

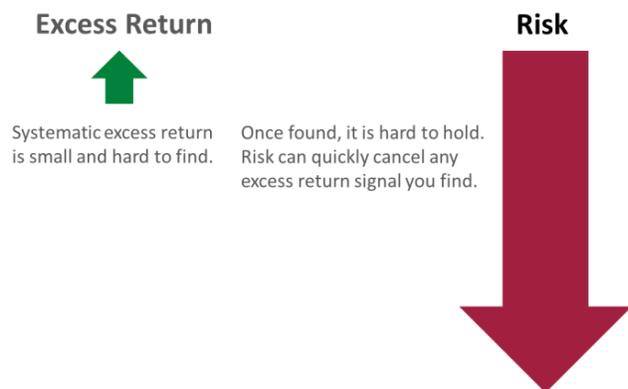
Price Sentiment Analysis and Indicators

Technical White Paper

Overview

Valspresso is a financial technology and investment strategy development firm built on the foundation of over a decade of research and innovation. Valspresso's goal is to secure the future of ordinary people by partnering with active managers to generate persistent Alpha, while reducing risk. By focusing on ordinary investors who are trying to save for retirement and their children's education and to reduce the suffering caused by market swings, our system produces a set of Indicators that have predictive power and allow managers such as you to easily generate persistent alpha, while reducing risk. Our unique approach is based on the idea of quantifying market psychology and deep fundamental analysis, which will be described later in this document.

The stock market is a noisy and often volatile environment which results in substantial risk for trading strategies. Active trading strategies seek to find and capture excess return, however, systematic excess return is small and hard to find. Once found, it is hard to hold. Market risk can quickly cancel any excess return signal you may find. In this paper we illustrate how Valspresso's Indicators and classification system can be used to predict and capture excess return by reshaping risk.



In this paper, we present a use case (see "Screener.2" classification) designed to be used as a starting point for building your own strategies. That strategy delivered rolling 3 year annualized Alpha between 2.7% and 8.0%.

The Indicators described here are particularly effective for fundamental managers, who need a starting point when analyzing multiple securities to build and deploy alpha-generating stock selection strategies. Additionally, the Valspresso Indicators are cost-effective for pension funds and other asset owners who may be looking to insource their active equity investment function but have limited capacity to hire a team of analysts.

This paper reviews the philosophy of Valspresso and its founder, gives a deep description of the Indicators, and reviews the evidence of the predictive and alpha-generating power of the signals. We encourage portfolio managers and analysts to evaluate the alpha value of our data feed for themselves via a free subscription to our historical data at FactSet: <https://open.factset.com/products/sentiment-and-fundamental-indicators/en-us>.

Contents

Overview.....	1
About Valspresso.....	3
The Company.....	3
Founder.....	3
Company History.....	3
Solution.....	4
Theoretical Framework of Market Psychology.....	4
Price Sentiment Analysis.....	5
Fundamental Analysis.....	12
Investment Philosophy.....	15
Testing Short-Term Behavioral Characteristics.....	16
Using Behaviors to Build Strategies.....	26
Other Use Cases.....	29
Data Processing.....	30
Sources of Data.....	30
Coverage.....	30
Frequency.....	30
Point-in-Time Data.....	31
Dictionary.....	31
Sentiment Indicators Data Feed.....	35
Product Overview.....	35
Version History.....	35
File Structure.....	35
Coverage and Frequency.....	36

About Valspresso

The Company

Valspresso, Inc. is a financial technology and investment strategy development firm based in Reston, Virginia. Its two decades of research and innovation culminated in the grant of a U.S. patent and the development and deployment of a suite of rules-based artificial intelligence software applications that automates the entire asset management process — from analysis to portfolio construction to trade execution — with a clear objective: reduce risk and improve returns.

Founder

Reginald Nosegbe, CPA, is the inventor of a theoretical framework that seeks to reduce the risk and improve the return of investing in the stock market. Reginald is an accountant, inventor, systems engineer, entrepreneur, student of economic history, and passionate advocate for investors trying to build and protect their investments. He has years of experience in risk management, regulatory compliance, internal controls, strategic consulting services, and investment management with firms including PricewaterhouseCoopers and Merrill Lynch.

Company History

As a student at the University of Virginia in the mid-1990s, Reginald became deeply concerned about extreme market risk and 10 major market failures, including the 1929 market crash. It was more than the failures themselves that concerned him. Rather, it was the impact that those failures had on the lives of so many people – people who worked for years to achieve financial stability and accumulate assets, only to have that security undone, sometimes in a matter of days.

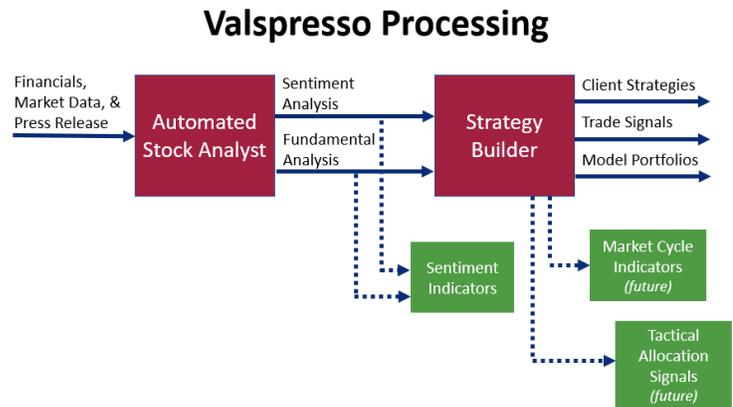
Motivated by this seemingly intractable problem in 1997, he designed an independent study course to create a mathematical model that would objectively quantify market sentiment. His initial findings were presented to a panel of professors and an audience of students. That resulted in an award for the research, analysis, and findings. He went on to get a master's degree in systems engineering from the University of Virginia. In 2007, Reginald filed a patent application for a "Stock Method for Measuring and Assigning Precise Meaning to Market Sentiment." The patent was granted in 2011.

Shortly after the crash of 2008, Reginald recruited a team of technology and financial professionals. Valspresso built an end-to-end suite of applications based on Reginald's theoretical framework. These applications were built on Valspresso's proprietary technology platform that consists of a robust point-in-time historical database, an expert system, diagnostic capabilities, and research tools. On top of that platform, Valspresso has built an automated analyst, a strategy development environment, investment operations modules, model portfolios, and proprietary indicators.

Solution

Valspresso's technology analyzes all publicly traded companies on the U.S. major exchanges every day. This daily automated analysis includes a patented price sentiment analysis as well as deep fundamental analysis.

Valspresso's Indicators empower subscribers to design and deploy alpha-generating strategies. Our proprietary technology and patented methods make it easy for analysts, and portfolio managers to integrate our indicators and signals into new or existing trading strategies with the objective of reducing risk and improving returns.



Theoretical Framework of Market Psychology

Valspresso's approach to investing is based upon our founder's research into market psychology. From that research, he developed the "Free and Competitive Market Equilibrium" hypothesis which is stated as follows:

"In a free and competitive market, the value that efficient and effective companies produce, shall, in the long run, approximately equal the value that the market (investors) demands."

Basic Assumptions

1. The market is free and competitive
2. Companies are efficient and effective
3. Companies' earnings can be independently verified
4. Companies will continue as a "going concern"
5. Investors expected return can be determined

Essentially, this hypothesis posits that when investors make buy and sell decisions on an open market, they are always trying to discover an equilibrium price for a security based upon expectation of earnings growth (or decline) and other emerging information. Investors, and the market in general, will sometimes overshoot this equilibrium price and sometimes undershoot, but will always seek equilibrium. Our Indicators are built to determine whether a stock's current price reflects under-reaction or over-reaction to market information, including recent earnings, and makes an implicit forecast. Our results show that this can be an effective way of building an active investment strategy.

Our hypothesis and assumptions enabled us to derive a series of mathematical models that decompose stock price into the portion that is attributable to underlying operations and the portion that is attributable to future expectations of growth, and to classify all companies traded on the

major exchanges as to whether the market is pessimistic, optimistic, or exuberant about their future operating performance (earnings). Price Sentiment Analysis is a key component of Valspresso's solutions.

Price Sentiment Analysis

Most sentiment analyses try to measure how people "feel" by analyzing what they "say," and use that as a proxy for what people will "do" regarding investment decisions. We all know that what people "feel", "say", and "do" are not always aligned. This is especially true in competitive marketplaces.

We believe that the best measure of investors' sentiment is the price they are willing to pay for a stock. Implicit in the price that investors pay for a stock on any given day is a forecast of a company's future operating performance (earnings). Valspresso was granted U.S. patent 7,966,241 for this method of measuring price sentiment.

Our primary metrics for price sentiment analysis are Sentiment Index (SI) and Sentiment Quotient (SQ). Our patent describes the following functions used to calculate SI and SQ:

Sentiment Index (SI)

$$SI = \frac{PPS \times R - EPS}{EPS}$$

When EPS is equal to 0, then 0.01 is used as a proxy for EPS

Sentiment Quotient (SQ)

$$SQ = \int_0^{SI} \frac{1}{(SI + 1)^2} dsi$$

Change in Sentiment Quotient (SQΔ)

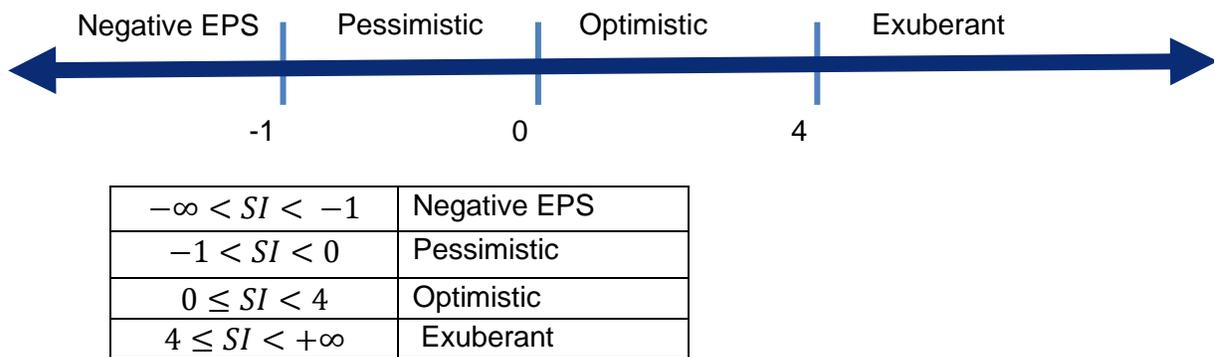
$$SQ\Delta = \int_{SI(TimeA)}^{SI(TimeB)} \frac{1}{(SI + 1)^2} dsi$$

Together, the above mathematical models play key roles in Valspresso's proprietary classification system and the development of alpha-generating strategies.

