

Evidence from Google Trends data: Can low Investor Sentiment still predict inefficiencies in the Stock Market?

Introduction

Trading strategies revolve around several sources of edge, ranging from discretionary macro ideas, fundamental based stock picking to more sophisticated ones such as factors exposure or high-frequency algorithms. One area which the current literature pays little attention to is sentiment trading, which can be established in several ways. In this article, we explore an academic paper [1] that focuses its attention on Google Trends search volume data, providing trading signals for individual stocks in the S&P 500, considered underpriced according to retail investor sentiment. The authors backtest the strategy from 2004 to October 2014, thus we are performing a continuation of this backtest for the past 10 years. We will focus our work on an equally-weighted strategy with weekly-rebalancing, performing robustness checks as well as a factor regression to detect abnormal returns generated by it. We will then analyse the changes encountered when switching to a longer rebalancing period of one month. Lastly, we will present a different portfolio weighting scheme than the one used by the paper and explain our findings.

Methodology

Our work starts by defining the trading universe. According to the paper studied, there are 122 liquid stocks in the S&P 500 index which have enough search data provided by Google Trends. As a proxy for liquidity and due to time constraints, we focused on the first 122 stocks ranked by current market cap. Out of them, we found 82 of them with available data, for which we manually imported individual weekly search data, for two sub-periods of 5 years. This is the maximum period allowed by Google Trends in order to get weekly data.

There are two aspects worth mentioning here. First, the most efficient way as provided by the current literature to get accurate search results is by typing the firm's name followed by the word "stock" in the navigator. This can be easily explained by the fact that people might search for Citigroup or Netflix due to pure interest in banking services or watching movies. Secondly, instead of showing the actual number of searches, Google provides a scaled score (GSV), ranging from 0 to 100, with 100 representing the maximum search volume over the selected period. This way, the strategy is based upon a relative measure rather than an absolute one. To get our signal, we will compare the current week's score to the median of the past 8 weeks, by calculating the abnormal search volume (ASV) according to the following formula:

$$ASV_t = \ln(GSV_t) - \ln(\text{Med}(GSV_{t-1}, \dots, GSV_{t-8}))$$

We distinguish between two different strategies. The "Main" one enters long a stock on Monday's open price if past week's ASV score is negative, indicating abnormally low attention. The "Loser" strategy requires past week's return of the stock to be negative as an additional condition, which indicates an increased probability of underpricing. In the first phase, both strategies equally-weight each stock in the tradable universe weekly.

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

Strategy Results & Robustness Checks

In this section, we are calculating the Sharpe Ratio (SR) of our strategy and its market outperformance, after running the backtest for two sub-periods, April 2015 to March 2020 and April 2020 to March 2025. By doing this split, we will be able to compare the performance between two different macro regimes, the low-rates environment and the post-COVID period.

In addition, we apply a t-test on the average weekly portfolio returns to check if the results are statistically different from 0, then we check the stationarity of our returns time series by using an Augmented Dickey-Fuller test and lastly, we correct for heteroskedasticity and autocorrelation of returns by applying Newey-West standard errors.

The returns series for both strategies and periods are significantly different from 0 and stationary. However, none of the excess returns relative to the market benchmark are statistically significant. Compared to the 2004-2014 period studied in the paper, this can be attributed to improved S&P 500 returns. Visual results can be seen in Figure 3 below, with very similar charts for the Main and Loser strategies.

