A News Sentiment Stock Screener for Discretionary Traders

May 2019

TABLE OF CONTENTS

The RavenPack Data Science Team:

Peter Hafez
Chief Data Scientist

Ricard Matas
Data Scientist

Francesco Lautizi
Data Scientist

Francisco Gomez
Data Scientist

Jose A. Guerrero-Colón
Data Scientist

www.ravenpack.com
New York – 300 Park Avenue, 12th Floor, New York, NY 10022 | TEL: +1 (646) 277-7339 | research@ravenpack.com

Executive Summary

In this paper, we introduce a news *sentiment stock screener* to construct small sized equity portfolios using RavenPack Analytics. The screener can help discretionary investors narrow down candidates from a larger investable universe. We select companies with a high number of news events and high sentiment polarization, a task that can be achieved using our SESI [1], or Sum Excess Sentiment Indicator. The main conclusions are:

- *News Sentiment* significantly outperforms *price momentum* as a stock screener, for which probability distributions clearly demonstrate a positive shift. This is more than a 4x increase in average performance during the 8+ year backtesting period.
- By creating targeted portfolios, we achieve high annualized returns (up to 80%) with attractive Information Ratios (IRs) of up to 4.0, but also high per-trade-returns (up to 33bps).
- When including a *discretionary overlay*, dropping 2 out of 22 daily trades, annualized returns and Information Ratios are boosted by up to 200%, depending on our skill at removing bad performers.

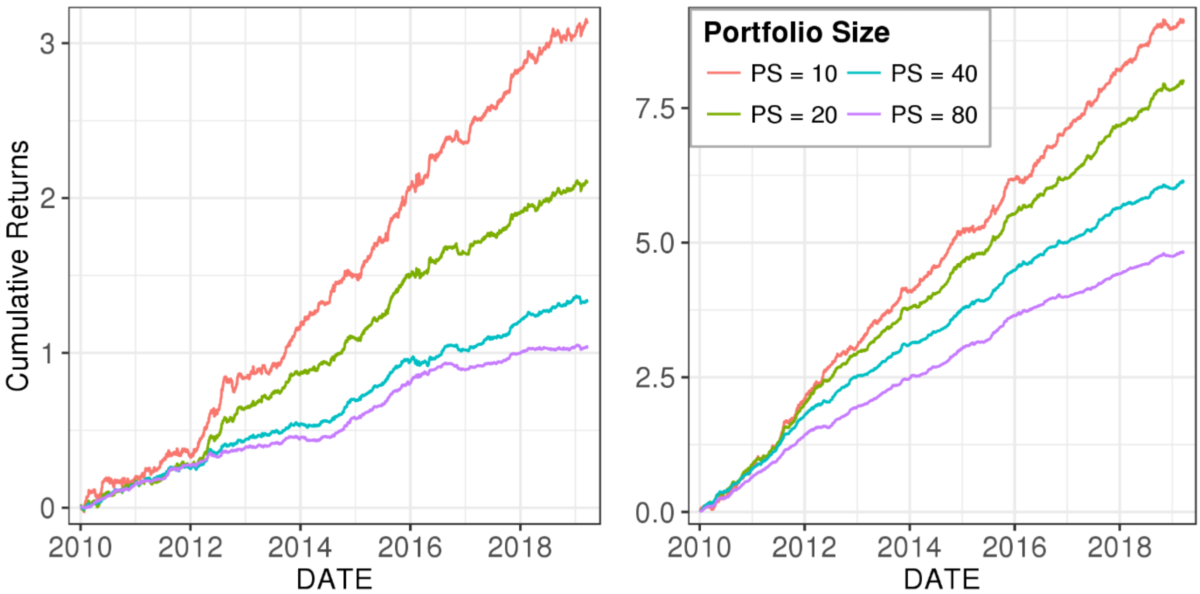

Figure 1: Cumulative Returns for portfolios restricted to extreme sentiment signals (US). Cumulative Returns of equal-weighted long/short extreme sentiment portfolios of different size (PS) and for several holding periods (HP). Results for Large-cap (left) and Small-cap (right) companies.

Source: RavenPack, May 2019

About RavenPack Data

RavenPack Analytics provides real-time structured sentiment, relevance and novelty data for entities and events detected in unstructured text published by reputable sources. Publishers include Dow Jones Newswires, Barron's, the Wall Street Journal, and over 19,000 other traditional and social media sites. RavenPack detects key news events and produces analytics on over 45,000 listed stocks, top products and services, all major currencies and commodities, financially relevant places and organizations, and key business and political figures. The dataset includes 18+ years of structured analytics data available for backtesting.

1. Introduction

Alternative data is the "new black" of the investment industry. Today, rarely anyone would question the potential of using new and unique data sources to seek alpha opportunities. However, new opportunities often come at a cost.

To leverage alternative data, quantitative trading firms are progressively driven to boost their investments in research and IT infrastructure. Once obtained, this data is evaluated with the objective of converting it into alpha strategies. Most of the alpha strategies using alternative data work in the framework of weak predictor aggregation. That is, when combining individual predictions, the individual accuracy does not need to be very high in order to obtain a stable and strong-performing strategy.

Two key points needed for these quantitative strategies to work include: (1) signal breadth and (2) signal frequency. The former implies trading large universes, involving hundreds or even thousands of equities; the higher the number is, the better the strategy works. The latter defines the frequency at which the strategy can be rebalanced using new insightful information; the more times this occurs, the better it works. As a result, investors will typically desire to consume more and more data, and expand into new markets and trading horizons.

RavenPack Analytics provides a source of alpha that directly falls into the above described scenario: a fast feed of information, involving thousands of entities, which has been shown to predict market movements [2-6]. Accordingly, the resulting strategies, with no additional considerations, may still require more effort, both in research and in trading: in research, because they may be coupled with other alternative data sources or traditional market factors, optimized under the desired portfolio constraints, etc.; in trading, because they are generally destined to be part of a large portfolio framework that combines many other alpha signals.

Instead of trying to generate as many signals as possible, as you would expect in a quantitative strategy, in this study we aim to reduce the signal volume and build sentiment strategies that can be more attractive to discretionary traders. Often, traditional investors don´t have the necessary resources to test and manage large universes or they prefer to focus on fewer equities, where they can apply more traditional human-based trading methodologies. Our objective is to offer an easier way for discretionary traders to incorporate alternative data into their portfolios, pointing them in the direction of new alpha sources, using a customized filtering layer that reduces the heavy resources involved in working with alternative data.

To this end, we propose a way to select trading signals that can be used to build portfolios or equity baskets of a desired size. This selection is based on choosing stocks with a high number of news events and high polarization in sentiment, information which, as described in previous research [1], is contained in the magnitude of the Sum Excess Sentiment Indicator (SESI). We already know that SESI's magnitude (or absolute value), the *Sentiment Strength*, can be used to select stronger signals, e.g. we showed in [1] how higher sentiment quintiles contain the most valuable information when predicting future stock returns. Quintile selection could be seen as a basic example of the universe reduction procedure that will be followed in the present research.

This paper is structured as follows: In Section 2, we provide a brief description of the data and our portfolio construction methodology. In Section 3, we compare random portfolios of fixed size using different selection criteria. In particular, we compare a strategy based on event sentiment with a fully random selection, as well as a basic price momentum strategy. Section 4 highlights the performance results for a set of *Extreme Sentiment Portfolios* (ESP), which only chooses entities with extreme SESI. To assess the statistical significance, the ESPs are compared to our random portfolio benchmark [7].

Typically, discretionary investors restrict their stock selection to a narrow universe, conditioning on properties such as liquidity and/or sector information. In Sector 5, we split our initial universe by market capitalization and sector, which will serve as a way to disclose the strength of the different contributions. Section 6 focuses on the generation of stock screens and studies the performance distributions of smaller portfolios. The resulting distributions provide an idea of the potential benefits of using our *Sentiment Stock Screener* to select a targeted universe from where to build your own portfolios. To better mimic a discretionary process, in Section 7, we include an additional overlay by introducing a filter that tries to approximate the ability of a trader to identify poor performers within the suggested portfolio. Finally, in Section 8, we provide our conclusions.

2. Data Description

To construct our SESI Indicator, we use RavenPack's news analytics suite including the Event Sentiment Score (ESS), the Event Relevance Score (ERS), and the Event Similarity Days (ESD). We use data covering the period from January 2010[1] through March 2019 and include all US companies belonging to the Russell 3000 index in a point-in-time sensitive fashion. When backtesting the performance of our strategy, we use excess returns.[2]

[1] To be conservative, we omit the financial crisis (during 2008-2009) from our analysis. Hence, we only show performance statistics from January 2010. Generally, the RavenPack data performed particularly well during the financial crisis, and thus including this period tends to inflate overall performance and bias our conclusions.

