Long-term performance of U.S. equity long/ short strategy

Based on Bloomberg news sentiment

Xin Cui, Quantitative Financial Research, Bloomberg L.P.

Lei Huang, Product Manager of Event-Driven Feeds, Bloomberg L.P.

Bloomberg

Abstract

We backtest the previously reported sentiment-driven equity strategy against 10 years of extended history.[1] By using cross-sectional ranking of aggregated sentiment scores, a diversified equity long/short portfolio can be constructed that exhibits low market exposure, low volatility and high Sharpe ratio. The performance is robust across both bull and bear markets. Portfolios based on small-cap stocks consistently outperform those based on large-cap ones. Elongating the aggregation window leads to a lower return but with a slower signal decay.

Contents

Introduction

While news sentiment has long been considered to be a strong driving factor of stock price movement, how to systematically extract sentiment-embedded alpha remains an open, unsolved problem for both academic researchers and financial practitioners. The challenges are many. For example, the aggregated news source from which sentiment is derived must be comprehensive enough to sufficiently represent the overall news flow in the marketplace. In addition, each news record needs to be point-in-time with the correct original time stamp — from which the story-level sentiment information can be reliably transformed into a stock-level time series. Last but not least, the text-to-sentiment derivation method, which is often implemented by NLP (natural language processing) techniques, must be specially built so that the output sentiment is indicative of any shift in investor confidence rather than just a general change of mood.

Bloomberg's real-time news sentiment product has been carefully designed to address most of the challenges outlined above. A previous paper proposed and tested three different portfolio construction techniques based on this sentiment dataset[1]. The study finds that the sentiment strategies outperformed the benchmark ETFs from January 2015 to August 2016 (20 months). However, the entire sample has been collected from the phenominal bull run in the U.S. stock market; this bull market saw the S&P 500 Index rally a whopping 267% from its March 2009 low. What remains unclear is whether the reported behavior can continue for a longer time. Investigating the performance during the subprime crisis from 2007 to 2009 is especially interesting; it was during this period that global stock markets were dominated by panic and volatility.

This paper reports on the extended study of the 10-year span from February 2007 to December 2016.
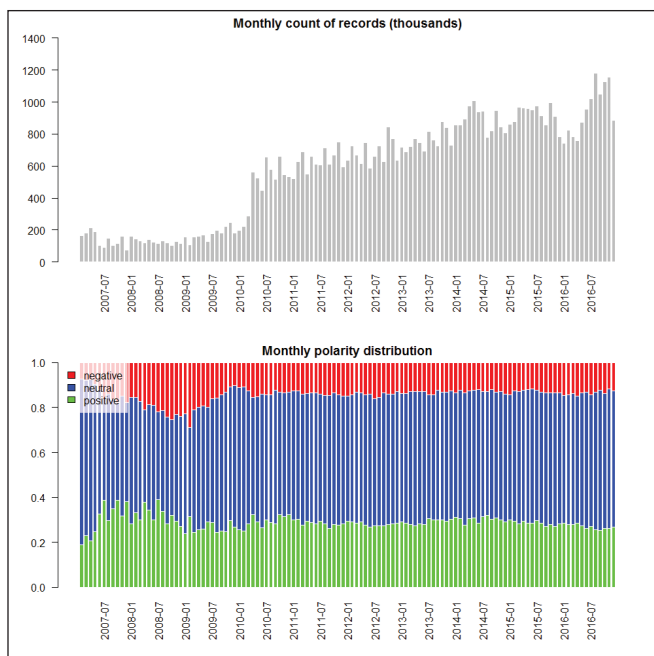
Bloomberg story-level news sentiment

Bloomberg story-level news sentiment tracks the per-story sentiment impact on any underlying company. It classifies news articles according to the answer to the question: "If an investor holding a long position in this security were to read this article, would his/her confidence in the holding increase, decrease or remain the same?"

For each news article, the classification is associated with a particular company and consists of two parts: score and confidence. The score is a directional indicator out of the <-1, 0, +1> triplet, which corresponds to negative, neutral or positive sentiment. The confidence is a numeric value between [0, 100], indicating the confidence of the model estimates. If multiple stocks are mentioned in the same story, each stock will be computed independently — it will get its own sentiment score and confidence values.

Story-level news sentiment from Feburary 7, 2007, to December 31, 2016, was used for this study. Bloomberg released the latest version (version 2) of the sentiment analysis engine on June 20, 2015; historical data before that date was backfilled by applying the same model to archived textual news.

The chart below shows the monthly volume and the polarity distribution of the datasets. Sentiment volume increased dramatically after the first quarter of 2010 because new sources were added. The polarity distribution is asymmetric: more stories classified positive than negative throughout the sample period.