2015-20 Performance Metrics	Main Strategy	Loser Strategy	S&P 500
Annualized Return	16.96%	14.24%	8.59%
Annualized Volatility	24.85%	25.41%	13.81%
Sharpe Ratio	0.66	0.54	0.59
Sortino Ratio	0.85	0.70	0.60
Hit rate	55.42%	54.86%	60.39%
Max Weekly Drawdown	-25.49%	-25.49%	-11.49%

Figure 1, Source: BSIC

2020-25 Performance Metrics	Main Strategy	Loser Strategy	S&P 500
Annualized Return	17.35%	23.13%	18.27%
Annualized Volatility	30.06%	30.36%	16.39%
Sharpe Ratio	0.56	0.75	1.08
Sortino Ratio	0.82	1.05	1.62
Hit rate	53.10%	53.97%	56.80%
Max Weekly Drawdown	-39.44%	-39.44%	-5.79%

Figure 2, Source: BSIC

The most relevant performance metrics are shown in the tables above. As can be noticed, the only improvement in risk-adjusted returns relative to the benchmark comes from the Main strategy in the first sub-period analyzed and it is marginal. For a risk-loving investor, the Loser strategy in the second sub-period also delivers impressive returns, but with the caveat of significant drawdowns.

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

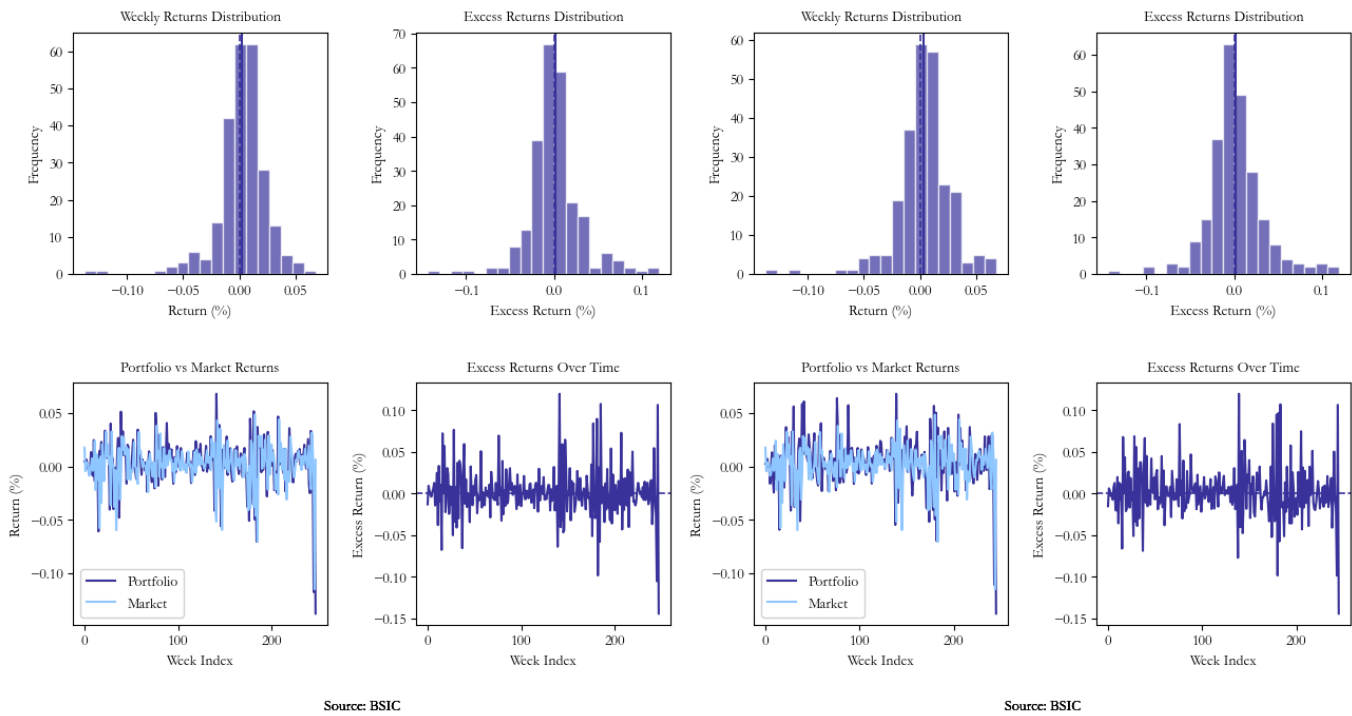


Figure 3, Left: Main Strategy, Right: Loser Strategy, Source: BSIC

Factor Analysis

The last step of our analysis consists in a factor OLS regression in order to check if the abnormal returns (alpha) are statistically significant for the two sub-periods. We have used WRDS to import the three Fama-French factors, the risk-free rate and the Carhart momentum factor. The equation used for our linear regression is:

$$r_{it} - rf_t = \alpha + \beta_i * mktrf_t + h_i * HML_t + s_i * SMB_t + u_i * UMD_t + \varepsilon_i$$

where r_{it} is portfolio's return, rf_t is the risk-free rate, α is the intercept of the regression, $mktrf_t$ is the excess market return, HML_t is the return of the "high minus low" portfolio, SMB_t is the return of the "small minus big" portfolio, UMD_t is the return of the "up minus down" portfolio, ε_i is the residual term and β_i, h_i, s_i and u_i are regression's betas.

Main 2015-20	Coefficient	P > z
Alpha (weekly)	0.26%	0.04**
Excess Market Return	10.72%	0.676
HML	17.95%	0.702
SMB	48.25%	0.145
UMD	0.78%	0.977

Loser 2015-20	Coefficient	P > z
Alpha (weekly)	0.33%	0.018**
Excess Market Return	1.48%	0.952
HML	32.88%	0.512
SMB	67.84%	0.054*
UMD	9.94%	0.741

Figure 4, Source: BSIC

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

Main 2020-25	Coefficient	P > z
Alpha (weekly)	0.44%	0.002***
Excess Market Return	-10.63%	0.455
HML	-33.65%	0.093*
SMB	-18.52%	0.342
UMD	-11.23%	0.580

Loser 2020-25	Coefficient	P > z
Alpha (weekly)	0.41%	0.007***
Excess Market Return	-16.74%	0.228
HML	-19.46%	0.312
SMB	-20.40%	0.349
UMD	-25.12%	0.161

Figure 5, Source: BSIC

As noticed above, alphas are statistically significant for both periods and both variations of the strategy, with annualized abnormal returns ranging from 13.52% to 22.88%. The exposure to the factors used is nonetheless insignificant, with marginal exceptions for the SMB of the Loser strategy between 2015 and 2020, and the HML of the Main strategy between 2020 and 2025. So far, the results are consistent with the current literature's findings.

Monthly Rebalancing

In this part, we continue our analysis by rebalancing the strategy monthly. Consequently, we will use monthly search scores for the entire period considered. The ASV is calculated similarly using the following formula:

$$ASV_t = \ln(GSV_t) - \ln(\text{Med}(GSV_{t-1}, GSV_{t-2}))$$

This helps reduce transaction costs and allows us to inspect the effect of under-pricings as indicated by the search volume data on a slightly longer time horizon than weekly. Same as before, we will differentiate between the Main and Loser strategies and calculate performance metrics and conduct factor analysis, which can be found below.

2015-25 Performance Metrics	Main Strategy	Loser Strategy	S&P 500
Annualized Return	17.72%	21.41%	13.42%
Annualized Volatility	15.80%	15.43%	15.52%
Sharpe Ratio	1.03	1.29	0.77
Sortino Ratio	1.19	1.77	1.00
Hit rate	68.75%	70.37%	68.14%
Max Monthly Drawdown	-18.03%	-10.60%	-12.51%

Figure 6, Source: BSIC

The Sharpe and Sortino ratios are clearly improved relative to the weekly rebalancing, mostly due to significant reduction in the volatility of the strategy. It can indicate some randomness present in the weekly signal, as well as the fact that short-term movements are mostly dictated by institutional investors. At the same time, retail investors usually seek medium to long-term returns when looking for potential stocks to add in their portfolios.