[2] We measure excess return relative to an equal weight portfolio across each universe under study.

The SESI indicator has been described in more detail in previous research [1][3]. Although we build daily strategies based on a daily signal, we consider only news published during the 4 hours prior to market open.[4] Consistent with this, we build a portfolio that trades from open to open, using Volume Weighted Average Prices (VWAP) averaged for 30 minutes immediately after market open. Details on how the VWAP prices are measured can be found in [8].

2.1 Random Portfolios

We define a random portfolio [7] of size 2N as an equal-weighted portfolio where the allocated assets are randomly selected. For simplicity, we will only work with balanced portfolios, with N short and N long positions when available.[5] The protocol followed to perform this random selection will characterize the type of random portfolio that we construct. In particular, we define 3 different types:

- *Fully Random Portfolio (Random):* The random selection is done from the entire universe with the allocation sign also being assigned randomly (keeping N of each sign), independent of any presently available information.
- *Random Price Momentum Portfolio (Price):* The random selection is done from the entire universe, but the allocation sign is chosen based on price momentum. For simplicity, we only use the sign of 1 day lagged returns, although the generalization to a more complex momentum prediction is straightforward.
- *Random Sentiment Portfolio (Sentiment):* The random selection is done from the sub-universe of assets with a non-zero event sentiment indicator. Each allocation sign follows the sign of the sentiment indicator itself. The assets to be allocated are drawn with a probability distribution proportional to the Sentiment Strength, i.e. the larger the magnitude of the Sentiment Indicator for a particular asset, the larger the probability that this asset will get selected.[6]

[3] For simplicity, we have changed the weighting function for the daily sentiment bias so that effectively we are performing a simple average of the all the daily Event Sentiment Scores after the usual filters have been applied (Event Relevance ≥ 90, Event Similarity Days ≥ 90 and Event Sentiment Score != 0).
[4] Generally, outside market hour news is most informative for return predictions. We refer to [8] for a more detailed description of why the 4 hour aggregation window is the preferred choice.
[5] This condition could be relaxed and one can easily turn into long-only or short-only portfolios.
[6] A Random Portfolio that only follows the sentiment indicator sign was also used, showing always lower performance. Hence, it will not be included in the paper.

2.2. Extreme Sentiment Portfolio

The Extreme Sentiment Portfolio (ESP) is formed by selecting the 2N assets with the highest Sentiment Strength [1] each trading day, N of them corresponding to the most extreme positive sentiment indicator values (long positions) and N to the most extreme negative sentiment (short positions). Obviously, the sentiment signal will not have 100% prediction accuracy. Hence, this portfolio is not necessarily the best performer of all possible N-sized equal-weighted portfolios. It is, nevertheless, the best available candidate based solely on the Indicator. Note that selecting entities based on the ranking of sentiment indicators can be easily done using the RavenPack's Self-Service Platform (see Appendix for more details).

While random portfolios have many possible realizations, the ESP is unique[7]. On a given day, if only a number M smaller-than-N equities have a SESI of a given sign, the corresponding number of equities selected will be M. Hence, the portfolio size will inevitably be smaller than 2N on that day. Regardless of this, Gross Exposure is always maintained at its maximum.

3. Random Portfolio Comparison

In this section, we compare the behavior of the different random portfolios defined in Section 2.1. For this, we generate multiple[8] realizations of each portfolio and measure the subsequent performance distributions. In particular, Figure 2 depicts the distributions of the annualized return and the Information Ratio for a 1 day holding period (HP) and portfolio size (PS) of 20 assets. It shows how the news sentiment strategy outperforms the rest, showing an important shift towards higher values across all measures. As a consequence, if we were to select entities without using any information other than price or daily event sentiment, we would largely benefit from using sentiment for betting on next day's return. Not only would we outperform a fully random selection, but also a simple price momentum-based choice. These results are consistent for all portfolio sizes (PS from 10-80) and holding periods (HP from 1-21).[9]

Generally, when increasing PS, we achieve higher Information Ratios due to increased diversification. Furthermore, we also observe a sharpening of the distribution, which mainly can be explained by a reduction in size of the sampling population. As we expand into longer HP's, we observe a continuous

[7] This is true except when there is more than one asset with the same sentiment signal in the boundaries of our selection. In this case, we will arbitrarily choose one of those and forget about the others – it is not a frequent situation, so the effect of this choice can be disregarded.

[8] We generate 1000+ portfolios. The resulting distributions are usually characterized by small but slow-decaying tails that will not be properly resolved. Nevertheless, the corresponding probabilities are very small and the central region can be safely approximated by a Gaussian distribution, used to reflect the statistical significance of a particular value.

[9] Results available upon request.

decrease in the distribution range, something which indicates that news sentiment provides stronger value for short-term than for long-term portfolios.[10]

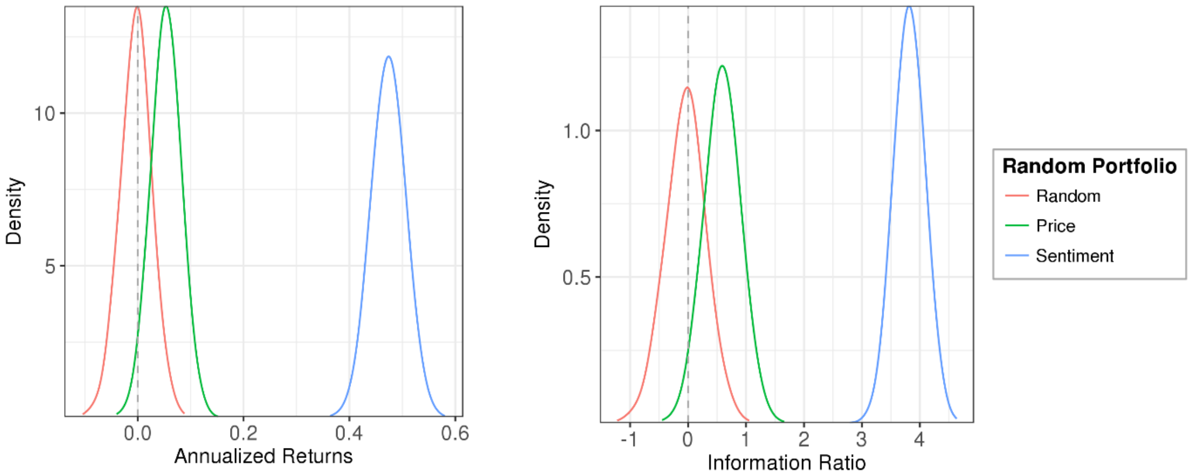

Figure 2: Performance distributions for Random Portfolios, US-R3000. Distributions of annualized returns (left) and Information Ratio (right) for different Random Portfolios, measured for HP=1 and PS=20.

Source: RavenPack, May 2019

4. Performance of Extreme Sentiment Portfolios

As shown in previous research, news sentiment is a source of alpha for equity trading [2-6]. Moreover, Sentiment Strength has proven to be a beneficial measure for selecting stronger signals among detected news events. In particular, we found that higher quintiles in Sentiment Strength outperform lower quintiles [1]. Encouraged by this result, we want to build dynamic portfolios that only select a small number of equities that rank high in Sentiment Strength. One way to proceed is to construct the so-called Extreme Sentiment Portfolio (ESP), defined in Section 2.2. These portfolios are interesting for discretionary investors as they can offer portfolio sizes that are manageable when taking a manual rather than an automated approach.

An ESP is formed by a particular selection of equities, in this case, with a non-zero sentiment signal. It corresponds directly to one of the many possible selections that constitute our same size random portfolios. Hence, the performance of the ESP can be assessed by comparing it to the random portfolio distributions[11] introduced in the previous section.

[10] To prevent path dependence, we allow for daily rebalancing even for long HP portfolios. Effectively, this means that the portfolio size increases with HP and that the empirical PS can be significantly larger than the fixed target PS (except for HP=1).

[11] Since the Random Sentiment Portfolio has been shown to outperform other random strategies, we will use it to build our benchmark distribution when we evaluate the ESP.