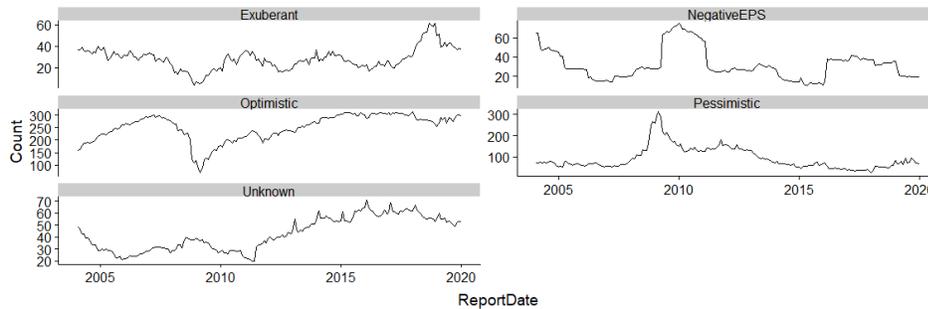
The input into Sentiment Index (SI) includes stock price (PPS), GAAP earnings per share (EPS), and required rate of return (R). Required rate of return is calculated using Beta (36-month Beta calculated daily), U.S. Treasury bill rate (3 month), and market rate of return (average annual return of S&P 500). By looking at the formula, it is easy to see what would happen to SI should an input increase or decrease. For example, an increase in price would increase SI (assuming all

other inputs are constant). This would indicate that the market is more optimistic about future performance. Conversely, if earnings increased, then SI would decrease. Understanding how the interplay of changing input contributes to the market psychology (or sentiment) of price has been a subject of our research for years.

One of the unique attributes of SI and SQ is that their values are on a scale that pivots around the fixed value of 0. Equating SI and SQ to 0 as a fixed frame of reference is the mathematical representation of Free and Competitive Market Equilibrium. This allows the automated analyst to objectively measure the amount of sentiment in companies' stock price and classify them as follows:



Number of S&P 500 Components per Fundamental Classification per Day



Number of Large Cap Stocks per Fundamental Classification per Day

Large cap stocks are those with market capitalization greater than \$4B



SI is used to quantify expected earnings growth. For example, an SI value of 4 means that the company's stock price reflects an expectation that earnings will grow by a factor of 4. An SI value of -0.5 means that the company's stock price reflects an expectation that earnings will decline by a factor of -0.5. The interpretation of SI depends on the sign of EPS and SI. The rules are as follows:

- When $SI > 0$ then
 - SI represents a factor by which earnings are expected to increase
- When $SI = 0$ then
 - Earnings are expected to remain constant
- When $SI > -1$ and $SI < 0$ then
 - SI represents a factor by which earnings are expected to decrease
- When $EPS < 0$ then
 - $SI \leq -1$
- When $SI \leq -1$ then
 - $-SI$ represents a factor by which earnings are expected to increase
- When $EPS = 0$ then
 - SI would be undefined; however, in that case we use 0.01 as a proxy for the value of EPS in order to be able to calculate an SI value.

SQ is used to quantify the portion of a company's stock price that is attributable to future earnings growth. For example, an SQ of 0.8 implies that 80% of the company's stock price is attributable to expectation of future earnings growth and 20% to current fundamentals. The interpretation of SQ depends on the sign of EPS and SQ. The rules are as follows:

- When $SQ > 0$ then
 - SQ represents the portion of stock price attributable to expectation of future earnings growth
- When $SQ < 0$ then
 - SQ represents the portion of earnings that is not reflected in the current stock price
- When $SI = -1$ then
 - SQ is undefined

Obviously, no one knows the future of a stock price, but Valspresso's Sentiment Indicators provide a useful and powerful framework to assess investors' optimism or pessimism reflected in a stock price. The uniform applicability and standardized values mean that they can easily be integrated into new or existing trading strategies with the objective of reducing risk and improving returns. Valspresso has used the sentiment framework to deliver sustained alpha and lower beta in both simulated and live portfolios.

Descriptive Statistics

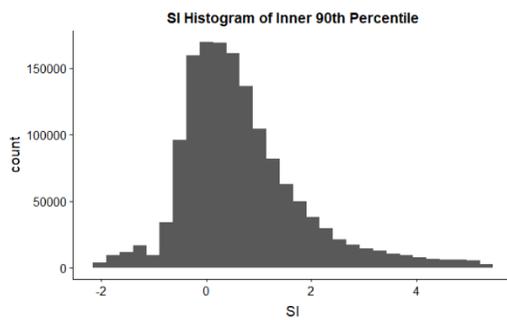
This section includes descriptive statistics for SI and SQ for the components of the S&P 500 from 2004 to 2019.

In some cases, outliers have been removed. Our definition of an outlier is the bottom and top 5%, therefore, the inner 90th percentile of data is included on the charts.

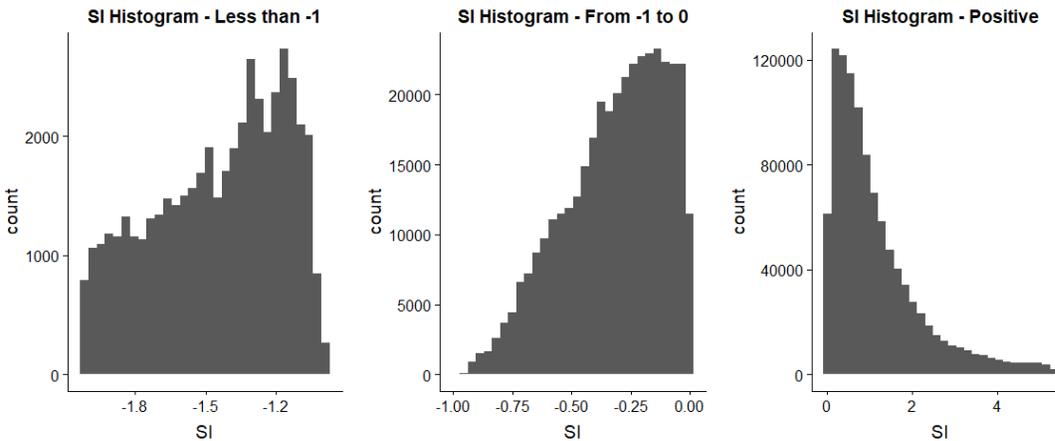
Sentiment Index (SI)

SI is not normally distributed so statistics such as mean and standard deviation are not meaningful. To understand the distribution of this indicator, refer to the charts below.

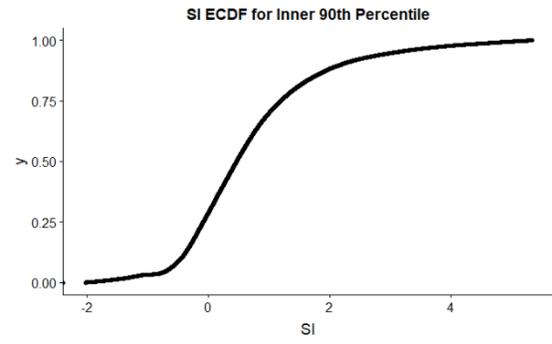
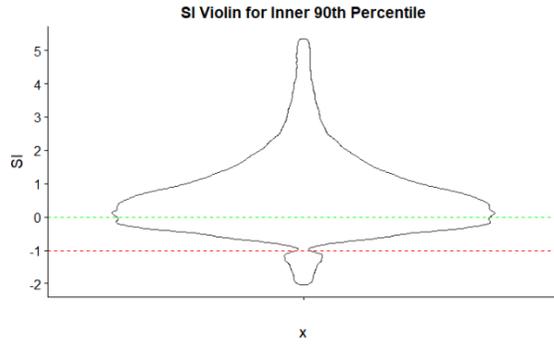
Below is a histogram of SI with outliers removed.



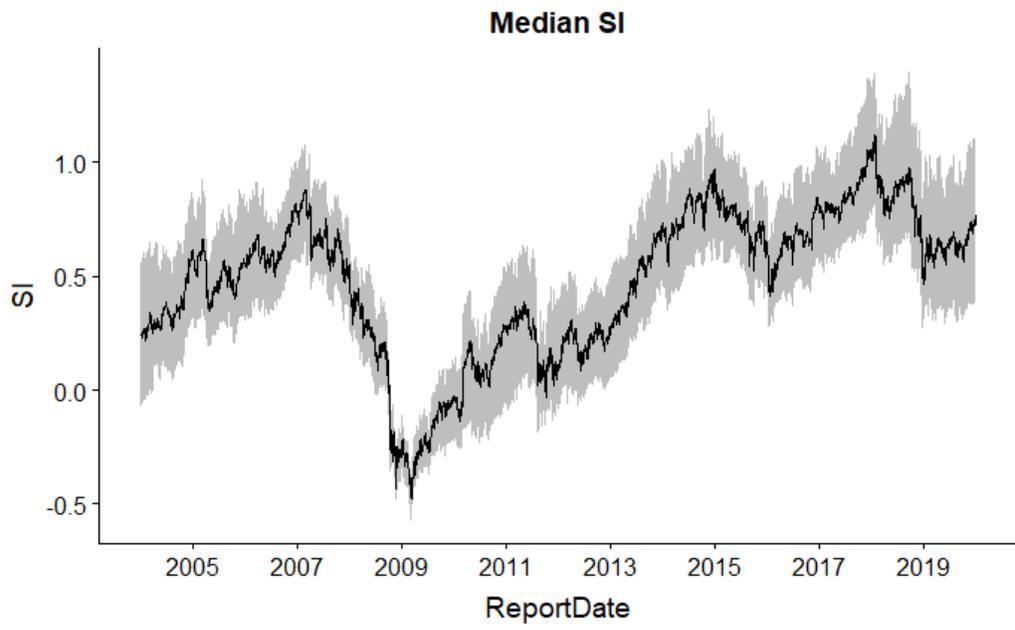
As SI has special cases, the histogram has been divided along those inflection points.