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

Main	Coefficient	P > z
Alpha (monthly)	1.31%	0.000***
Excess Market Return	101.37%	0.053*
HML	-53.90%	0.320
SMB	-232.96%	0.006***
UMD	68.51%	0.147

Figure 7, Source: BSIC

Loser	Coefficient	P > z
Alpha (monthly)	1.62%	0.000***
Excess Market Return	103.85%	0.06*
HML	-144.16%	0.405
SMB	67.12%	0.042**
UMD	39.72%	0.467

The alphas are once again statistically significant, as expected, providing annualized abnormal returns of 15.72% and 19.44% for the Main, respectively Loser strategy. However, for this rebalancing horizon, the strategy also has statistically significant exposure to the SMB factor, a potential indication of the size effect being present in investors' decision process.

A different Approach to Portfolio Weighting

Lastly, we tried implementing a new method of allocating weights to the individual stocks with a negative ASV signal. Instead of the original equally-weighted strategy with weekly rebalancing, we decided to allocate more the more negative each stock's score is. Each week, for each stock i with a negative ASV score, its weight is calculated by the formula:

$$w_i = \frac{|ASV_i|}{\sum_{i=1}^n |ASV_i|}$$

where the denominator represents the sum of absolute values of negative ASV scores in that week. In addition, we normalize each weight by the stock's realized volatility and rescale the normalized weights such that they add up to 1. We do this in order to limit the returns' exposure to factors other than the sentiment idea discussed in the introduction. By performing the same analyses as before, we acknowledge that this weighting scheme performs very poorly in the first sub-period, but it produces similar results to the equal-weighted one between 2020 and 2025.

2015-20 Performance Metrics	Main Strategy	Loser Strategy	S&P 500
Annualized Return	3.59%	9.01%	8.59%
Annualized Volatility	12.50%	16.57%	13.81%
Sharpe Ratio	0.25	0.51	0.59
Sortino Ratio	0.26	0.54	0.60
Hit rate	31.85%	43.67%	60.39%
Max Weekly Drawdown	-15.31%	-15.31%	-11.49%

Figure 8, Source: BSIC

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

2020-25 Performance Metrics	Main Strategy	Loser Strategy	S&P 500
Annualized Return	25.01%	18.11%	18.27%
Annualized Volatility	16.63%	18.72%	16.39%
Sharpe Ratio	1.47	0.94	1.08
Sortino Ratio	2.08	1.36	1.62
Hit rate	52.26%	50.00%	56.80%
Max Weekly Drawdown	-5.83%	-7.26%	-5.79%

Figure 9, Source: BSIC

Main 2015-20	Coefficient	P > z
Alpha (weekly)	0.07%	0.496
Excess Market Return	7.40%	0.712
HML	25.55%	0.599
SMB	22.03%	0.406
UMD	-0.05%	0.998

Figure 10, Source: BSIC

Loser 2015-20	Coefficient	P > z
Alpha (weekly)	0.15%	0.235
Excess Market Return	2.43%	0.925
HML	22.87%	0.682
SMB	67.12%	0.044*
UMD	9.62%	0.746

Main 2020-25	Coefficient	P > z
Alpha (weekly)	0.43%	0.002***
Excess Market Return	-5.33%	0.693
HML	-22.91%	0.237
SMB	-26.05%	0.169
UMD	0.75%	0.970

Figure 11, Source: BSIC

Loser 2020-25	Coefficient	P > z
Alpha (weekly)	0.34%	0.021**
Excess Market Return	-17.82%	0.142
HML	-15.95%	0.397
SMB	-38.48%	0.077*
UMD	-21.40%	0.227

In the 2015-20 period, returns are not statistically different from zero and neither are the alphas statistically significant. However, the picture changes for the second sub-period, with performance metrics significantly improved compared to the equally-weighted strategy, making the Main strategy an attractive alternative to the benchmark. The annualized abnormal returns are high enough at 22.36% and 17.68% for the Main and Loser strategies, once again statistically significant.