In Figure 3, we compare the performance of the ESPs with the distribution benchmark based on the Random Sentiment Portfolios. The ESP is represented by a vertical straight line highlighting its unique value. We find that the ESP results are highly statistically significant with strong consistency across all holding periods (from 1 to 21 days). While in this example, we have focused on a PS of 20 assets, similar behavior is found for other portfolio sizes. However, as we increase PS, we observe a clear monotonous shift towards smaller annualized returns and larger Information Ratios.[12] This pattern becomes particularly clear in Figure 4, where we compare the cumulative returns across all HP and PS. From the figure, it can also be observed that the performance is fairly consistent across the entire backtesting period.

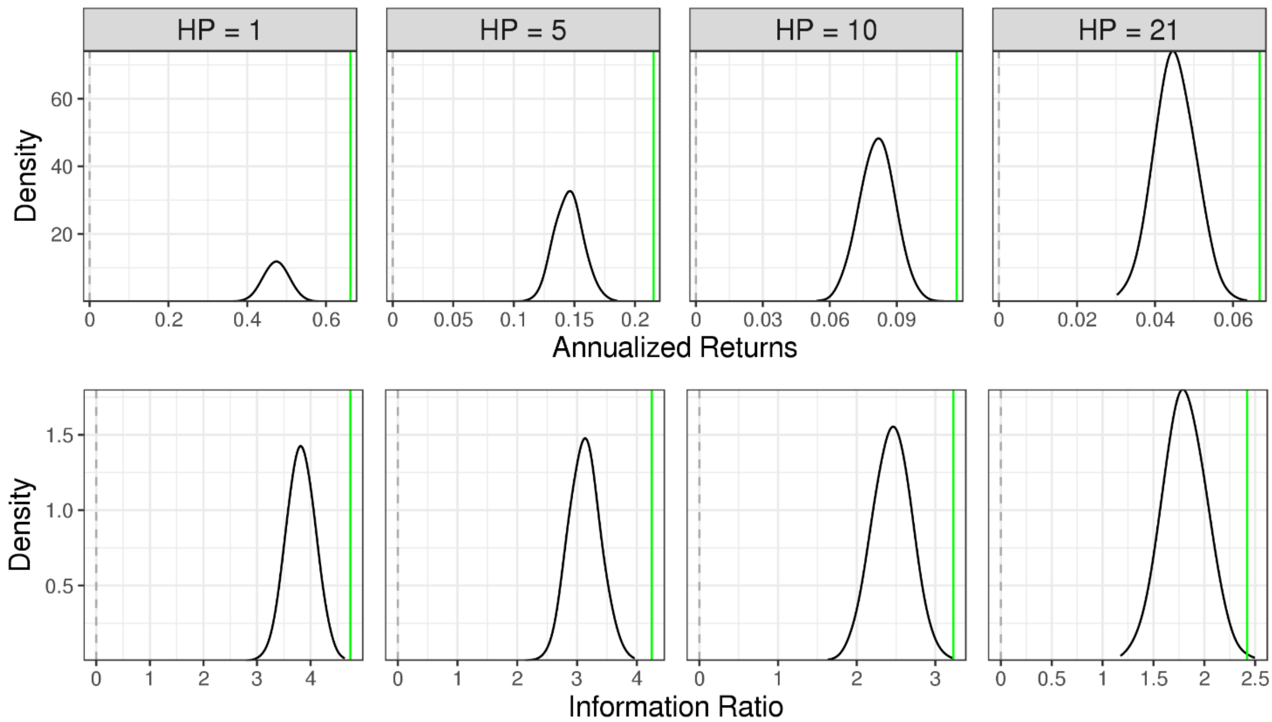

Figure 3: Performance of the ESP and Random Portfolios (US-R3000). Annualized returns (top) and Information Ratio (bottom) of the Extreme Sentiment Portfolio (vertical green line) and corresponding distributions of the Random Sentiment Portfolio for several holding periods (HP) and PS=20.

Source: RavenPack, May 2019

[12] Results for PS=10, 40 and 80 are available upon request.

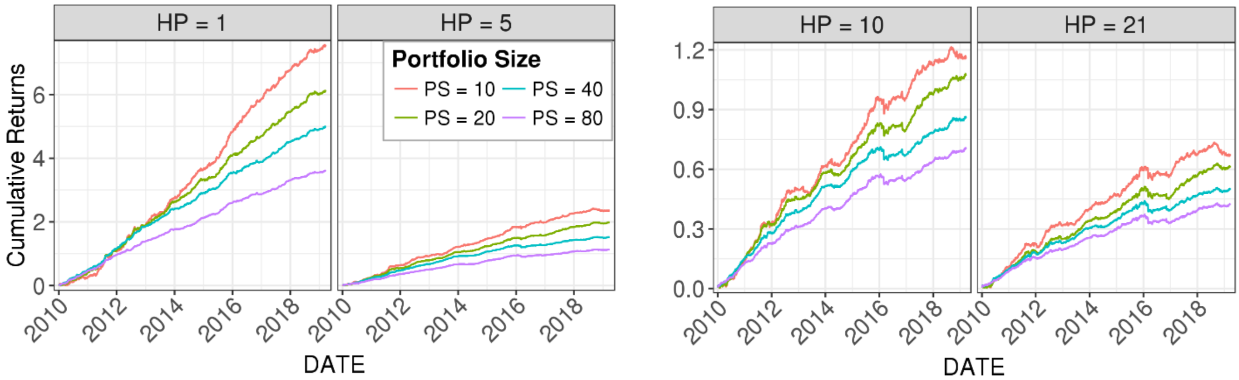

Figure 4: Cumulative Returns of the ESP and RSP (US-R3000). Cumulative Returns of the Extreme Sentiment Portfolio for several holding periods (HP) and portfolio sizes (PS).

Source: RavenPack, May 2019

As a direct consequence, by creating more targeted portfolios, it is possible to create portfolios that are particularly appealing when taking transaction costs into account. Figure 5 plots the 1 day average Per Trade Returns as a function of portfolio size. Without trading costs, we are able to achieve values between 16 - 33bps.

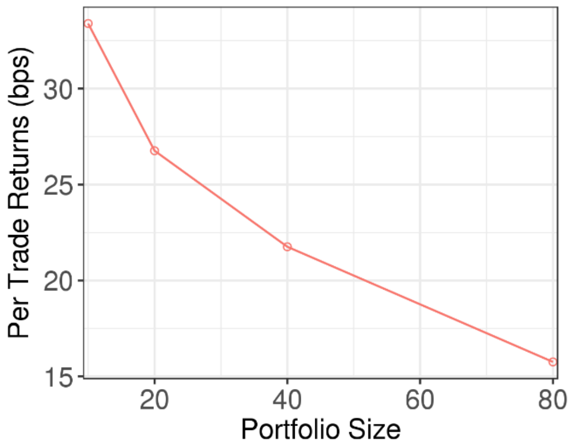

Figure 5: Per Trade Returns of the ESP (US-R3000). Per Trade Returns of the Extreme Sentiment Portfolio for several portfolio sizes and for 1 day holding period.

Source: RavenPack, May 2019

Overall, we find that news sentiment can be a powerful tool for stock selection with strong outperformance of several benchmarks constructed using random, price-based selection. Note that so far we have based our study on the entire Russell 3000 universe. In the following section, we will restrict the stock selection to narrower universes, which will serve as a way to test the robustness of the results, as well as to disclose the possible differences between the subsequent contributions.

5. Market Capitalization and Sectors

Typically, discretionary traders create targeted portfolios based on strict mandates that make them easier to manage (e.g. by focusing on particular market capitalization or sector exposure). In this section, we will first split performance by sector to get a better sense for the signal robustness for these targeted mandates[13]. Afterwards, we will split the initial universe in two market cap regions: Large/Mid-cap and Small-cap, and investigate the differences between these.

Table 1 summarizes the performance statistics of the Extreme Sentiment Portfolios when restricting the stock selection to a particular sector. Notably, all sectors have positive performance, both for short and long holding periods (1-21 days)[14]. Nevertheless, in some sectors the available volume of news is too small to consistently select 10 daily stocks. This can lead to average portfolio sizes that are smaller than the target size (PS). Interestingly, the two worst performing sectors suffer from this, with only 6 daily positions on average for a 1 day holding period; hence, selecting stocks based on extreme sentiment is less effective.