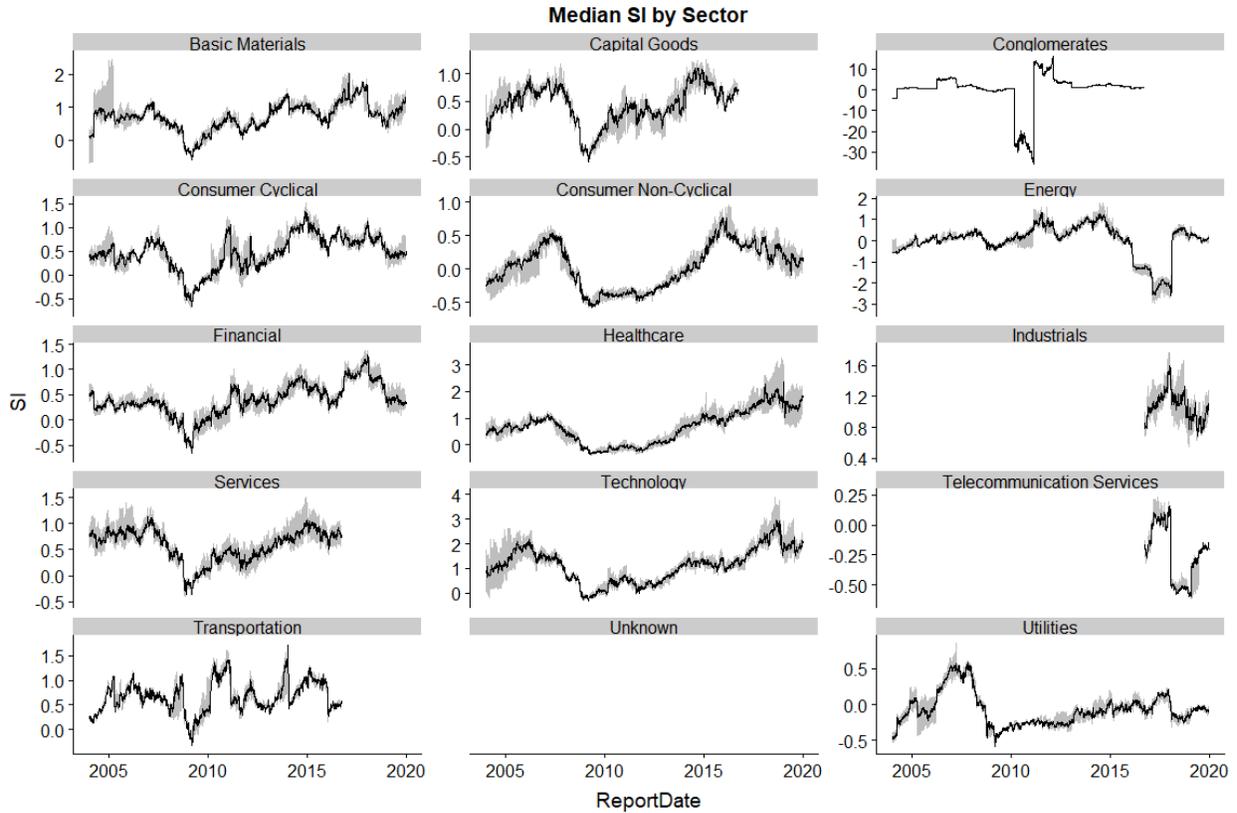
A violin plot of SI is helpful for seeing the distribution behavior. Horizontal lines at 0 and -1 are included to visually indicate the inflection points of the SI function as described above. Outliers have been excluded. To the right, you will find the Empirical Cumulative Distribution Function for SI with outliers excluded.



Below is the SI median over time. The lower part of the grey ribbon is the 40th percentile, while the upper part is the 60th percentile.

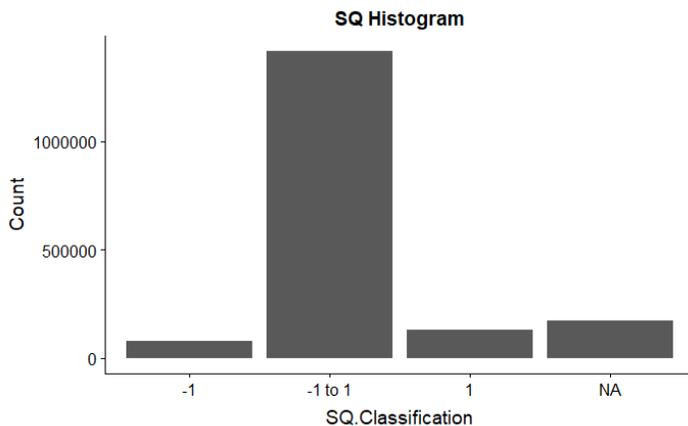


The same data as above has been grouped by sector to see how the various sectors have behaved historically.

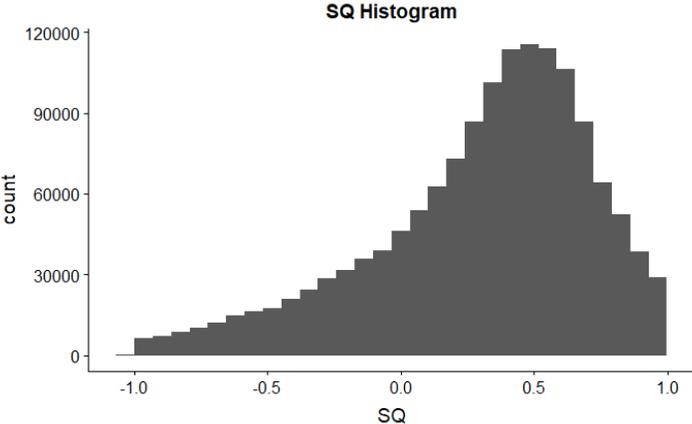


Sentiment Quotient (SQ)

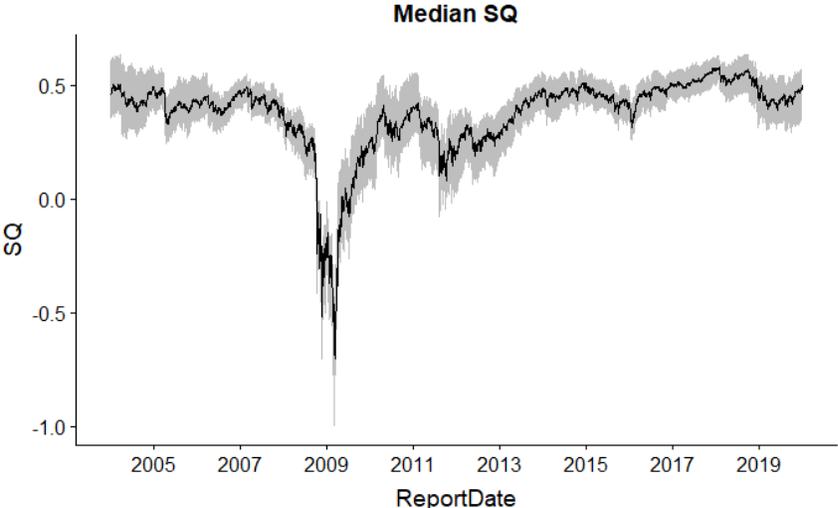
SQ is not normally distributed, so statistics such as mean and standard deviation are not meaningful. In addition, there are special cases for SQ at -1 and 1. Values for SQ fall into 4 categories: -1, between -1 and 1, 1, and NA. An SQ of NA indicates there was insufficient information or history to calculate an SQ.



It only makes sense to visualize the distribution for values between -1 and 1 (non-inclusive).



Below is the SQ median over time. The lower part of the grey ribbon is the 40th percentile, while the upper part is the 60th percentile.



Fundamental Analysis

To provide context and veracity for the use of Sentiment, Valspresso's Automated Stock Analyst performs deep financial analysis of publicly traded U.S. companies. This context is important to manage your strategy's risk. Let's say that one of your strategy's holdings is a stock with a high expected earnings growth. If that company's financial quality is poor, it is less likely to achieve or sustain that growth and therefore should be considered riskier. In that case, additional risk mitigation measures would need to be employed to achieve investment objectives.

Our system automatically assesses profitability, solvency, and management by inspecting each public company's financial statements and press releases filed with the Securities and Exchange Commission (SEC) and classifies the quality of companies' financials as **High**, **Medium**, or **Low**.

The concept of "Simplicity" underlies our approach to fundamental analysis. That is, the expert rules that drive our analysis must be: 1) transparent, 2) simple enough for both individual and institutional investors to understand, and 3) rigorous and capable of handling the complexity of automated investment decision making.

Our assessment of the health of companies begins with automated fundamental analysis using Valspresso's proprietary automated analyst. On a daily basis, the software analyzes the financial statements of publicly-traded U.S. companies based upon 3 simple but fundamental questions:

1. Is the company profitable from its core operations?
2. Can the company pay its bills from cash on hand?
3. Is the company structured to reward investors?