Sentiment aggregation

The story-level news sentiment can be aggregated for the same stock to form company-level sentiment flow. For a given stock at time t, a daily average sentiment $\overline{S}_t$ is calculated using a 24-hour window:

$$\overline{S}_t = \frac{\sum_{i=1}^{N} S^i \cdot C^i}{N}$$

where N represents the number of stories that mentioned the target stock; S^i is the sentiment label with value 1, 0, or -1; and C^i is the corresponding confidence for a given story i.

Portfolio construction

For this study, the sentiment aggregation was performed every day before market open for a universe of stocks. From this daily snapshot, stocks could be sorted into their cross-sectional percentile rankings. Intuitively, we expect stocks to outperform on (relatively) bullish sentiment and underperform on (relatively) bearish sentiment. This hypothesis can be explored by a momentem-style long/short portfolio according to sentiment rankings. In the following section, portfolios will be constructed using different parameter selections to study their effects.

Discussion

S&P 500 Universe vs. Russell 2000 Universe

First, we study strategies based on the daily average sentiment $\overline{S}$. The portfolios are constructed with these variations:

- **High-Minus-Low 5% portfolio (HML 5%)** – long (short) the top (bottom) 5% of stocks ranked by sentiment scores. Stocks in long and short portfolios are equally weighted.
- **High-Minus-Low 1/3 portfolio (HML 1/3)** – long (short) the top (bottom) 1/3 of stocks ranked by sentiment scores. Stocks in long and short portfolios are equally weighted.
- **Proportional portfolio (Proportional)** – long (short) stocks with positions proportional to the difference of the sentiment score from its cross-sectional mean. If the sentiment score is above the mean, take a long position; if it is below, take a short position. The further away from the mean, the greater should be the position.

Since the news coverage in the early years (2007-2009) is relatively low, a possible outcome is having only a few stocks in our daily holdings, which presents a challenge for the diversification of the idiosyncratic risk. Therefore, in our back-testing, we impose the requirement that both the long and short legs have, at minimum, 10 stocks; if either leg has fewer, we do not form the portfolio on that day. It turns out that this condition is only triggered for a few days in the early years (2007-2009).

All positions are entered at market open and liquidated at market close. Daily portfolio return is a weighted sum of those of individual stocks (detailed formula provided in the previous paper[1]). Equity index ETFs are used as benchmarks, with the overnight close-to-open gaps stripped away for consistent comparison of the cumulative performance.

We back-tested the strategy on S&P 500 and Russell 2000 stocks. Index membership is point-in-time. The results are reported below. We find that over the long term, sentiment-driven portfolios significantly beat their respective benchmarks on a risk-adjusted basis. The performance is robust in both bull and bear markets. In fact, portfolios based on an S&P 500 universe performed notably better from 2007 to 2009 than in the subsequent years. All portfolios demonstrate low negative market exposure (beta) and low volatility, with an annualized Sharpe ratio in the range of 1 to 2 for S&P 500 stocks and high 3 for Russell 2000 stocks (without transaction cost). More diversified portfolios turned in a better performance.

Illustration for S&P 500 stocks
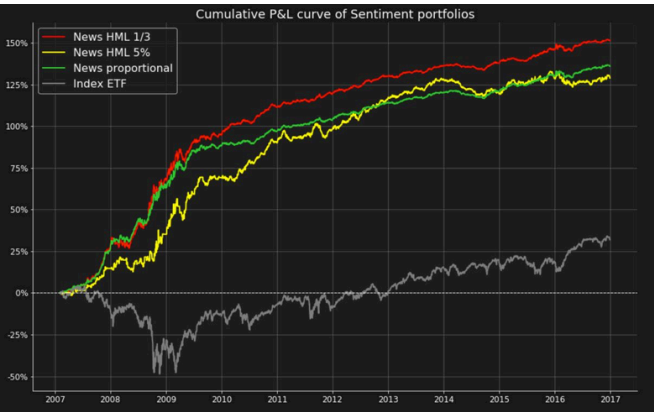
	Beta	Ann. Ret.	Ann. Vol.	Sharpe	Avg # Short	Avg # Long
HML 5%	-0.06	13.0%	10.6%	1.23	16.0	16.0
HML 1/3	-0.19	15.3%	8.5%	1.80	68.0	104.0
Proportional	-0.14	13.7%	7.2%	1.89	172.0	139.0
Index ETF	1.00	3.2%	16.4%	0.20		

