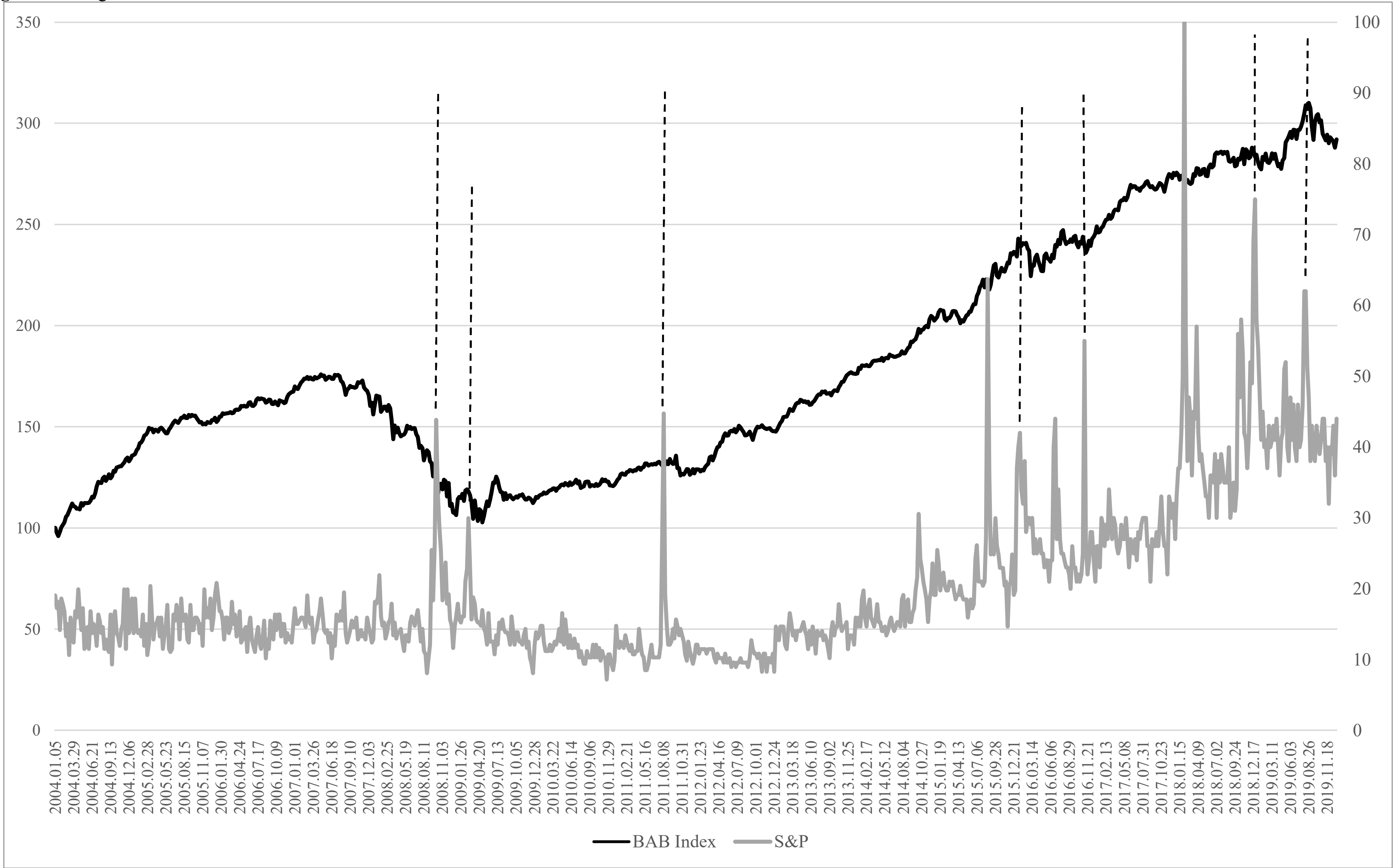
Table 9: BAB and Investor Attention relation in the remaining G7 countries