Each of the above questions is answered daily by programmatically inspecting each public company's financial statements filed with the Securities and Exchange Commissions (SEC). These include filings such as 10-K, 10-Q, & 8-K. Financial statements tell a complicated and nuanced story, which is often missed when simply looking at the top and bottom lines. Valspresso's Automated Analyst looks deeply at related information on the financial statements to cross-check those high-level values to compose a holistic understanding of the health of the company. The factors of financial health underpin the long-term success of any business. Regardless of the size of companies or the industry sector in which they operate, those health factors are relevant. The automated assessments are applied uniformly to all companies in all sectors and across all market conditions. Valspresso covers all publicly traded U.S. companies and assesses their financial health every day.

Question 1 inspects values such as net income and cash flow. If those values indicate the company is profitable and solvent then question 1 is considered "passed," otherwise it is considered "failed."

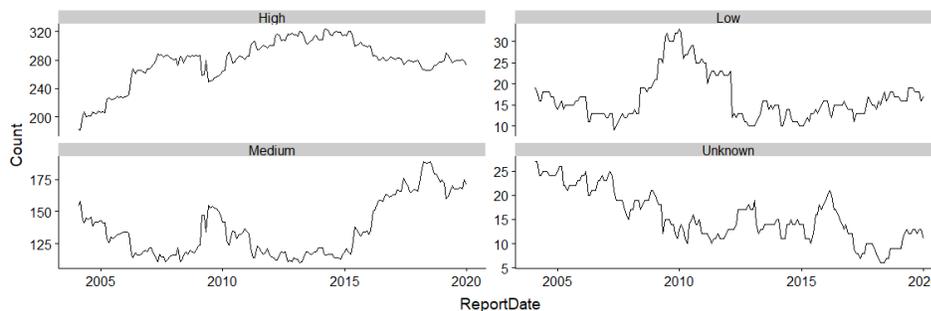
Question 2 inspects values such as assets and liabilities. If those values indicate the company has sufficient liquidity then question 2 is considered "passed," otherwise it is considered "failed."

Question 3 inspects values such as equity, net income, cash flow, and assets. If those values indicate sufficient return on equity then question 3 is considered “passed,” otherwise it is considered “failed.”

Using the pass/fail status to those questions as input on a daily basis, the software classifies each company at various points in time as:

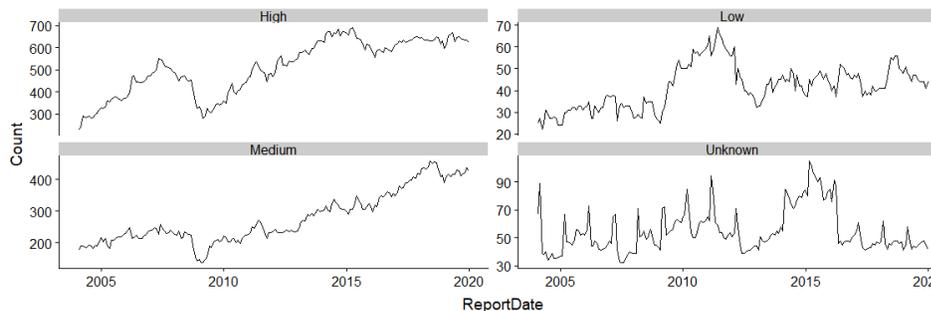
- “Green” or “**High Quality**” (passed all three questions)
- “Yellow” or “**Medium Quality**” (failed one question)
- “Red” or “**Low Quality**” (failed two or more questions)

Number of S&P 500 Components per Fundamental Classification per Day



Number of Large Cap Stocks per Fundamental Classification per Day

Large cap stocks are those with market capitalization greater than \$4B

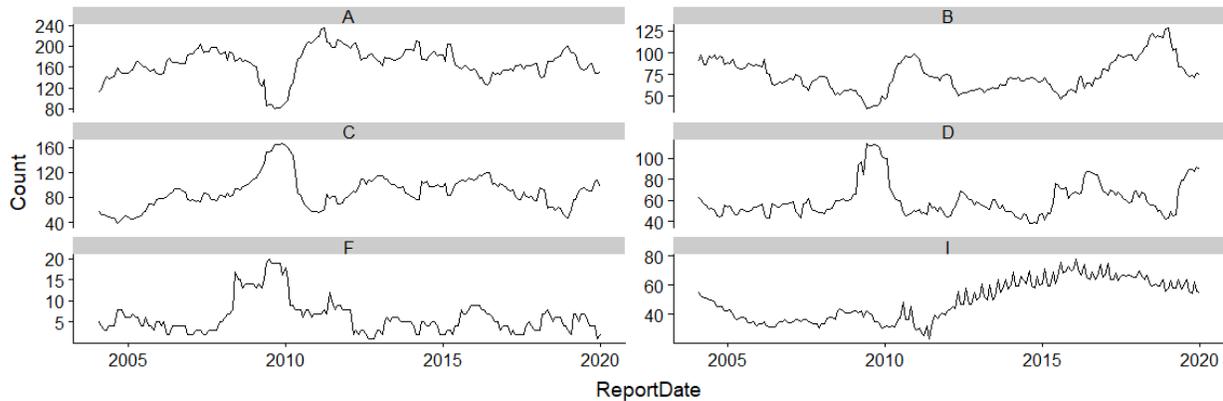


Assessing the quality and velocity of change in companies’ GAAP EPS plays an important role in our classification system’s persistent alpha-generation capabilities. Another field provided in the data feed is the EPS Growth Classification (EC) which classifies each company at points in time as **Growing, Flat, or Declining**.

Fundamental Grade

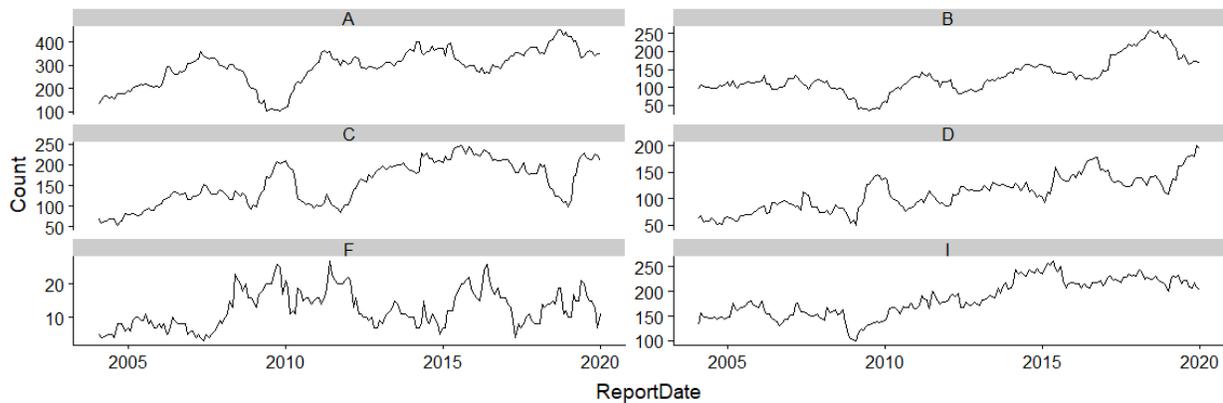
By combining the fundamental classification (FC) and the EPS Growth Classification (EC), the system assigns a Fundamental Grade (FG). Just like in school, grades range from “A” to “F”, with “A” being the best. A company can also be given an “I” for incomplete. It means that there was incomplete or insufficient financial history for the system to grade that company.

Number of S&P 500 Components per Fundamental Grade per Day



Number of Large Cap Stock per Fundamental Grade per Day

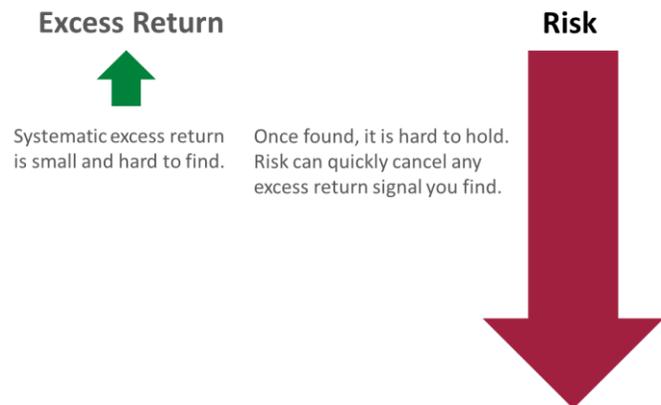
Large cap stocks are those with market capitalization greater than \$4B



Investment Philosophy

We believe that price sentiment analysis paired with deep fundamental analysis is a powerful tool for predicting future relative performance and risk. Behaviors are probabilistic, which means that given a stock's price sentiment and fundamental classification, we can measure the probability of outperforming or underperforming the relevant benchmark. Using those probabilities, portfolio managers can build strategies designed to deliver Alpha with lower Beta.

The stock market is a noisy and often volatile environment which results in substantial risk for trading strategies. Active trading strategies seek to find and capture excess return, however, systematic excess return is small and hard to find. Once found, it is hard to hold. Market risk can quickly cancel any excess return signal you may find. Our approach to creating Alpha generating strategies is to first manage the risk associated with a set of holdings. As will be described later in this document, our indicators and classifications demonstrate the ability to reduce risk. Once risk is reduced and manageable, only then can excess return be captured in a systematic way.



The goal of any active strategy is to outperform its passive benchmark. From a probabilistic point of view, the goal is to beat the passive benchmark more than 50% of the time. The expectation is that over short periods of time, a successful strategy should outperform only slightly better than 50% of the time, but as the time horizon increases, that probability should increase.

Over the short-run, our classifications have predictive power; this translate, over the long-run, to higher Alpha when implemented as a strategy. Later in this document, we present backtesting results for a selected subset of classifications.

Testable Hypothesis – Short-Term Performance

We expect stocks in each classification to have a statistically significant probability of outperforming or underperforming the relevant benchmark over the next X trading days.