Illustration for Russell 2000 stocks
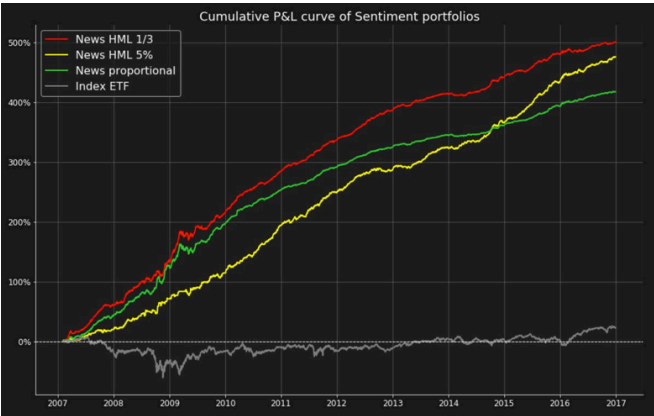
	Beta	Ann. Ret.	Ann. Vol.	Sharpe	Avg # Short	Avg # Long
HML 5%	0.00	48.0%	12.8%	3.76	25	25
HML 1/3	-0.08	50.6%	13.0%	3.89	69	165
Proportional	-0.07	42.2%	10.8%	3.92	298	204
Index ETF	1.00	2.2%	22.2%	0.10		

Decomposition of long/short legs

The performance during the subprime mortgage crisis is especially interesting. In such a highly volatile environment, stock correlation rose — with price actions driven by prevailing pessimism rather than by company fundamentals. That crisis was a very challenging period for fundamentals-based investment strategies.

To understand the driver of the return, we break the portfolio down into two legs — one with the long-only stocks and one with the short-only stocks — and analyze their individual performance. High-Minus-Low 1/3 (HML 1/3) portfolios are used as the example. For S&P 500 stocks, we find that the return from 2007 to 2009 came predominantly from the short leg, which delivered negative returns in later years of the bull market. Similar behavior was observed for Russell 2000 stocks, but here the short leg delivered positive returns during the entire test period.

Another observation: the performance of an individual leg tends to have a much higher volatility/drawdown. However, when long and short legs are combined, most of this choppiness cancels out and the overall portfolio enjoys a much smoother return. This is one of the hallmarks of market-neutral strategies.
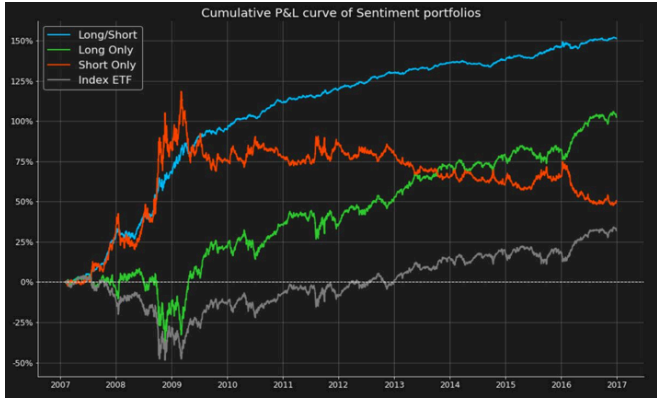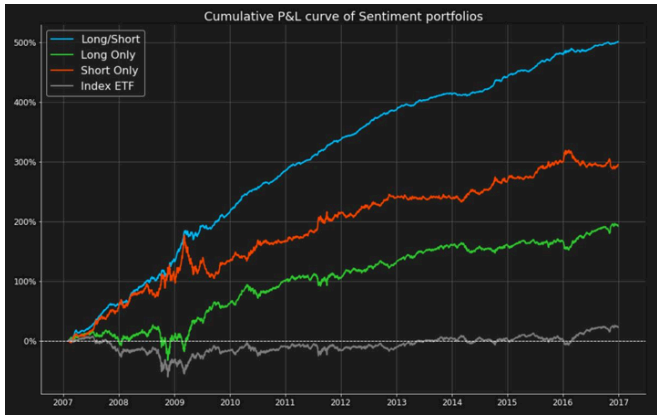
Illustration for S&P 500 stocks

Illustration for Russell 2000 stocks

Signal strength & alpha decay

In this section, we discuss two aspects of the sentiment signal:

- Strength of signals derived from various moving-average smoothing windows
- Duration that the signal can last

To study the signal strength, we formed smoothed sentiment based on S_t with different half-lives. The shorter the half-life, the more overweighted recent daily scores are against older ones. Specifically, we chose half-lives of 1 day, 3 days, 7 days and 14 days to study the effect.

The stardard time-weighted exponential moving-average method is used to calculate the smoothed sentiment M_t:

$$M_t = 2^{-\frac{\Delta}{\tau}} M_{t-1} + (1 - 2^{-\frac{\Delta}{\tau}}) \cdot \overline{S}_t$$

$$M_0 = \overline{S}_0$$

where τ is the decay half-life, and Δ is the time elapsed between the new observation and the immediately preceding one. Time difference is measured in number of days.