	HP=1		HP=21	
SECTOR	**AR**	**IR**	**AR**	**IR**
All Sectors	81.80%	4.02	7.30%	2.02
Financial Services	45.20%	3.76	6.60%	2.87
Healthcare	45.10%	2.06	6.00%	1.37
Durables	40.80%	2.65	3.40%	1.07
Materials&Proc.	37%	2.26	0.50%	0.12
Energy	30.90%	1.55	5.50%	1.03
Cons. Discretionary	30.70%	1.80	4.20%	1.22
Technology	26.30%	1.62	1.80%	0.51
Cons. Staples	21.40%	1.27	5.00%	1.27
Utilities	13.50%	0.88	4.10%	1.12

Table 1: Portfolio properties separating by sector (US-R3000). Annualized returns (AR), Information Ratio (IR) and portfolio size (PS) of Extreme Sentiment Portfolios when separating the universe by sectors. Results for a desired portfolio sizes of 10 stocks and two different holding periods (HP).

Source: RavenPack, May 2019

[13] In order to evaluate our strategies we always use excess returns in the particular universe of study.
[14] Note that, as explained previously, the long holding periods involve larger portfolio sizes even though we are only selecting 10 new possible trades daily.

In Section 4, we evaluated the performance of the Extreme Sentiment Portfolios (ESP) using the Russell 3000, i.e. including all market caps. In this section, we repeat our analysis, but after splitting the universe into: (1) Large/Mid-cap (Russell 1000) and (2) Small-cap (Russell 2000).

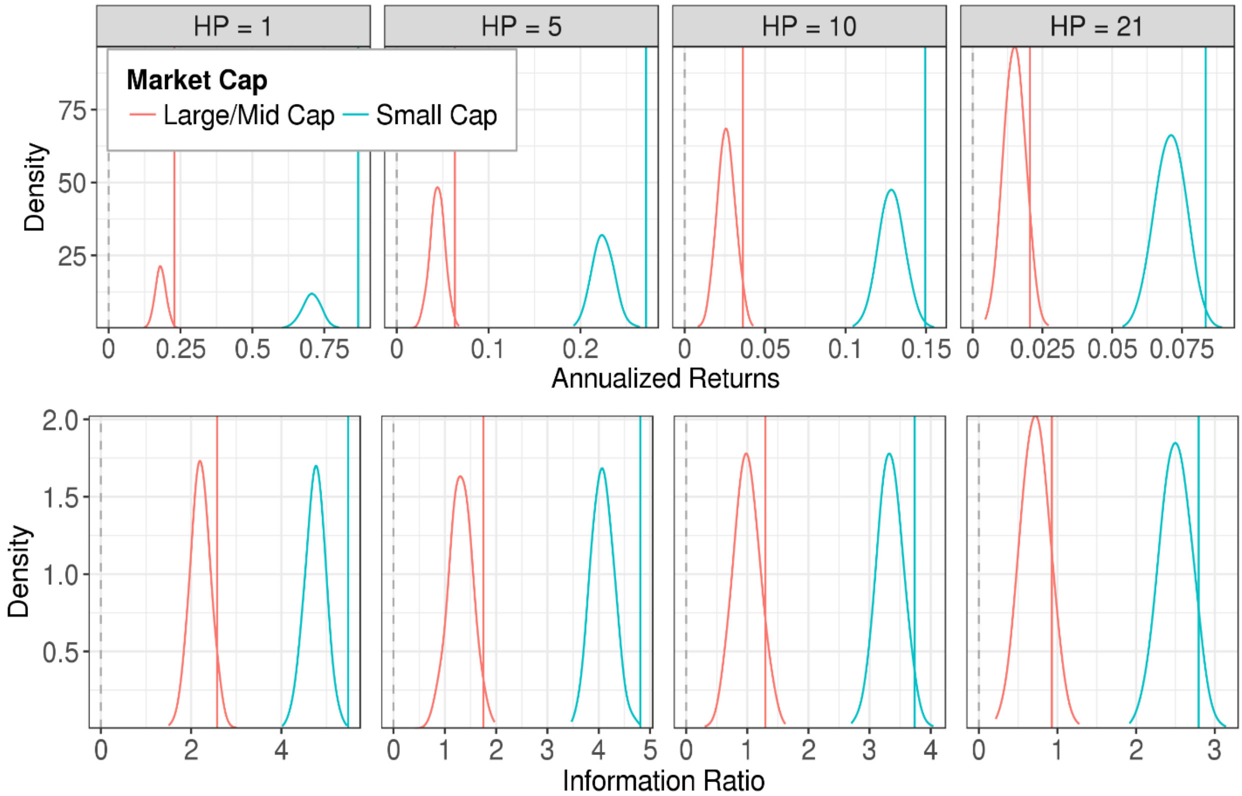

Figure 6: Portfolio performance for ESP's and RSP's separated by market cap (US-R3000). We compare the annualized returns and the Information Ratio of the Extreme Sentiment Portfolios (vertical color lines) with the corresponding distributions of the Random Sentiment Portfolios when separating the companies by market cap. Results for several holding periods (HP) and PS = 20.

Source: RavenPack, May 2019

Figure 6 highlights the annualized return and Information Ratio for ESPs with PS=20, split by market cap. These are compared with their Random Sentiment Portfolio distribution benchmarks. By splitting into Large/Mid-cap and Small-cap portfolios, we see a clear delineation in performance, with both higher annualized returns and Information Ratios across the latter universe. Results are consistent for different portfolio sizes[15].

Figure 7 shows the Cumulative Returns across the two market cap universes, for various portfolio sizes. Performance is consistent across the entire backtesting period. Clearly, Small-caps show the strongest performance[16], but attractive results are achieved across both universes.

[15] Results for other PS available upon request.
[16] Note that we take a simple approach in our evaluation, applying zero liquidity constraints and trading costs.

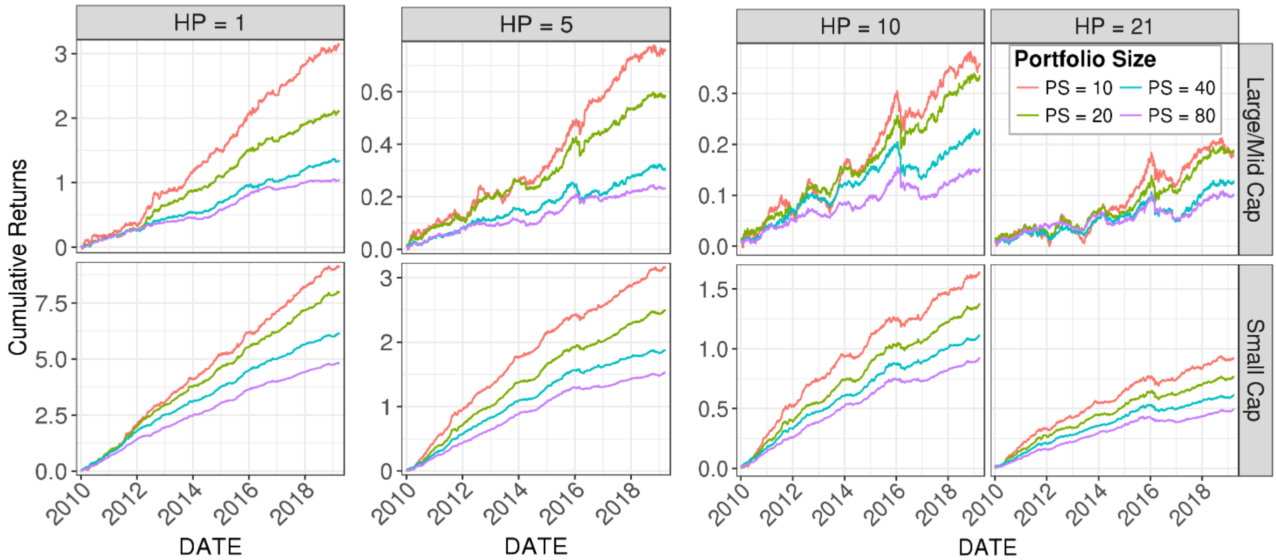

Figure 7: Cumulative Returns for ESP's and RSP's separated by market cap (US-R3000). Cumulative Returns of Extreme Sentiment Portfolios for different portfolio sizes (PS) and holding periods (HP) when separating the universe by market capitalisation.

Source: RavenPack, May 2019

While the ESPs show strong statistical significance across both portfolios for short holding periods of 1-5 days, only the Small-cap performance is significant at 10-21 day HPs. However, the sentiment selection represented by the RSPs in Figure 6 still provides statistically significant performance compared to random selection across all universes and holding periods: all the distributions show a clear positive shift. Compellingly, the lower significance of the ESP performance for Large/Mid-cap means that a trader could take advantage more easily of our sentiment selection by adding additional criteria in order to beat the ESP, shown by the larger area of the RSP distribution that surpasses the ESP line.