The table shows the results for the regression between local BAB returns and investor attention to the given stock market index (S_i) in G7 countries excluding the US. The GSV are from the terminology of the local stock market index (e.g., "FTSE" for the UK), with the exception of Italy, for which I use "stock market" due to missing data. The factors of the Fama-French model are those for developed countries excluding the US and are available on Kenneth French's website. The remaining control variables were defined in the previous sections. The sample period is from 01.01.2004 to 12.31.2019. The numbers in parentheses are t-statistics from the Newey-West standard error estimator. The coefficients in bold are significant at the 10% level. N=830.

	UK		GER		JPN		FRA		CAN		ITA	
	(3)	(4)	(3)	(4)	(3)	(4)	(3)	(4)	(3)	(4)	(3)	(4)
S_{t-1}	**-0.0099**	**-0.0098**	-0.0010	-0.0022	-0.0021	**-0.0026**	**-0.0015**	**-0.0015**	-0.0017	0.0001	-0.0011	-0.0011
	(-2.56)	(-3.41)	(-0.47)	(-1.05)	(-1.51)	(-1.88)	(-1.90)	(-2.04)	(-0.75)	(0.03)	(-1.48)	(-1.60)
S_{t-2}	**-0.0070**	**-0.0062**	-0.0014	-0.0025	0.0018	0.0007	-0.0008	**-0.0017**	0.0007	0.0018	**-0.0015**	**-0.0015**
	(-1.82)	(-1.72)	(-0.59)	(-1.04)	(0.95)	(0.36)	(-0.75)	(-1.88)	(0.29)	(0.93)	(-1.79)	(-1.99)
S_{t-3}	**-0.0074**	**-0.0066**	**-0.0052**	**-0.0056**	0.0009	0.0004	-0.0012	**-0.0018**	0.0019	0.0021	-0.0002	-0.0006
	(-2.09)	(-1.96)	(-2.15)	(-2.36)	(0.41)	(0.24)	(-1.21)	(-1.95)	(0.88)	(1.18)	(-0.18)	(-0.74)
S_{t-4}	-0.0026	-0.0017	-0.0036	-0.0036	0.0001	0.0007	-0.0009	-0.0010	0.0011	0.0015	-0.0008	-0.0009
	(-0.63)	(-0.58)	(-1.41)	(-1.47)	(0.09)	(0.52)	(-0.99)	(-1.33)	(0.54)	(0.91)	(-0.93)	(-1.25)
MRP		-0.067		-0.098		**-0.269**		**-0.155**		**0.235**		-0.039
		(-1.18)		(-1.38)		(-6.14)		(-2.71)		(6.78)		(-1.11)
SMB		**0.775**		0.236		0.084		**0.682**		**0.264**		**0.590**
		(7.21)		(1.47)		(0.86)		(6.67)		(2.87)		(6.26)
HML		-0.105		**-0.337**		**-0.328**		-0.268		-0.040		**-0.395**
		(-0.87)		(-1.70)		(-2.29)		(-1.64)		(-0.37)		(-3.34)
RMW		-0.015		-0.204		0.268		-0.135		-0.014		0.037
		(-0.10)		(-0.74)		(1.33)		(-0.63)		(-0.08)		(0.20)
CMA		**0.287**		**0.485**		**0.858**		**0.634**		**0.382**		**0.309**
		(1.95)		(2.49)		(5.07)		(3.44)		(2.77)		(2.16)
Intercept	0.000	**0.005**	**0.003**	**0.006**	**0.002**	**0.004**	**0.004**	**0.007**	**0.003**	**0.005**	0.001	**0.002**
	(-0.05)	(4.6)	(2.88)	(3.48)	(2.12)	(3.45)	(4.14)	(4.97)	(4.12)	(4.72)	(1.16)	(2.14)
Business Cycle Control	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Lagged TED spread Control	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