In this document, we test this hypothesis using X=20. 20 trading days is approximately one calendar month. We have also tested where X=30 and X=40 and found similar results.

Backtesting – Long-Term Performance and Risk

Over any 3-year rolling period, strategies that hold companies in classifications with short-term predictive power, tend to outperform the market and other classifications. Backtesting results are shown in a later section.

Testing Short-Term Behavioral Characteristics

Using the Indicators described earlier, one can create different classifications to be used in building strategy holdings. In this section, we define some simple classifications and show how each classification behaves in terms of risk and performance.

Classifications

Fundamental Grade – a classification will be created for each grade.

Sentiment Quotient – the range of SQ (-1 to 1) will be subdivided and each resulting subset will become a classification.

Behavioral Characteristics

The behavior of each classification is analyzed to understand risk and performance. While backtesting can be a useful tool for understanding long-term risk and performance, it can obfuscate the underlying behavioral characteristics of an indicator or classification. Backtesting reporting primarily relies on performance metrics that generally have significance when applied to return series of at least 36 months. In the next section, we will cover our backtesting methodology. However, in this section, we apply a more straightforward statistical analysis in order to demonstrate the predictive nature and power of our classifications over the short-term. Once empowered with this information, you can then use it to build your own strategies.

Experimental Objective

To assess the predictive power of our classifications, our classifications need to make predictions and those predictions need to be compared against outcomes.

Prediction - “Company X will outperform its peers over the next 20-trading-days”

Outcome - “Company X outperformed or underperformed over those 20-trading-days”

For this experiment, we will measure “win rate”, which will be defined as the rate at which predictions are correct (win). It is the rate at which a classification (treatment group) outperforms its peers (control group) over a 20-trading-day period. Outperforming will be defined as delivering a higher return than 50% of the peer group at the end of a given 20-trading-day period. If a company is in the top 50% of peers, it is considered a “win” (correct prediction) or “loss” (incorrect prediction). By aggregating all “wins” and “losses” for a classification (treatment group) we can measure the “win rate” for that classification. This can then be used to estimate the probability of outperformance and the statistical significance of that estimate. A probability that deviates significantly from 50% is considered to have “information”. The industry standard measurement of information is the information coefficient (IC) which is simply the win rate converted to a scale from -1 to +1. In our experiment, a positive IC is associated with a win rate greater than 50% and indicates predictive power in selecting outperforming stocks. Conversely, a negative IC indicates a predictive power in selecting underperforming stocks.

Methodology for Analysis

To understand performance, we need to be able to analyze companies on a per-day basis. Because the stock market fluctuates with a great deal of noise on a daily basis, we cannot use the magnitude of daily returns as a basis for comparison. We need to scale daily returns to relatively stable metrics, regardless of potentially dramatic daily swings.

Percentile Rank (Peer Ranking)

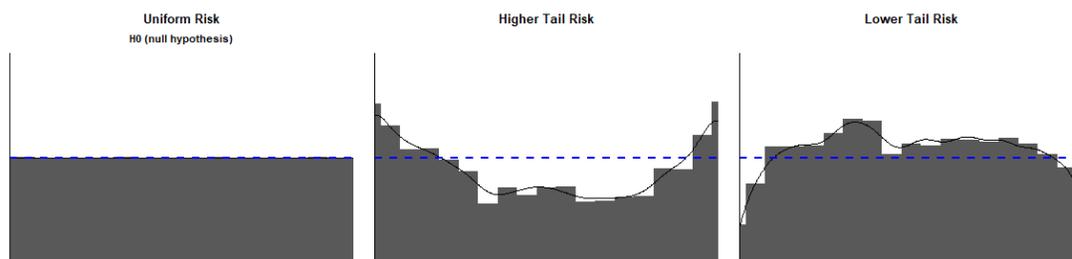
We rank all companies in the S&P 500 on a daily basis by their leading (future) 20-trading-day return. This provides us with a basis for peer ranking. Peer ranking is calculated every day. Those ranks are converted to percentiles. The percentile rank becomes our unit of analysis for assessing risk and performance. The scale is from 0 to 1 with a median of 0.5. Given this design, there is a 50% chance of underperforming and a 50% chance of outperforming peers. Given this design, all things being equal, a randomly selected company for a randomly selected day theoretically has a 50% probability of outperforming its peers over the next 20 trading days (assuming equal weighting).

Due to the ranking nature of this metric, the distribution will naturally be uniform. Our Null Hypothesis (H_0) is that companies will follow this uniform distribution with 50% probability of outperforming. As we test an Alternative Hypothesis (H_a) we will test for significant deviation from 50% probability of outperforming. A statistically significant deviation above 50% would indicate that the test group is outperforming its peers (the other components of the S&P 500).

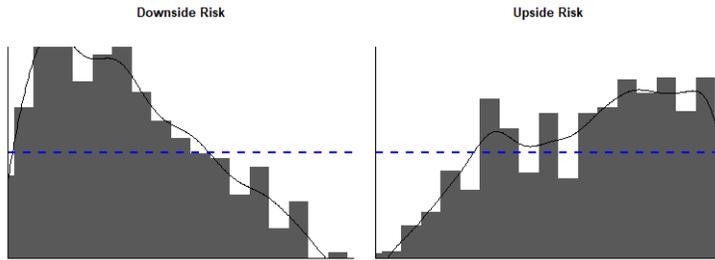
Assessing Risk

Since, by design, the null hypothesis is a uniform distribution, there is an equal probability that a company will be ranked anywhere from 0 to 1. This means there is an equal “risk” that a company will underperform (downside risk) or over-perform (upside risk).

A non-uniform distribution implies a change in the risk profile. The histograms below demonstrate the null hypothesis’ uniform distribution as well as other example hypotheses that deviate from uniform distribution. The middle histogram is U-shaped with higher density in the tails. That classification of companies therefore has a higher downside and upside risk at the tails of the distribution, which we will call “higher tail risk.” The right histogram exhibits an “inverted U” shape which means lower density in the tails. That classification of companies has “lower tail risk”.



Distributions can also be skewed to one side or the other, which would indicate a higher probability of either downside or upside risk.



Assessing Performance

Company Day

To assess performance of our classifications, we test the probability of outperforming and its statistical significance of deviation from the 50%. A significant deviation above 50% indicates that the test group is outperforming its peers (the other components of the S&P 500). To estimate the probability of outperforming and its significance, we primarily use the Binomial Test (`binom.test`). A $p\text{-value} < 0.05$ suggests that population outperformance probability is NOT 50%. This test is applied to the total sample (of a classification) and a sample by year. This metric is used to answer the question, **“If I know today’s classification of a company, what is the risk and performance probability distribution for the next 20 trading days?”** This unit of analysis is a **Company Day**, meaning that each data point (observation) is a company on a particular day.

Results for “Company Days”

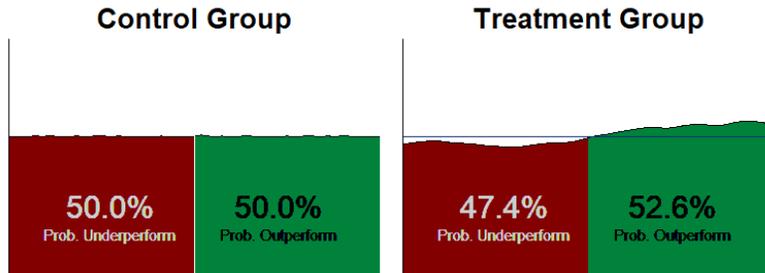
This section provides the statistical tests used to answer the question, **“If I know today’s classification of a company, what is the risk and performance probability distribution for the next 20 trading days?”** The unit of analysis is a **Company Day**, meaning that each data point (observation) is a company on a particular day.

With this unit of analysis, we can test the probability of outperforming and the associated risk. Because returns are converted to percent ranking, the distribution is uniform and the probability of outperforming is 50%. Any deviation from uniformity would imply higher or lower tail risk, while deviation 50% probability would imply underperformance or over-performance relative to peers (components of the S&P 500). More details are provided in the prior section.

Interpreting the Charts

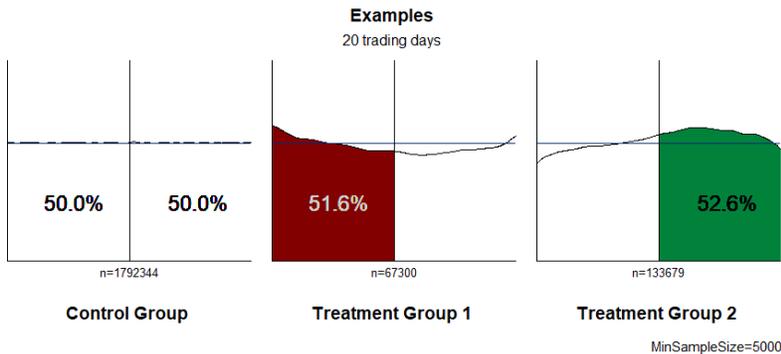
To assess risk and performance, a few different types of charts are used. Those charts are described here and guidance is provided on how to interpret them.