In the following section, we expand further on the stock selection idea by introducing a stock screener that defines a subset universe from which we build our final portfolios. By introducing this additional step, we expect to shift the annualized return and Information Ratio distributions of our sentiment portfolios further into positive territory, potentially moving beyond the fixed selection defined by the Extreme Sentiment Portfolio.

6. Sentiment Stock Screener

In Section 4, we selected stocks using Sentiment Strength and proved that the resulting portfolios had significantly better performance than the randomly generated ones. On top of this unique selection, a discretionary trader will likely take into account many other factors, quantitative or qualitative, when deciding to build a final portfolio. As a result, he may decide to trade only a subset of the suggested equities. In this case, we use the extreme sentiment selection as a Sentiment Stock Screener and build daily baskets of entities, stock screens, where each constituent company is a potential candidate to be included in the final portfolio of smaller size.

A stock screen results in a sub-universe of size S that is formed by the same stock selection protocol used when forming the ESPs. Consequently, if we include all equities listed in this sub-universe in our final portfolio, we would be building a portfolio that coincides exactly with the ESP of size PS=S. Nevertheless, the stock screen is a set of equities that can be freely selected to be part of a smaller size portfolio[17].

In selecting the optimal stock screen, we are faced with a trade-off between flexibility and cost. The larger the screen (i.e. S) is, the more flexibility we have for introducing additional criteria, factors, and constraints that can benefit our final selection. However, this added flexibility increases the cost of fundamental research, as we need to allocate more resources towards covering a larger set of stocks. Subsequently, we investigate the effects of using stock screens of different sizes by comparing the performance of the Random Sentiment Portfolios (RSP) that can be generated from each of them.

By investigating Random Portfolios, we are able to generate performance distributions which represent the most likely outcomes of a trader randomly selecting assets from the stock screens. When selecting assets, we follow the Random Sentiment Portfolio (RSP) approach described previously, where stocks with higher Sentiment Strength have a higher probability of being selected[18].

Figure 8 below summarizes the results for the RSPs generated from the different stock screens, in particular the annualized returns and Information Ratio. Amongst all available stock screens, there are two limiting cases: (1) when the stock screen is of the exact same size as our final portfolio (i.e. S = PS), represented by a vertical line indicating a single possible outcome. (2) When the stock screen is of unrestricted size and, thus, includes all available assets with non-zero sentiment, labeled as S = ALL in Figure 8, which corresponds to the distributions already presented in Figure 3.

[17] By construction, the stock screen is balanced, i.e. with the same number of long and short positions.
[18] As before, the other random portfolios show lower performance even for small stock screens, thus we take advantage of sentiment also for this random selection.

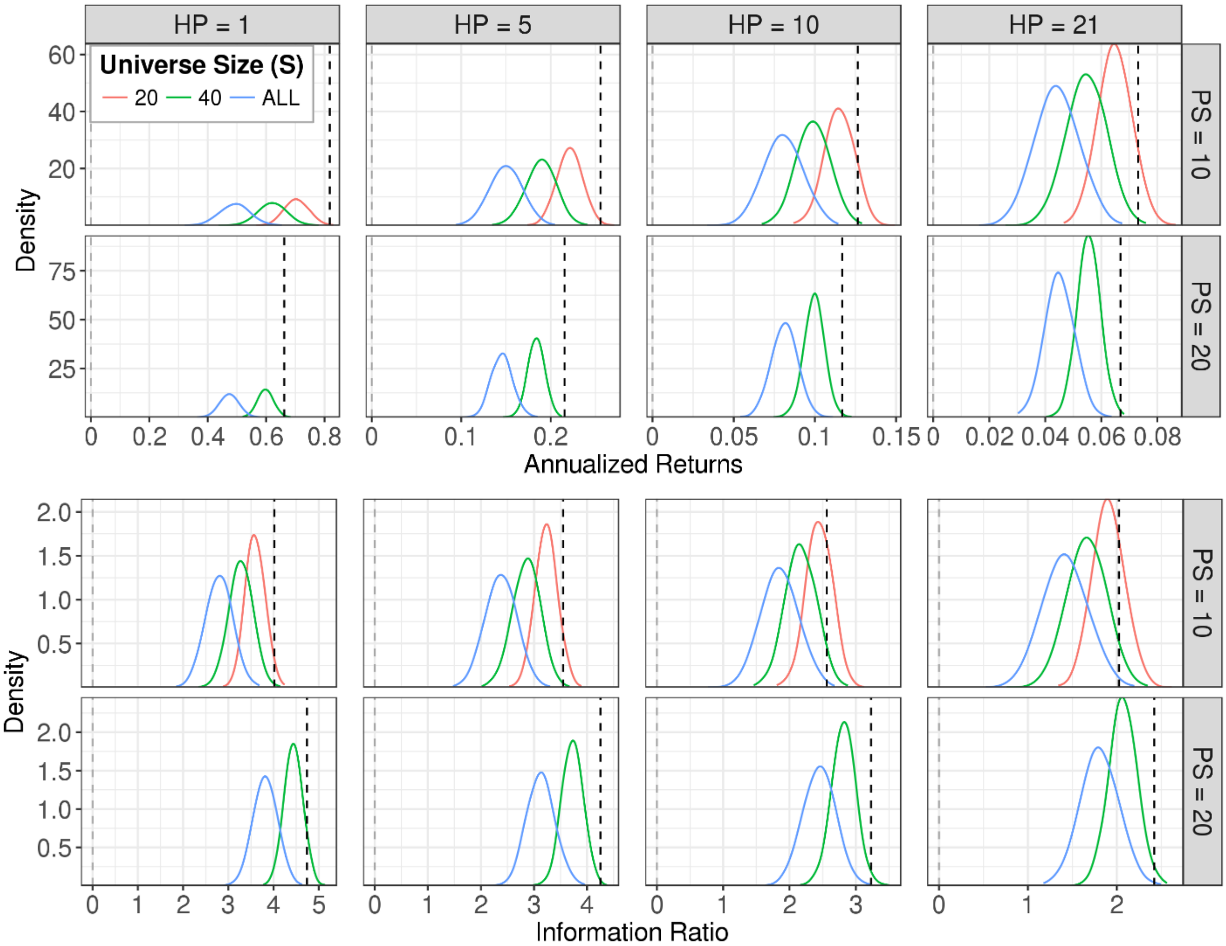

Figure 8: Performance for RSPs built from different size stock screens (US-R3000). Distributions of the annualized returns (top) and Information Ratios (bottom) of the Random Sentiment Portfolios constructed from stock screens of different size (S), for several holding periods (HP), and portfolio sizes (PS). The vertical black line represents the distribution for BS=PS (corresponding to the ESP).

Source: RavenPack, May 2019

As shown in Figure 8, the transition between the two limiting cases is monotonic, with a continuous increase in performance as we move from larger to smaller portfolio sizes. That is, by applying more restrictive stock screens, we also have a higher probability of constructing better performing portfolios. Nevertheless, the distributions also get narrower as S decreases, which is the result of a loss of flexibility in our selected portfolio, and, therefore, the monotonicity has to be broken at some point. This will become more evident as S moves closer and closer to PS, where the distribution, eventually, will collapse into a single outcome. The observed behavior is present across all holding periods.

By implementing a stock screener, we are able to combine the benefits of sentiment selection, ensuring the universe is tilted towards better performing portfolios. This has the advantages of a discretionary process and the added flexibility also provides an opportunity to outperform the ESP. The extent to which the ESP can actually be outperformed depends on the power of the discretionary overlay, as well as on the potential values captured by the area of the tail of the RSP distribution on the right side of the ESP line. From Figure 8, we can observe that this "tail area" increases with the holding period (HP), leaving greater opportunities to outperform the ESP. The same behavior is observed when splitting the universe by market cap, with Large/Mid-Caps showing greater potential value than Small-caps when applying the Sentiment Stock Screener combined with a discretionary overlay (addressed in next section).

Although larger size stock screens allow for more flexibility, as previously described, they are more expensive to implement, and they are also associated with greater risk, due to their higher distributional variance. While they may result in greater opportunities to add value beyond the ESP, the probability of significantly underperforming the lower risk alternatives is also increased. Therefore, we are confronted by a double-balance between resources and risk, on one side, and flexibility, on the other. The optimal choice depends on the investor´s risk profile and the level of confidence it has to pick stocks.