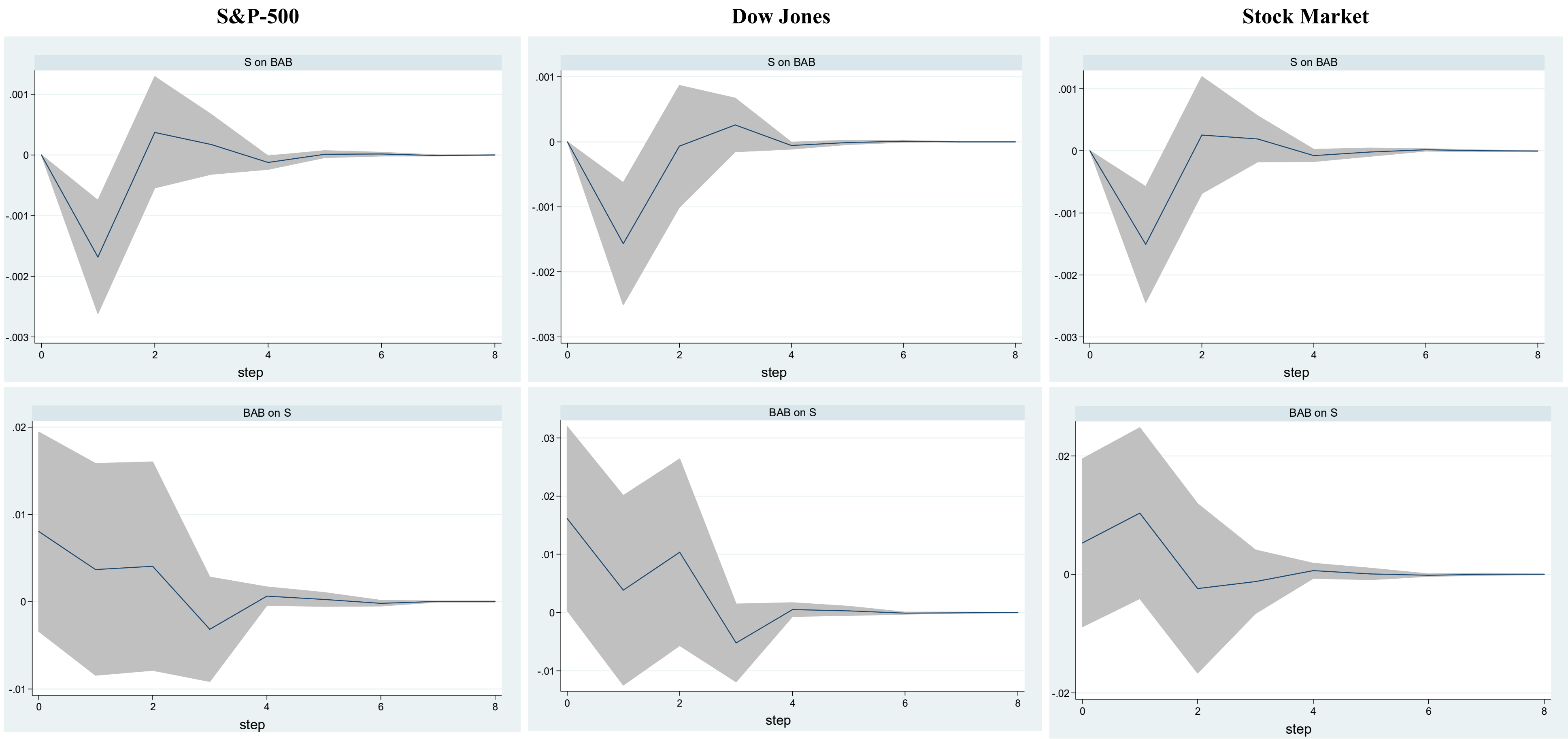
Lagged BAB Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	3.2%	29.4%	4.0%	8.5%	1.8%	27.6%	0.7%	22.9%	6.4%	13.6%	1.5%	19.7%

Figure 1: Google Search Volumes of S&P-500 searches and BAB cumulative returns.