Performance Distribution 1



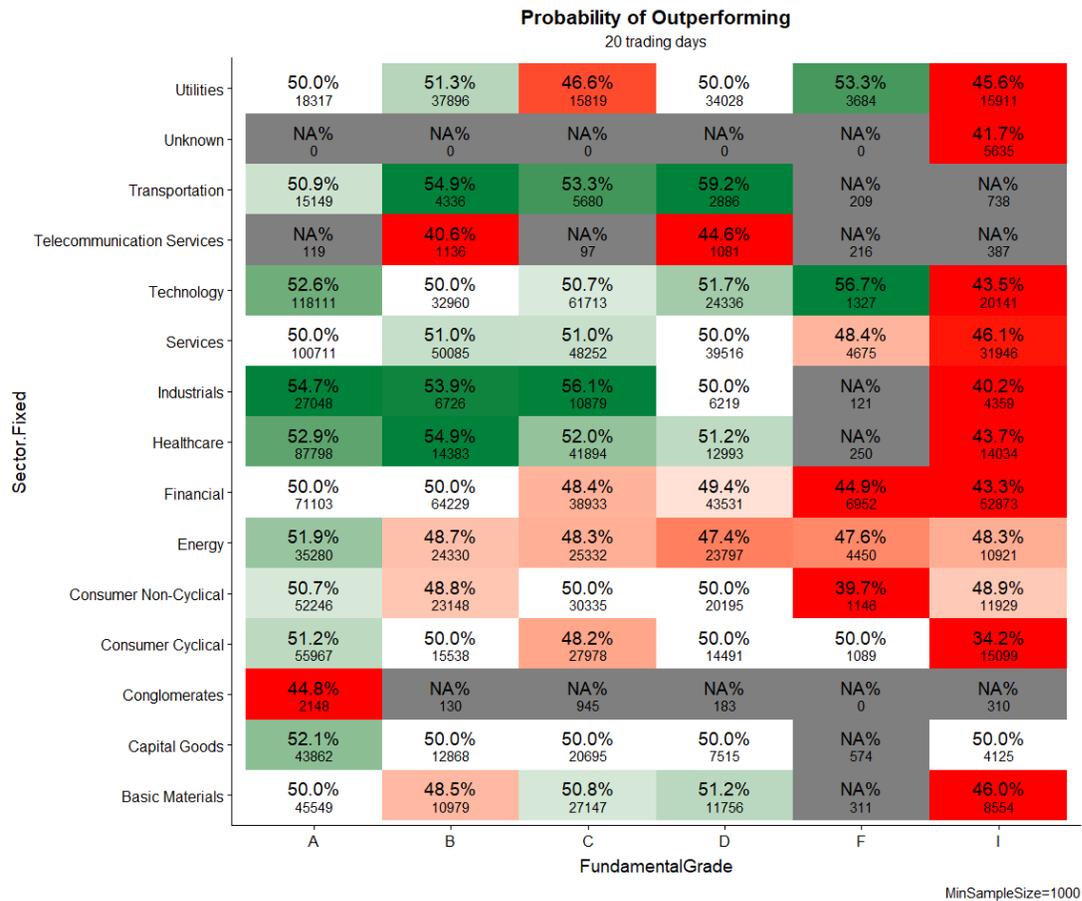
The “Performance Distribution 1” chart shows the distribution of ranked performance. The green portion of the chart shows the probability of outperformance, while the red portion shows the probability of underperformance. By design, the Control Group (Market) is a uniform distribution with a 50% probability of outperforming. When rendering this chart, the binom.test is applied. If the probability of outperforming is significantly different from 50% then the probability estimate is rendered, else 50% is rendered. For example, you would not see a probability of outperforming of 50.01% because it is not statistically significant. So, in that case, it is always rendered as 50%. In summary, **if you see a number other than 50% you can be assured that it is a statistically significant number.**

Performance Distribution 2



The “Performance Distribution 2” chart is an easier to interpret version of “Performance Distribution 1”. The difference is that the area under the curve is only colored when there is a statistically significant deviation from 50%. Because of this, only one side of the chart is colored at a time. This makes it easier to quickly identify scenarios that outperform or underperform. If there is no significant deviation from 50%, the chart will show 50% on both sides as is shown in the Control Group example. Treatment Group 1 shows when there is significant underperformance, while Treatment Group 2 shows outperformance.

Performance Heat Map



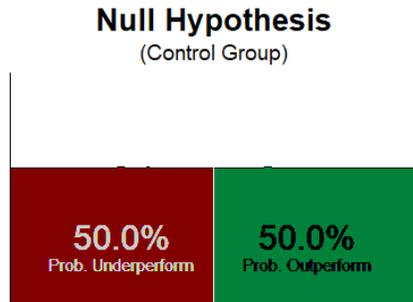
The “Performance Heat Map” chart above is an even simpler rendering that only shows the probability of outperforming. If the number is significantly above 50%, then it is colored green, with darker colors indicating larger magnitude. If the number is significantly below 50%, then it is colored red, with darker colors indicating smaller magnitudes. NOTE: the outperformance value is shown. i.e. it is the number that would normally correspond to the right side of the performance distribution charts described above. If the number does NOT significantly deviate from 50% (e.g. 50.01%), then it is simply rendered as 50% with a white background. Underneath the outperformance percentage is the number of records for that classification. If the number of records is considered too small, then NA is rendered indicating that there was insufficient information to estimate the probability. This threshold is called “MinSampleSize” and is primarily a judgement call for the analyst. The configured value for MinSampleSize is documented in the lower right corner of each heat map.

The main advantage to this chart is that many classifications can be rendered in a dense format. Obviously, the major disadvantage is that risk information is lost. As you will see in our analysis below, we will use the heat map to identify opportunities for outperformance among many possible options. Once we have identified some candidates, we will explore them in more detail using the performance distribution to understand the risk profile.

Null Hypothesis

Here are the results for the Null Hypothesis (Control Group), to confirm that it is indeed uniform with 50% probability of outperforming.

Total



By Year

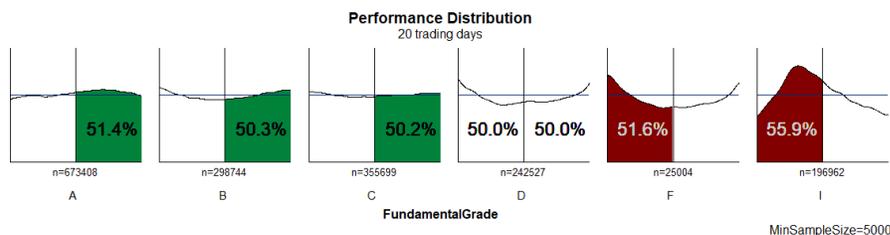
Tests by year also confirm that the Null Hypothesis is consistently uniform with 50% probability of outperforming. (Charts not included)

Fundamental Grade

Total

Fundamental Grade performance varies based on the grade. In general, “A” is the best and “F” is the worst. “Best” and “worst” are determined by inspecting probability of outperformance as well the distribution of performance, which is a measure of risk. It is clear that “A” has the highest probability and lowest downside risk combination.

A company can also be given an “I” for incomplete. It means that there was incomplete or insufficient financial history for the system to grade that company. “I” companies have a bias toward underperformance.



By Year

The probability of outperformance for Fundamental Grade “A” tends to hold across most years. As market conditions vary, outperformance may also vary over the short-run.

Probability of Outperformance

20 trading days

ReportDateYear	A	B	C	D	F	I
2019	53.3% 41916	51.9% 22114	51.8% 21870	51.4% 18122	50.0% 1172	33.0% 14937
2018	51.2% 42579	51.2% 28696	50.0% 17028	50.0% 14001	54.2% 1332	43.5% 15441
2017	57.2% 40189	51.5% 22974	51.3% 20492	47.1% 16073	NA% 745	31.4% 16521
2016	50.9% 35857	49.1% 17229	53.1% 27129	53.8% 19278	54.1% 1609	39.0% 16951
2015	50.0% 42687	51.9% 14426	48.8% 25302	47.2% 15732	39.6% 1337	52.6% 16754
2014	50.9% 47142	52.5% 17504	50.7% 23514	50.0% 10875	NA% 983	42.4% 15511
2013	53.8% 45057	50.0% 15373	50.0% 24891	51.0% 13818	50.0% 1278	35.3% 14059
2012	51.4% 46212	50.0% 14452	50.0% 25348	50.0% 14663	NA% 719	44.4% 12118
2011	52.3% 53309	47.8% 19136	49.1% 17950	47.2% 12697	43.4% 2119	47.3% 8368
2010	50.5% 42982	50.0% 20755	49.2% 24317	50.0% 15258	50.0% 2264	50.0% 8698
2009	48.6% 26752	46.7% 11094	50.0% 36750	51.8% 25067	51.8% 4255	50.0% 9427
2008	50.0% 44329	51.8% 14981	49.0% 23457	50.0% 14323	41.3% 2929	49.0% 9856
2007	51.3% 48303	51.9% 16893	50.0% 19650	43.9% 13128	NA% 737	48.5% 8475
2006	50.0% 41699	48.2% 18213	50.0% 21327	52.2% 13330	NA% 919	50.0% 8427
2005	49.0% 39442	50.9% 21573	46.7% 14500	52.6% 12784	50.0% 1278	53.2% 9239
2004	49.3% 34953	48.3% 23331	52.7% 12174	49.1% 13378	50.0% 1328	53.4% 12180