7. Introducing a Proxy Discretionary Overlay

On top of the sentiment signal, there are many other factors or sources of information that can be used to condition the portfolio selection, whether to improve the return, lower the risk, or simply facilitate the execution of the strategy. Obviously, one could include these conditions using a quantitative approach, subsequently increasing the complexity of the model. Nevertheless, we will try to mimic what could be considered a *discretionary overlay*, but in a systematic way, using probability distributions.

The proposed strategies will use the simple models based on news sentiment that we have used so far. However, we will include an additional allocation step that tries to reflect the expertise of the trader, by randomly avoiding some bad performing stocks in the selection process. Like this, we build a Random Sentiment Portfolio with Discretionary Overlay (RSPDO), which will allow us to show the impact on performance if a trader was able to successfully identify and discard a bad performing stock at the pre-allocation stage. The applied procedure aims to proxy a discretionary overlay in a probabilistic way, similar to what we did previously when constructing the RSPs. That is, instead of achieving only a single outcome, we get an entire distribution of outcomes, each representing a portfolio with varying discretionary ability to identify bad future trades.

The portfolio construction, using a discretionary overlay, can be split into two steps:

- First, we select N+1 long and N+1 short daily trades from an initial stock screen (Section 5). We will do so using the Random Sentiment Portfolio selection, which considers Sentiment Strength as a probability weight to pick stocks. We will first use the largest available stock screen (S=ALL) and, afterwards, compare stock screens of smaller size.
- Second, we randomly drop 1 long and 1 short position amongst the selected trades, leaving us with a final portfolio of PS=2N. In the case where we assign an equal "drop probability" to all stocks, a discretionary overlay provides zero edge – and thus corresponds to our benchmark distribution. However, in order to mimic trader skill, the random drop will be performed using the knowledge of the next-day return of each trade. To do so in a well-controlled manner, we will only use the next-day returns to split stocks into two baskets: bad performers – negative (positive) returns for long (short) positions and good performers – otherwise. Once the split is done, we increase the benchmark probability to drop a stock from the bad performing bucket (and, consequently, decrease the probability from the good performing one) linearly with the "skill" probability, p_s.

The skill probability goes from 0, corresponding to the benchmark uniform probabilities, to 1, when the drop is always selected from the bad performing bucket, if non-empty[19]. The probability of being dropped is kept constant across all bucket constituents, i.e. within the bad vs. good performing bucket. This means that the trader does not necessarily choose the worst performer, introducing like this what we could consider a wider range of "trader expertise". That means if, on average, our initial portfolio selection has 50% good trades and 50% bad trades, a trader with zero skill, p_s=0, would select a bad trade 50% of the time and a good trade the other 50% of the time. However, if we introduced a skill probability of 25%, it implies that the trader will correctly drop a bad trade 62.5% of the time vs. 37.5% for a good trade. In reality, our initial portfolio selection is not balanced between good and bad trades. Using sentiment selection, we achieve an accuracy or hit rate greater than 50%, which implies that the real edge added by the skill probability is even smaller[20].
The aim of this exercise is to offer an idea of the potential value-add of including a discretionary overlay, which assumes a positive contribution to the stock selection process. These strategies are not intended for implementation since they clearly introduce look-ahead bias, i.e. the stock drop is

[19] If no stocks are found in the bad performing basket on a given day, we still drop a random stock from the good performing one. This maintains the desired portfolio size but reduces the effective skill probability.
[20] In the most extreme case, where our initial sentiment selection has 100% accuracy, there will be no way of adding value using trader skill. In fact, the trader might have a negative impact as the best performing stock might get removed.

executed using future trade performance. For this reason, we stay conservative by only dropping one long and one short trade. Further improvements could be achieved by allowing for additional drops per day, by increasing the probability of dropping the worst performing trades or, especially when no bad performers are found in a given day, reducing the probability to drop the best performing trades.

Nevertheless, the results from our Random Sentiment Portfolios with Discretionary Overlay already show promising results: if the trader is able to successfully include new valuable information in the stock selection process, a statistically significant boost in performance can be achieved compared to Extreme Sentiment Portfolios (vertical line) across all HPs and PSs. Figure 9 compares the annualized return and Information Ratio distributions for different skill probabilities, built from the stock screen with S = ALL (using all sentiment signals).

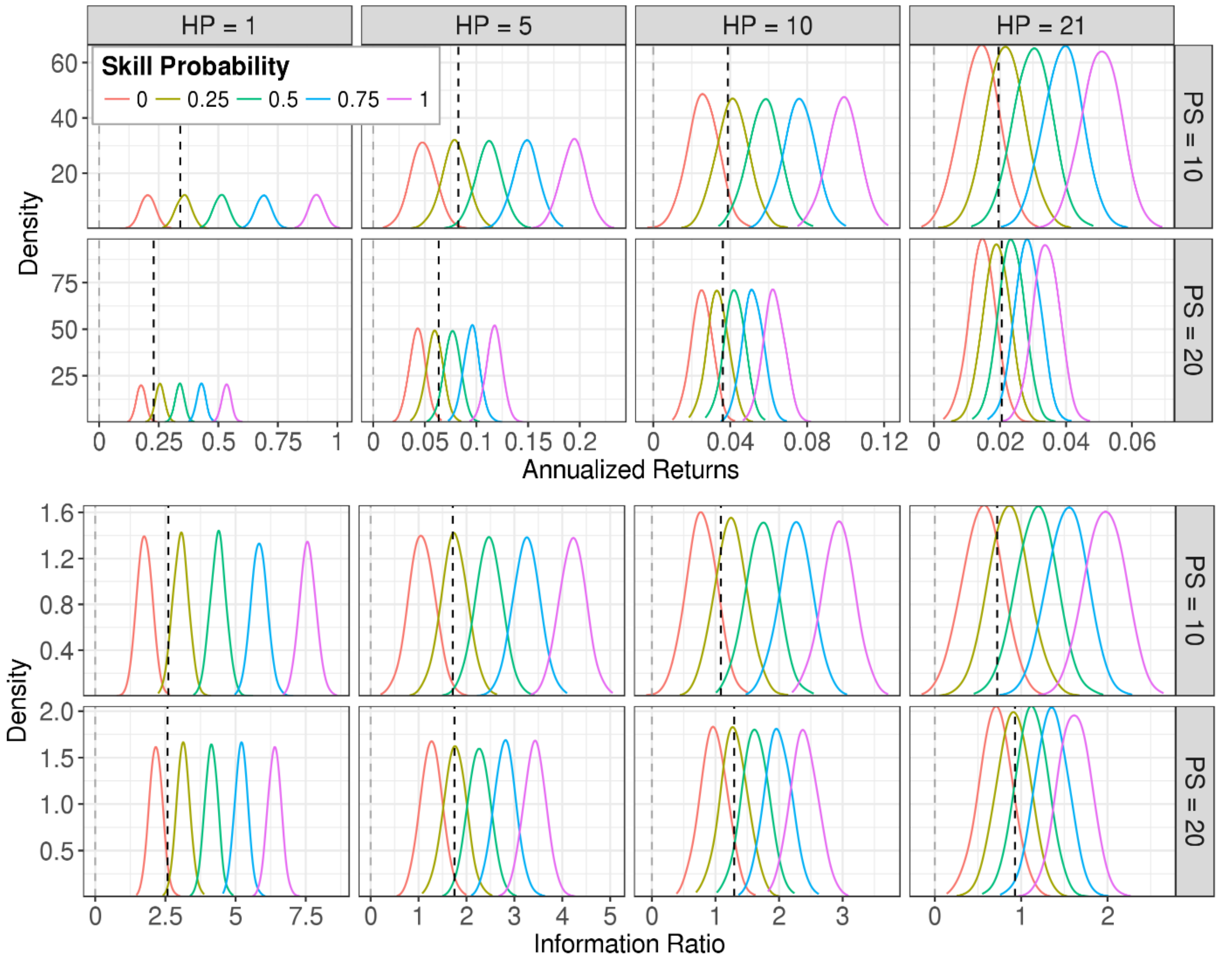

Figure 9: Performance measures for RSPs with discretionary overlay (US Large/Mid-Cap). Distributions of the annualized returns (top) and Information Ratios (bottom) of the Random Sentiment Portfolios for several holding periods (HP) and portfolio sizes (PS). We compare the strategies with the discretionary overlay (RSPDO) using different skill probabilities. The vertical dashed black line represents the performance of the Extreme Sentiment Portfolio (ESP).