Note: The figure shows the weekly Google Search Volumes for searches on the S&P-500 (gray line, right axis) and the time-series of the cumulative returns of BAB (black line, left axis), starting from 100. The dataset is from 01.01.2004 to 12.31.2019. The dashed vertical lines indicate GSV peaks that are followed by BAB decreases.

Figure 2: Impulse response function for the VAR of BAB and Investor Attention (S) to market indices.



Note: The figure plots the impulse response to one standard deviation of the VAR model between BAB returns and Investor Attention (S) on the S&P-500 (left column), Dow Jones (medium column) and Stock Market (right column). The first row exhibits the influence of attention on BAB, and the bottom row the influence in the opposite direction. The confidence interval of 95% is represented by the gray area. The data are weekly from 01.01.2004 to 12.31.2019.

Appendix A

This appendix contains some of the main results using the log-difference between the GSV on the given date and the median of the previous 8 weeks to calculate investor attention (*S*), following the classical approach employed in Da et al., 2011. As can be seen in Tables A.1 and A.3, the results are not biased by this choice.

Table A.1: VAR analysis between investor attention and BAB
The table presents the same analysis summarized in Table 2 using the alternative operationalization for S_i described above. The number of lags was based on the AIC. The standard errors are reported below the coefficients. The data is weekly from 01.01.2004 to 12.31.2019. In the lower section, the table reports the p-values of the Granger Causality tests beside the respective chi-squared in brackets. N = 830.

	S_{t-1}	S_{t-2}	S_{t-3}	S_{t-4}	BAB_{t-1}	BAB_{t-2}	BAB_{t-3}	BAB_{t-4}	Intercept	Adj. R^2
S&P 500										
BAB_t	-0.01***	0.006*	0.002	-0.001	-0.173***	0.019**	0.087	0.027	0.001**	5.3%
	(0.003)	(0.003)	(0.003)	(0.003)	(0.035)	(0.035)	(0.035)	(0.035)	(0.000)	
S_t	0.514***	0.065*	-0.003	-0.135***	0.447	0.446	-0.327	0.21	0.007	30.2%
	(0.035)	(0.039)	(0.039)	(0.035)	(0.416)	(0.42)	(0.419)	(0.412)	(0.006)	
Dow Jones										
BAB_t	-0.008***	0.004	0.001	-0.002	-0.178***	0.013	0.086**	0.035	0.001***	6.1%
	(0.002)	(0.002)	(0.002)	(0.002)	(0.035)	(0.035)	(0.035)	(0.034)	(0.000)	
S_t	0.661***	0.061	0.016	-0.178***	0.569	0.762	0.039	-0.325	0.015*	47.2%
	(0.034)	(0.042)	(0.042)	(0.035)	(0.585)	(0.592)	(0.591)	(0.579)	(0.008)	
Stock Market										
BAB_t	-0.009***	0.006**	0.000	-0.002	-0.179***	0.018	0.08**	0.03	0.001	5.6%
	(0.002)	(0.003)	(0.003)	(0.002)	(0.035)	(0.035)	(0.035)	(0.035)	(0.000)	
S_t	0.712***	-0.057	0.032	-0.106*	0.893*	0.028	0.031	-0.315	0.011	45.3%
	(0.035)	(0.043)	(0.043)	(0.035)	(0.522)	(0.529)	(0.527)	(0.519)	(0.007)	
Granger causality test										
			S&P 500			Dow Jones			Stock Market	

S does not cause BAB	0.01	[13.14]	0.00	[19.92]	0.00	[15.46]
BAB does not cause S	0.49	[3.41]	0.65	[2.47]	0.52	[3.26]

Table A.2: Influence of investor attention on BAB returns

The table presents the analysis summarized in Table 3 using the alternative operationalization for S_i described above. The sample period is from 01.01.2004 to 12.31.2019. The numbers in parentheses are the t-statistics from the Newey-West standard error estimator. The number of lags for S_i and BAB was based on the AIC. The coefficients in bold are significant at the 5% level. N=830.