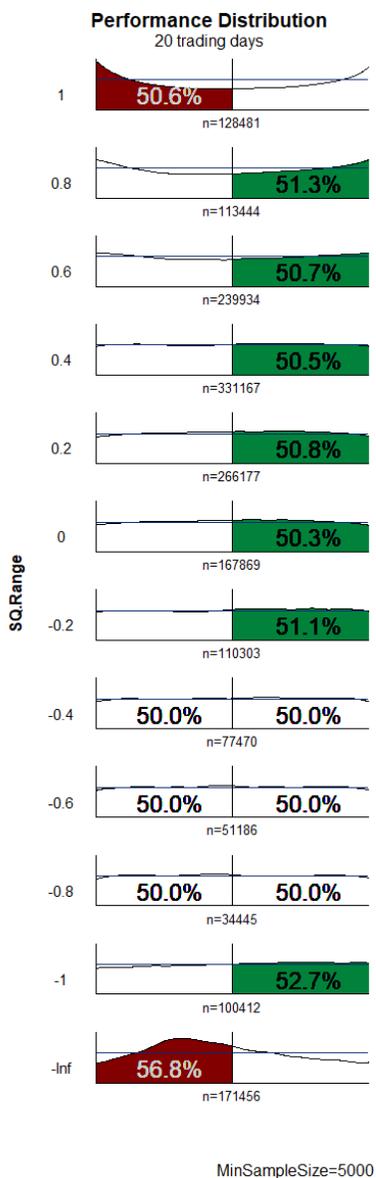
MinSampleSize=1000

Sentiment Quotient

Total

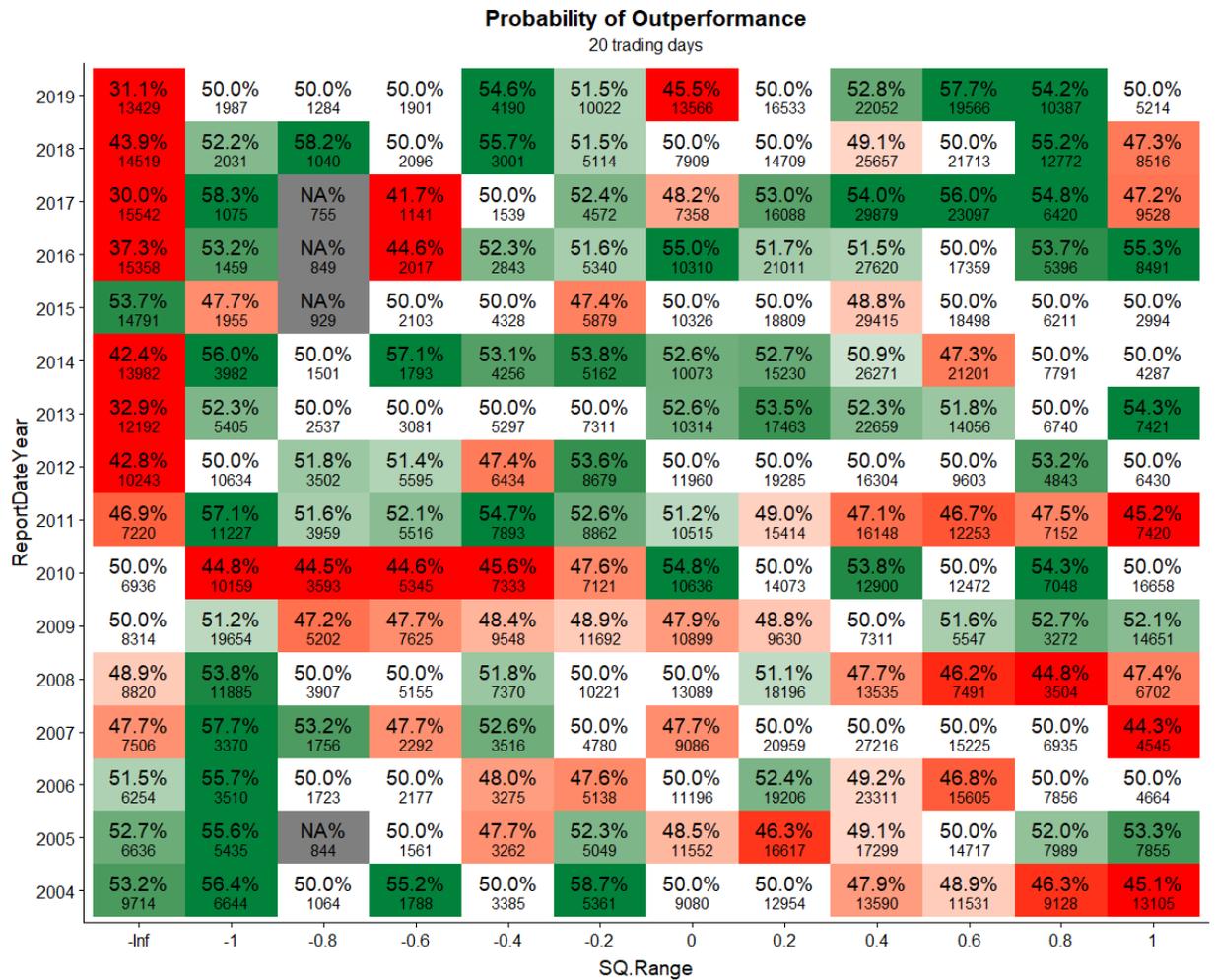
Sentiment Quotient performance varies based on the value. Unlike Fundamental Grade, there is not a “best” to “worst” scale. SQ quantifies the portion of a stock’s current price that is attributable to future operating performance (future earnings). There are certain “goldilocks” regions. Based on risk tolerance different strategy designers may choose different regions to achieve their investment objectives.

A stock can also have a blank SQ, indicating that there was incomplete or insufficient financial or pricing history for the system to assign a sentiment quotient. These blank SQ stocks, have a bias toward underperformance. On the chart, blank SQ stocks are represented as “-inf” in order to maintain the numerical units of the y-axis.



By Year

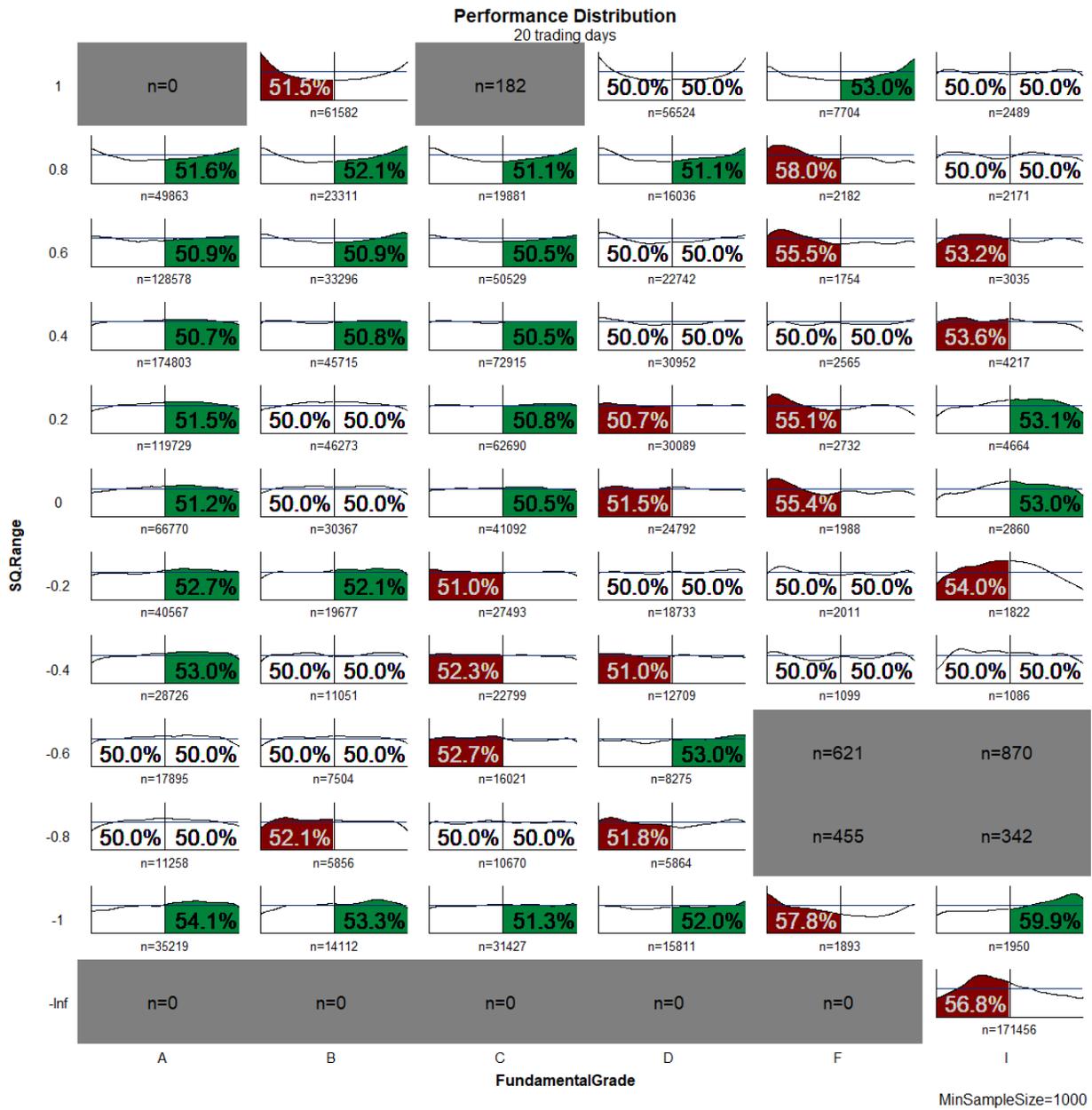
The probability of outperformance tends to hold across most years. As market conditions vary, outperformance may also vary over the short-run.



MinSampleSize=1000

Composite Classifications

By combining Fundamental Grade and Sentiment Quotient, one can develop powerful stock screeners for custom strategies. Each strategy designer may choose different regions based on risk tolerance and investment objectives.



Using Behaviors to Build Strategies

The analysis of behavioral characteristics described in the previous section demonstrates the predictive power for certain classifications. While some of the statistically significant probabilities may seem small in absolute terms, those small differences add up quickly over time when implemented as a strategy. The methodology used above removes magnitude of underperformance or overperformance in terms of return and therefore does not reflect the magnitude of a potential strategy's performance. The methodology only analyzes peer ranking on the basis of leading (future) 20-trading-day returns, which does not tell us much about long-term buy and hold strategies, nor does it tell us much about strategies that adapt to ever-changing market conditions and company classifications

In this section, an example strategy is defined and backtested for a single classification, "Screener.2."

Performance of Example Strategy

The classifications described earlier are extremely powerful tools for building alpha-generating strategies. Let's choose a simple, straightforward, and data-driven set of classifications for demonstration and educational purposes. Because this is for demonstration and educational purposes only, it may not meet all your organization's real-world requirements for a strategy. Valspresso is happy to discuss your organization's requirements and how the indicators might work in one of your strategies.

Criteria

For this example strategy, we define a strategy that is simply a stock screener that exclusively uses Valspresso classifications. This strategy is called **Screener.2**.