We tried implementing just the relative weighting method, without volatility normalization, however, the performance metrics were much worse.

To get a visual understanding of the difference between the two weighting schemes, we have attached below the cumulative performance charts of the 4 strategies versus the benchmark, for both sub-periods.

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

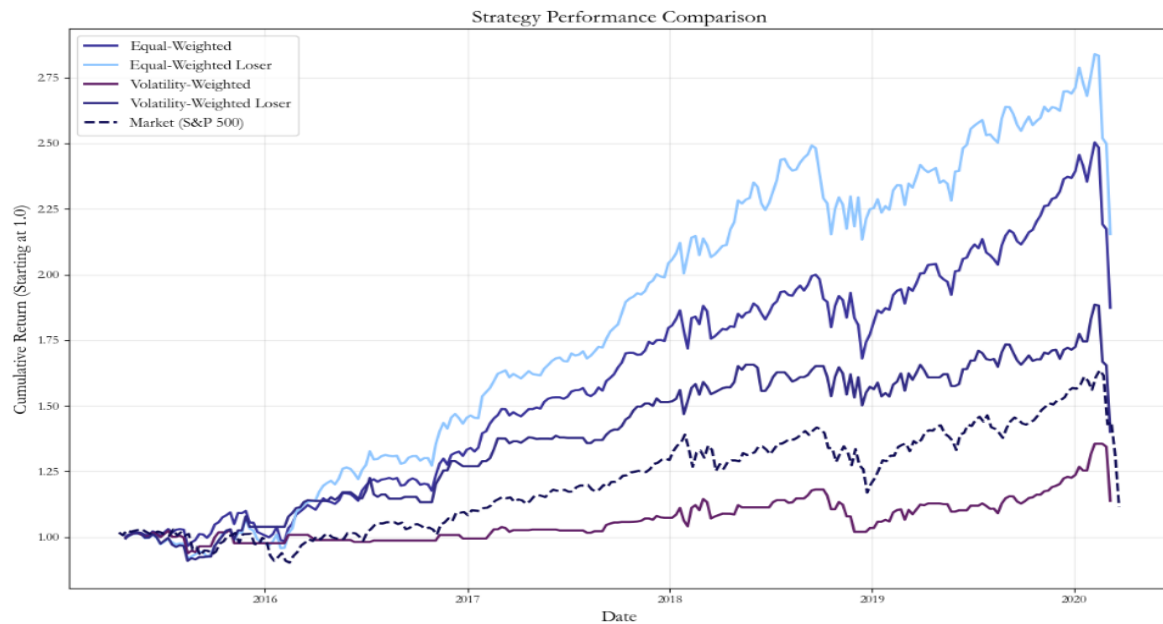


Figure 12, Period: 2015-2020, Source: BSIC

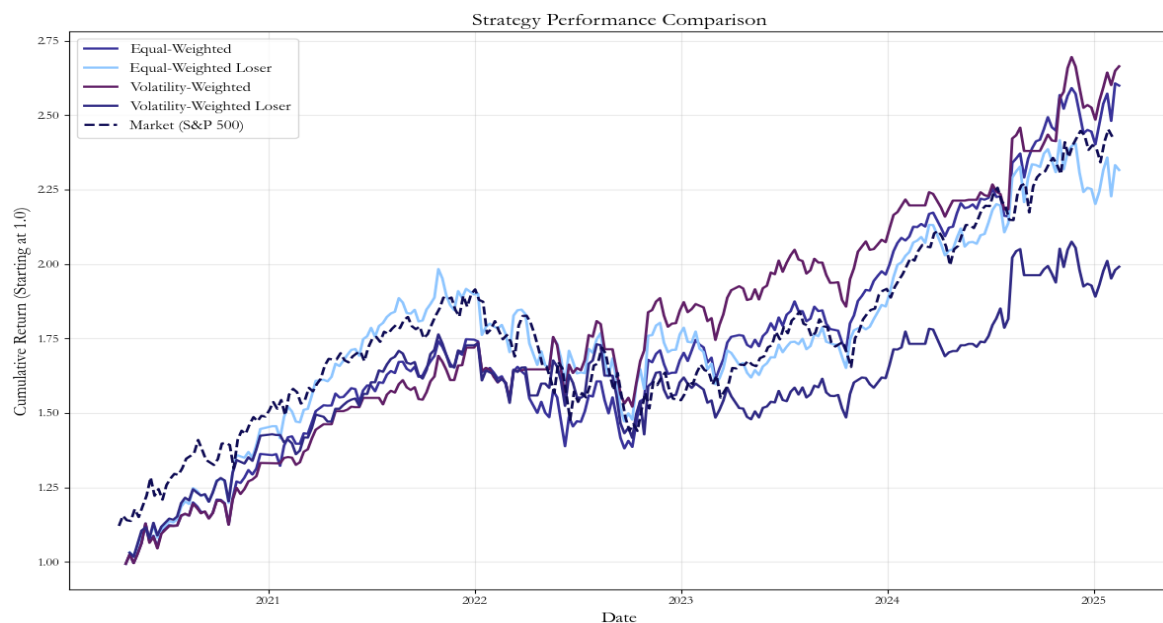


Figure 13, Period: 2020-2025, Source: BSIC

Accounting for Transaction Costs

The above results would be purposeless without a rigorous analysis of strategy's transaction costs. When buying the security, the difference between ask and mid price is paid by the investor, while the difference between bid and mid price is paid when selling it. For simplicity, we will assume the transaction costs calculated in the paper as our baseline, which should hold true given market makers have grown more competitive over the past 10 years. As our

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.

tradable universe consists of about two thirds of the stocks used in the paper, the assumed transaction costs are about 0.0702% from returns for every single rebalancing period. That would account for an annualized 3.65% TCs for the weekly-rebalancing equal-weighted strategy and 0.84% TCs for the monthly-rebalancing equal-weighted strategy, both figures being significantly below strategies' returns. The transaction costs for the relatively-weighted strategy cannot be precisely calculated, but they surely make the strategy worse off in the 2015-20 sub-period and are not great enough to counterbalance the impressive results obtained in the second half of our study.

Conclusion