	S&P		Dow Jones		Stock Market	
	(3)	(4)	(3)	(4)	(3)	(4)
S_{t-1}	**-0.0103**	**-0.0097**	**-0.0085**	**-0.0078**	**-0.0088**	**-0.0078**
	(-2.17)	(-2.67)	(-2.34)	(-3.22)	(-2.21)	(-2.86)
S_{t-2}	0.0062	0.0043	0.0037	0.0025	0.0057	0.0041
	(1.41)	(1.11)	(1.16)	(0.98)	(1.92)	(1.57)
S_{t-3}	0.0021	0.0005	0.0006	0.0001	-0.0003	-0.0012
	(0.64)	(0.19)	(0.24)	(0.03)	(-0.79)	(-0.39)
S_{t-4}	-0.0014	-0.0015	-0.0019	-0.0020	-0.0020	-0.0018
	(-0.40)	(-0.49)	(-0.92)	(-1.06)	(-0.79)	(-0.78)
SENT		**0.027**		**0.027**		**0.028**
		(3.95)		(4.03)		(4.04)
MRP		-0.017		-0.031		-0.028
		(-0.39)		(-0.75)		(-0.64)
SMB		**-0.201**		**-0.194**		**-0.191**
		(-3.41)		(-3.31)		(-3.19)
HML		**-0.163**		**-0.150**		**-0.147**
		(-2.26)		(-2.11)		(-2.07)
RMW		**0.271**		**0.271**		**0.275**
		(3.82)		(3.95)		(4.00)
CMA		0.129		0.123		0.127
		(1.41)		(1.36)		(1.39)
Intercept	**0.001**	-0.016	**0.001**	-0.017	**0.001**	-0.017
	(2.40)	(-1.66)	(2.70)	(-1.92)	(2.61)	(-1.80)
Business Cycle Control	No	Yes	No	Yes	No	Yes
Lagged BAB Control	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	4.4%	19.2%	5.2%	20.0%	4.7%	19.4%

Appendix B

This appendix contains the Spearman correlation between the variables of Model (4) in the case of the S&P-500 index, as presented in Table B1. The results demonstrate that investor attention S exhibits a small correlation with the remaining control variables, suggesting that they capture different dimensions of BAB. The asterisk indicates significance at the 5% level.

Table B.1: Spearman correlation between BAB, investor attention and control variables.

	BAB	**S**	**SENT**	**MRP**	**SMB**	**HML**	**RMW**	**CMA**	**DEF**	**RREL**	**TERM**	**DY**
BAB	1.00											
S	0.08*	1.00										
SENT	0.20*	0.02	1.00									
MRP	-0.21*	-0.13*	0.10*	1.00								
SMB	-0.15*	-0.04	0.08*	0.32*	1.00							
HML	-0.12*	0.00	0.05	0.17*	-0.04	1.00						
RMW	0.26*	0.08*	-0.02	-0.23*	-0.19*	-0.12*	1.00					
CMA	-0.04	0.08*	0.03	-0.06	-0.04	0.33*	-0.09*	1.00				
DEF	-0.09*	-0.02	-0.17*	0.02	0.02	-0.02	0.01	0.08*	1.00			
RREL	0.05	0.02	0.09*	0.01	0.03	0.08*	-0.02	-0.01	-0.29*	1.00		
TERM	0.05	-0.03	0.09*	0.01	0.04	0.02	0.04	-0.02	0.12*	-0.25*	1.00	
DY	-0.07*	-0.02	-0.39*	-0.02	-0.01	-0.04	0.01	0.05	0.64*	-0.34*	0.29*	1.00

Online Appendix

See the online appendix at: http://bit.ly/bab-attention.