Source: RavenPack, May 2019

Clearly, a consistent shift of the entire distribution towards better performance is observed when increasing the skill probability, with larger shifts for shorter holding periods and smaller portfolio sizes. Improvements grow approximately linearly with the skill probability. Compared to the zero skill benchmark, the gain in returns for PS=10 with short HPs ranges from 70% to 340% with increased skill level (i.e. p_s=0.25 and p_s=1). Similarly, we observe a gain in returns across long HP ranging from 44% to 210%. As expected, the relative improvement when dropping a fixed number of stocks becomes smaller as we increase the portfolio size. Furthermore, we notice that in order to beat the Extreme Sentiment Portfolio, in most cases, we need at least a skill probability larger than 0.25. Although Figure 9 focuses on Large/Mid-Cap stocks, our general conclusions hold also for Small-caps, however, with less significant distributional shifts due to an already strong performance from the pure sentiment-driven approach[21].

In the previous section, we found that reducing the size of the stock screen increases the expected performance of the final portfolio. Nevertheless, this comes with the trade-off of having less flexibility to add value applying a discretionary overlay. To understand this relationship better, in Figure 10, we compare two different PS = 10 scenarios, setting S = 12, 20 vs. S = ALL[22], with the latter scenario already considered in Figure 9.

As can be observed from Figure 10, starting off with a smaller sub-universe, the effect of the discretionary overlay is not undermined. In fact, it may even show a slightly larger relative boost. This, together with the fact that performance distributions are already shifted to the right for the smaller stock screen, as described in Section 6, puts more emphasis on the conclusion that the Sentiment Stock Screener can be a simple, but effective way of applying alternative data into an existing discretionary investment process. As a consequence, when using smaller stock screens the edge needed to surpass the performance of the Extreme Sentiment Portfolio is smaller. In particular, for the smallest Stock Screen that still gives some flexibility to the trader to drop 2 stocks, with S=12, very small skill probabilities show distributions shifted above the Extreme Sentiment Portfolio, i.e. with p_s=0.25 the performance distributions are already well above the black dashed vertical line.

[21] Results for Small-cap are available upon request.

[22] We pick 5 long trades and 5 short trades from a sub-universe of either 12 or 20 extreme sentiment stocks (S=12,20) or from the entire universe of stocks with sentiment (S = ALL).

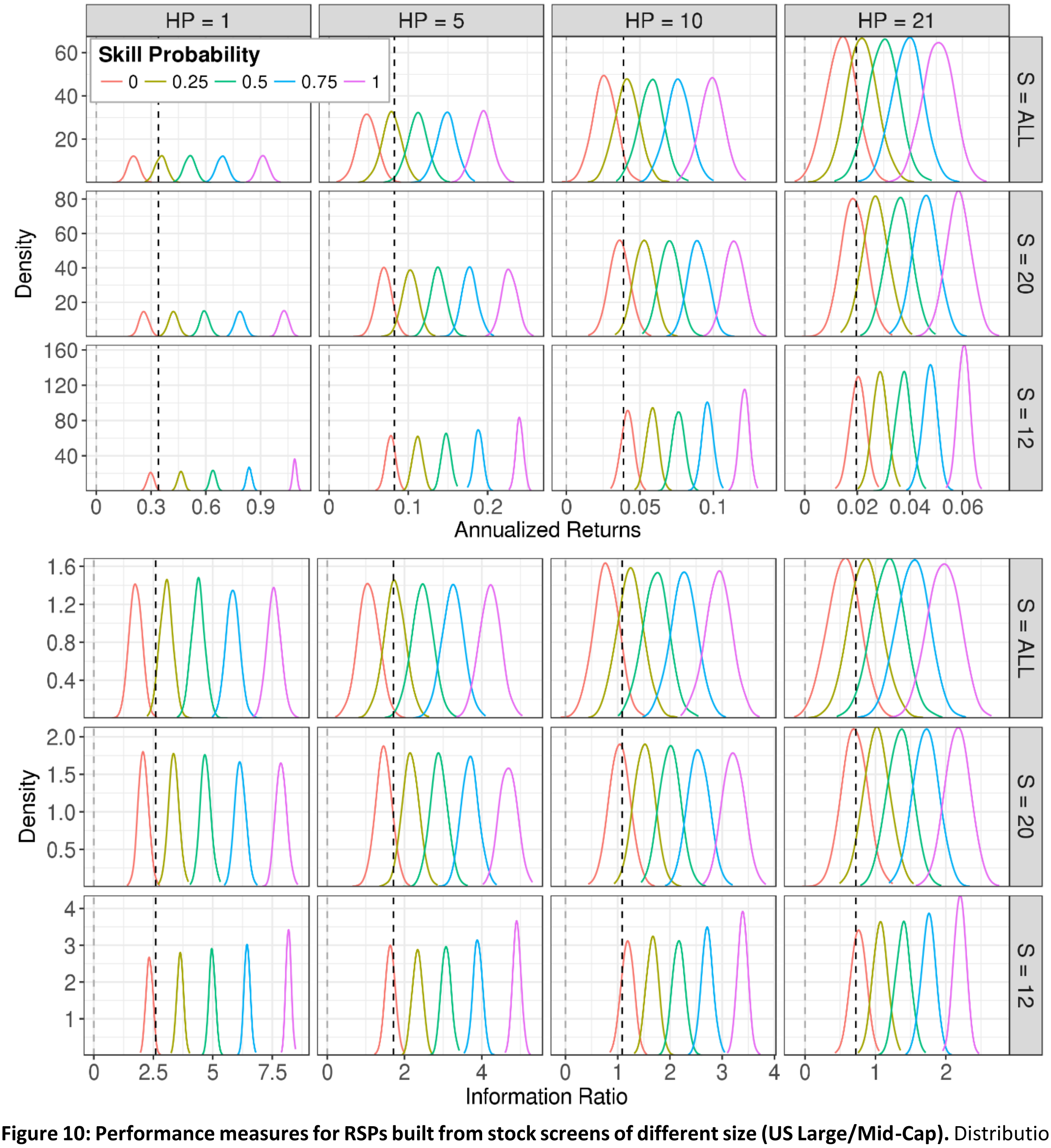

Figure 10: Performance measures for RSPs built from stock screens of different size (US Large/Mid-Cap). Distributions of the annualized returns and Information Ratios of the Random Sentiment Portfolios with discretionary overlay for several holding periods (HP), built from stock screens of different size. We compare different skill probabilities for PS = 10. The vertical dashed black line represents the performance of the Extreme Sentiment Portfolio (ESP).

Source: RavenPack, May 2019

8. Conclusions

In this paper, we introduced a Sentiment Stock Screener using RavenPack Analytics. The objective was to conveniently reduce the stock universe from which a discretionary trader should select his final portfolio. Our stock selection process picks stocks with a high number of news events and high sentiment polarization using RavenPack's Sum Excess Sentiment Indicator, (SESI), which was introduced in previous research [1]. In order to study the performance of our selection, we introduced several benchmark portfolios, utilizing different random selection criteria. Consequently, results could be summarized in a probabilistic fashion, providing information on their statistical significance.

By selecting only stocks with stronger SESI signals, forming what we called the *Extreme Sentiment Portfolios* (ESP), we were able to obtain portfolios that performed significantly better than any of the Random Portfolio benchmarks based on either price or sentiment momentum. We also found that *per trade returns* consistently increased when decreasing portfolio size via more aggressive sentiment filtering, an interesting scenario from a discretionary point of view when considering trading costs.

While our initial analysis is done using the Russell 3000 constituents, our conclusions still hold when separating our universe based on market cap, i.e. Russell 1000 vs. Russell 2000, and on market sectors. However, small-caps clearly outperform when constructing simple portfolios without taking liquidity or trading costs into account.

By using a probability-driven selection approach, we showcased how an approximate *discretionary overlay* could add significant value to our Sentiment Stock Screener, by selecting stocks with extreme sentiment values. Specifically, to mimic *trader expertise*, we allowed the allocation process to manually drop one long and one short position from our final portfolio, with the knowledge that these could potentially contribute negatively to performance. We introduce a skill probability that controls the amount of accuracy added in the *drop* step compared to a random selection. The resulting portfolios are non-tradable due to look-ahead bias; however, they provide interesting insights into the potential value of combining the Sentiment Stock Screener with trader skill or expertise. As an example, we found that dropping only 2 out of 22 daily trades, annualized returns and Information Ratios can be considerably boosted. For example, for a 0.5 skill probability, they can increase by up to 100%.