In this paper, we aimed to provide the reader with a basic understanding of how a simple mechanism like Google Trends can serve as trade idea generator. We presented two versions of the signal required to enter the strategy, two rebalancing methods and two approaches to portfolio weighting. Further areas of research can include dynamic rebalancing based on the macroeconomic regime, inclusion of the VIX index as an additional signal as suggested by the paper, inclusion of more sentiment proxies or more systematic portfolio weighting methods.

References

- [1] Storms K., Kapraun J. & Rudolf M., "In Search of Alpha – trading on limited Investor Attention", 2015
- [2] Storms K., Kapraun J. & Rudolf M., "Can Retail Investor Attention enhance Market Efficiency? Insights from Search Engine Data", 2015
- [3] Fang L.H. & Peress L., "Media Coverage and the Cross-Section of Stock Returns", 2009
- [4] Barber B.M. & Odean T., "All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors", 2008
- [5] www.google.com/trends

TAGS: Google Trends, S&P 500, Fama-French Factors, Alpha

All the views expressed are opinions of Bocconi Students Investment Club members and can in no way be associated with Bocconi University. All the financial recommendations offered are for educational purposes only. Bocconi Students Investment Club declines any responsibility for eventual losses you may incur implementing all or part of the ideas contained in this website. The Bocconi Students Investment Club is not authorised to give investment advice. Information, opinions, and estimates contained in this report reflect a judgment at its original date of publication by Bocconi Students Investment Club and are subject to change without notice. The price, value of and income from any of the securities or financial instruments mentioned in this report can fall as well as rise. Bocconi Students Investment Club does not receive compensation and has no business relationship with any mentioned company.