To study the alpha decay, we used the same portfolios formed on the same day as we observed the signal, either $\overline{S}_t$ or M_t, but delayed the order execution with a different number of days. Positions are still entered at open and exited at close within the same day for consistency. Specifically, we chose to delay 1 day, 2 days, 5 days and 10 days to study the effect.

High-Minus-Low 1/3 (HML 1/3) portfolios are used in the example.

The graphs below compare the risk-adjusted performance of all the combinations of smoothing/delay parameters. Each colored curve represents portfolios of the same signal smoothing (half-life), but different execution delays. We tested both the S&P 500 and the Russell 2000 stock universe.

Our study shows that smoothing with a longer window tends to reduce the signal strength. On a relative basis, the reduction is more prominant for large caps than for small caps, suggesting that the information filtration process is more transient in the former. All the curves show downward sloping — steeper for shorter delays and flattened out thereafter. This suggests that one should take immediate action once signal is observed, while the holding period could be slightly longer than one day when transaction cost is taken into account.

These observations are consistent with the general understanding that large-cap stocks price in new information faster than small-cap stocks because of market efficiency and liquidity advantage. That said, from a practical portfolio management standpoint, signal smoothing is still beneficial even for large-cap stocks as it can stabilize the daily fluctuations — directly leading to a lower turnover ratio and the associated transaction cost.
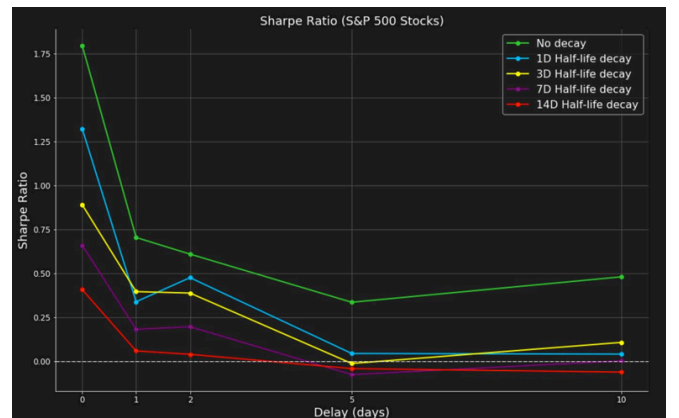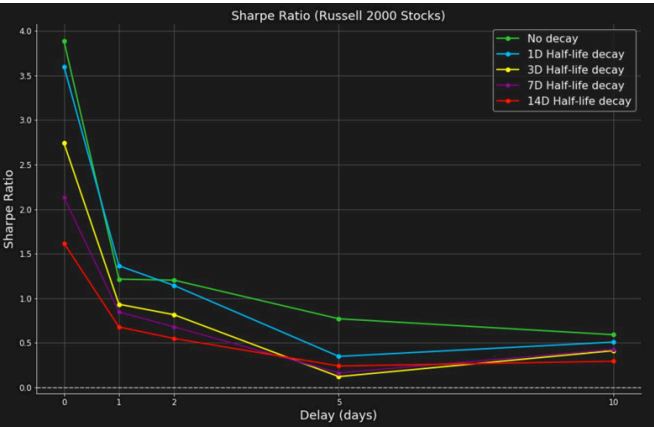
Illustration for S&P 500 stocks

Long-term performance of U.S. equity long/short strategy
based on bloomberg news sentiment

Illustration for Russell 2000 stocks

Sharpe Ratio (Russell 2000 Stocks)

Final remarks

Nonconventional datasets, especially those with strong economic or causal linkage to the stock fundamentals, are finding their way into more and more systematic trading models. While the purpose of this study is not to provide a ready-to-go trading strategy that can be immediately implemented to generate superior investment performance, the findings do provide strong support for the long-term persistence of sentiment alpha. The basic portfolio construction method also lays down a foundation that can be further optimized to address practical issues unique to individual portfolios.

Reference

[1]Xin Cui, Daniel Lam, and Arun Verma, Embedded Value in Bloomberg News & Social Sentiment Data, 2016.

Learn more

Learn more about how we can help your firm react more strategically using sentiment. Visit bloomberg.com/professional or reach us at eprise@bloomberg.net.

Take the next step.

For additional information, press the <HELP> key twice on the Bloomberg Terminal®.

bloomberg.com/professional

Beijing +86 10 6649 7500	**Hong Kong** +852 2977 6000	**New York** +1 212 318 2000	**Singapore** +65 6212 1000
Dubai +971 4 364 1000	**London** +44 20 7330 7500	**San Francisco** +1 415 912 2960	**Sydney** +61 2 9777 8600
Frankfurt +49 69 9204 1210	**Mumbai** +91 22 6120 3600	**São Paulo** +55 11 2395 9000	**Tokyo** +81 3 3201 8900