In conclusion, extreme sentiment selection provides an effective way of reducing the number of companies that need specific trader attention, especially for short horizon trading strategies. The optimal size of the stock screen depends on the trader's capacity to perform deeper fundamental analysis of the constituent companies, his ability to add value to the pure sentiment driven strategy, and the risk he is willing to take when introducing his criteria. While we have found meaningful results, this study was limited to simple models and portfolios. In future research, we plan to improve the sentiment selection process by using information at a more granular level and increasing flexibility by allowing the Sentiment Stock Screener to be more customizable.

9. References

[1] Peter Hafez and Ricard Matas (May 2018), "Effects of Event Sentiment Aggregation: Sum Vs Mean", RavenPack Quantitative Research.

[2] Peter Hafez and Mads Koefoed (April 2017), "Introducing RavenPack Analytics for Equities", RavenPack Quantitative Research.

[3] Peter Hafez and Mads Koefoed (May 2017), "A Multi-Topic Approach To Building Quant Models", RavenPack Quantitative Research.

[4] Peter Hafez and Jose A. Guerrero-Colón (November 2017), "Hedging Sentiment Signals with MSCI Barra Risk Models", RavenPack Quantitative Research.

[5] Peter Hafez and Francesco Lautizi (February 2017), "Abnormal Media Attention Impacts Stock Returns", RavenPack Quantitative Research.

[6] Peter Hafez and Francesco Lautizi (June 2018), "News Sentiment Everywhere: Trading Global Equities", RavenPack Quantitative Research.

[7] Patrick Burns (1972). "Random Portfolios for Evaluating Trading Strategies". Advances in Computational Management Science, vol 9. Springer.

[8] Peter Hafez and Francisco Gómez (May 2018), "How Intraday News Patterns Dictate the Implementation of Sentiment Strategies", RavenPack Quantitative Research.

Appendix – Building a News Sentiment Stock Screener using the RP Self-Service Platform

Below, we describe the steps needed in order to carry out a daily selection of entities that resembles the stock screener proposed in this paper, when using the RavenPack Self-Service Platform. The objective is to rank entities cross-sectionally, on a daily basis, and select only the ones with higher sentiment indicator strength.

Generate the Dataset & create the Sentiment Indicator

1) From **ravenpack.com** open the menu by clicking on the top-left icon (3 horizontal bars);
2) From the **Datasets** section in the menu, click the add icon (+);
3) You will be prompt to the dataset builder, where you will see a horizontal Tab menu under the name (*Untitled Dataset*) of your still-empty dataset. It is suggested that you give it a sensible name so you can save it for later use;
4) Inside **Filters**, Select the **Entities** of interest by Entity Name (you can also add *rp_entity_id* or upload a .csv file with a list of entities of interest);
5) We only consider highly relevant events. In the **Relevance** section we select $EVENT\ RELEVANCE \geq 90$;
6) We apply a **Novelty** filter by only including events with $EVENT\ SIMILARITY\ DAYS \geq 90$;
7) In the **Sentiment** section, we filter out neutral content by selecting events with $EVENT\ SENTIMENT\ SCORE\ \neq 0$;
8) Create the Indicator of interest using the **Functions** tab:
   a) In this case, select **Sum** function;
   b) In the **Name** field, enter a name (say SUM_I);
   c) In the **Class** field, select *Daily* (daily average);
   d) In the **Input Field**, select *EVENT_SENTIMENT_SCORE*;

Note that the Sum Sentiment Indicator that we have built here is not the same indicator used as part of the paper. The difference resides in the fact that SESI does include a term that removes a daily sentiment bias, which is not an operation currently supported in the platform.

Cross-sectional Ranking & top/bottom selection

1) Add a new function going to the **Functions** tab:
   a) Select the **Ranking** function;
   b) In the **Name** field, enter a sensible name (say RANK_P);
   c) In the **Order By** field, select the Indicator created above (SUM_I);
   d) In the **Direction** field, select DESCENDING;

   This will rank daily and in descending order all the entities with sentiment indicator.
2) Add a condition using the **Conditions** tab:

a) Select the Less Than **Comparison** (drop-down list);
b) In **Field**, select the name of the ranking function (RANK_P);
c) In the **Value** field, select the number of entities that you want for the Stock Screener (for example 20 for the Top 20 sentiment stocks);

This will filter the Universe daily, so that only the desired number of entities will be downloaded.

For the selection of the entities with the lowest Sentiment Indicator values, we can proceed in the same way, but choosing ASCENDING in step 1.d instead, which ensures that the Indicator is ranked in the contrary order. Note that both top and bottom entities have to be downloaded in separate datasets.

Download the dataset

1) Under the Dataset Preview, select the **start** and **end date** with the desired timestamp. The time will mark the range used for the daily aggregation: for example, if the market of interest closes at 20:00 UTC, by selecting 19:30 UTC every daily bucket will include data from the previous date after 19:30 UTC until 19:30 UTC on the current day, i.e. 30 minutes before market close. Note that you can also select the timezone that better suits you;
2) Click on the **Download** button on the top-right of the page and select the format you want to download;
3) When the dataset is ready, you will receive an email with a link to download the data.

Terms of Use

This White Paper is not intended for trading purposes. The White Paper is not appropriate for the purposes of making a decision to carry out a transaction or trade. Nor does it provide any form of advice (investment, tax, legal) amounting to investment advice, or make any recommendations regarding particular financial instruments, investments or products. RavenPack may discontinue or change the White Paper content at any time, without notice. RavenPack does not guarantee or warrant the accuracy, completeness or timeliness of the White Paper

You may not post any content from this White Paper to forums, websites, newsgroups, mail lists, electronic bulletin boards, or other services, without the prior written consent of RavenPack. To request consent for this and other matters, you may contact RavenPack at research@ravenpack.com.

THE WHITE PAPER IS PROVIDED "AS IS", WITHOUT ANY WARRANTIES. RAVENPACK AND ITS AFFILIATES, AGENTS AND LICENSORS CANNOT AND DO NOT WARRANT THE ACCURACY, COMPLETENESS, CURRENTNESS, TIMELINESS, NONINFRINGEMENT, TITLE, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE OF THE WHITE PAPER, AND RAVENPACK HEREBY DISCLAIMS ANY SUCH EXPRESS OR IMPLIED WARRANTIES. NEITHER RAVENPACK NOR ANY OF ITS AFFILIATES, AGENTS OR LICENSORS SHALL BE LIABLE TO YOU OR ANYONE ELSE FOR ANY LOSS OR INJURY, OTHER THAN DEATH OR PERSONAL INJURY RESULTING DIRECTLY FROM USE OF THE WHITE PAPER, CAUSED IN WHOLE OR PART BY ITS NEGLIGENCE OR CONTINGENCIES BEYOND ITS CONTROL IN PROCURING, COMPILING, INTERPRETING, REPORTING OR DELIVERING THE WHITE PAPER. IN NO EVENT WILL RAVENPACK, ITS AFFILIATES, AGENTS OR LICENSORS BE LIABLE TO YOU OR ANYONE ELSE FOR ANY DECISION MADE OR ACTION TAKEN BY YOU IN RELIANCE ON SUCH WHITE PAPER. RAVENPACK AND ITS AFFILIATES, AGENTS AND LICENSORS SHALL NOT BE LIABLE TO YOU OR ANYONE ELSE FOR ANY DAMAGES (INCLUDING, WITHOUT LIMITATION, CONSEQUENTIAL, SPECIAL, INCIDENTAL, INDIRECT, OR SIMILAR DAMAGES), OTHER THAN DIRECT DAMAGES, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. IN NO EVENT SHALL THE LIABILITY OF RAVENPACK, ITS AFFILIATES, AGENTS AND LICENSORS ARISING OUT OF ANY CLAIM RELATED TO THIS AGREEMENT EXCEED THE AGGREGATE AMOUNT PAID BY YOU FOR THE WHITE PAPER. JURISDICTIONS DO NOT ALLOW THE EXCLUSION OR LIMITATION OF LIABILITY FOR DAMAGES OR THE EXCLUSION OF CERTAIN TYPES OF WARRANTIES, PARTS OR ALL OF THE ABOVE LIMITATION MAY NOT APPLY TO YOU.

These Terms of Use, your rights and obligations, and all actions contemplated by these Terms of Use will be governed by the laws of New York, NY, USA and You and RavenPack agree to submit to the exclusive jurisdiction of the New York Courts. If any provision in these Terms of Use is invalid or unenforceable under applicable law, the remaining provisions will continue in full force and effect, and the invalid or unenforceable provision will be deemed superseded by a valid, enforceable provision that most closely matches the intent of the original provision.