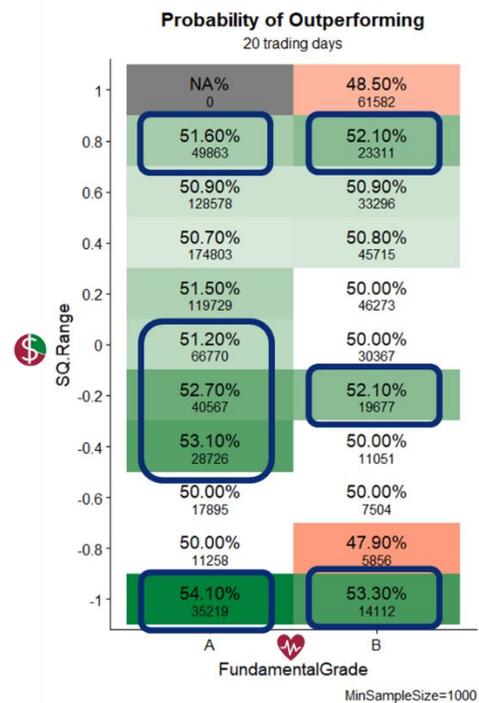
Only hold stocks where all of the following are true:

1. Fundamental Grade must be A or B
2. SQ value must be in a region with a good outperformance probability and reduced risk

If the company no longer adheres to the above classification, sell it.

By looking at the composite classifications we choose the classifications designated by the blue squares.

So, in summary, the strategy holds stocks that are fundamentally healthy, and its SQ (measure of amount of belief) is predictive of outperformance. When a stock no longer adheres to these criteria it is sold.



Backtesting

By applying only those rules to the components of the S&P 500 from 2004 to 2019, backtested results demonstrate the power of Valspresso's sentiment classifications.

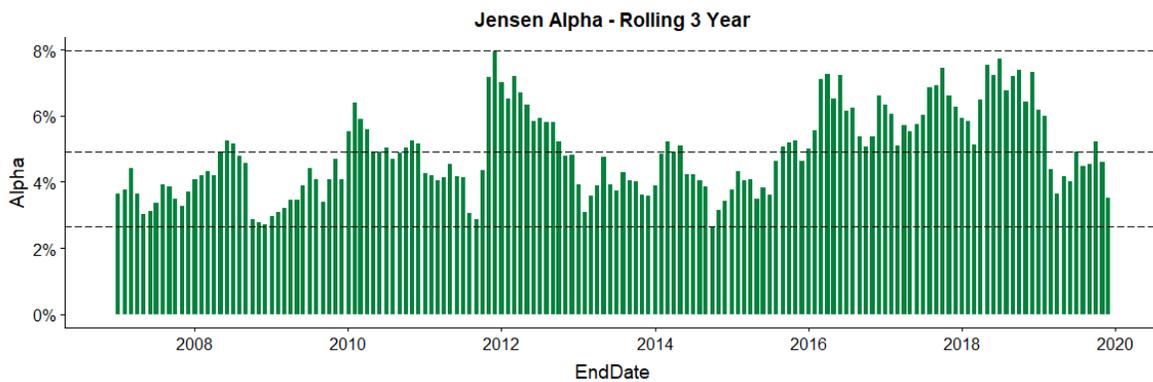
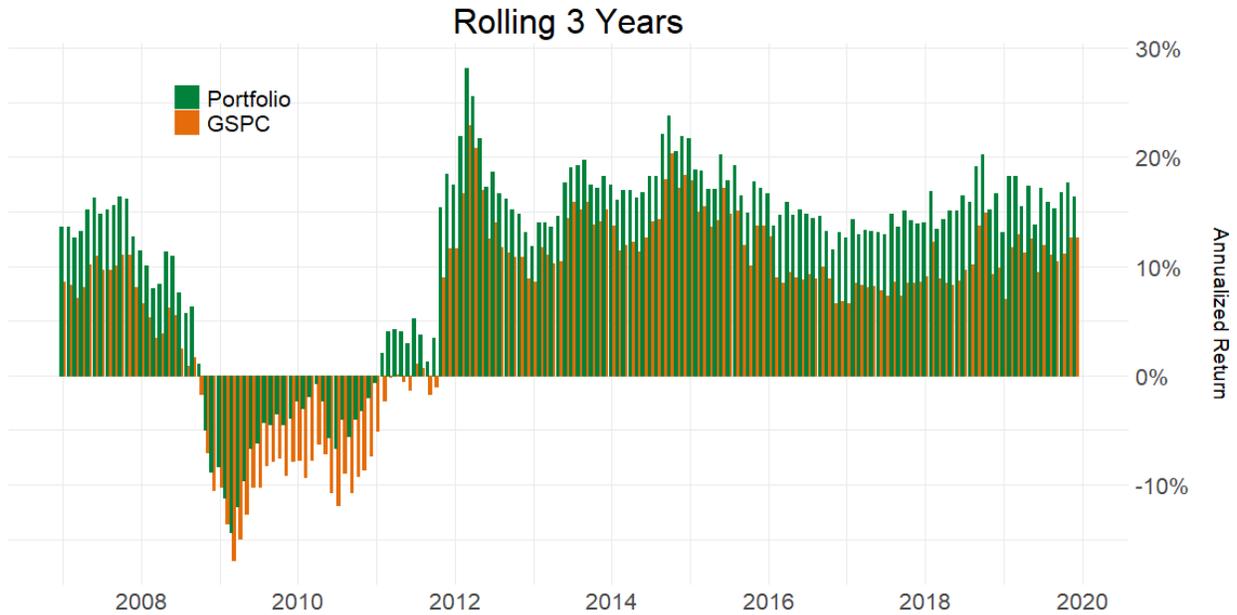
Methodology

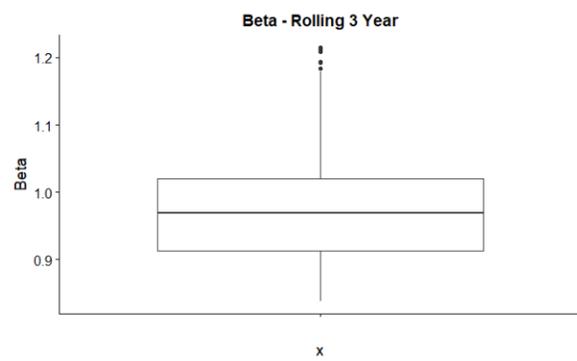
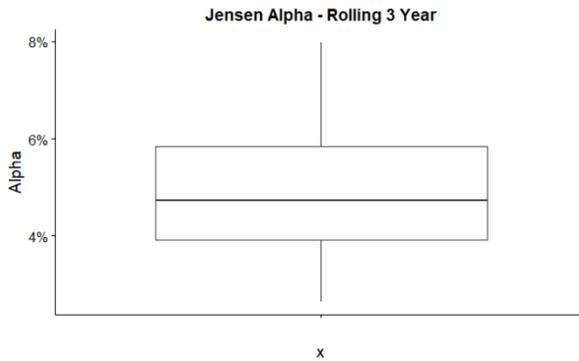
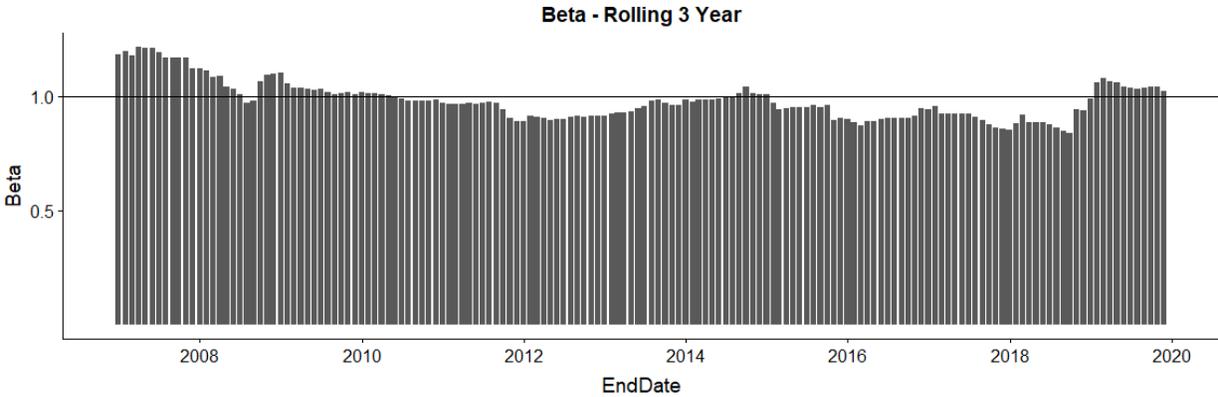
To demonstrate the contribution of Valspresso's classifications, we created a simplified backtesting system. This system was intentionally kept very simple to more easily test and compare the performance against publicly-available benchmarks without concern for confounding factors in the backtesting methodology.

Data consists of Valspresso's Sentiment data for S&P 500 companies from 2004 to 2019 and daily closing prices for each component of the S&P 500. All companies matching the holding rule (described above) on a given day will be held in the simulated portfolio. Holdings are equal-weighted. Portfolio is 100% stock. No money is simulated. (i.e. no account value, no trading costs, no management fees, etc.) Based upon those daily holdings, the backtesting system simply uses daily returns and percent allocations to calculate daily weighted returns for the portfolio. Those daily returns are then used to build all performance reports. The primary performance measure is 3-year rolling Alpha.

Performance vs. S&P 500

Since the backtested holdings are components of the S&P 500, let's review the performance using the S&P 500 price index as a benchmark. The signals for the Screener.2 strategy for rolling 3-year periods from 2004 to 2019 delivered an average **annualized Alpha of 4.9% with a range of 2.7% to 8.0%**. This result is delivered with an average beta less than 1 and information ratio greater than 1.





The signals for our Screener.2 strategy are available for subscribers to independently reproduce the results and perform their own analyses.

Other Use Cases

The above use case is by no means the only possible use for Valspresso's indicators. Valspresso has created several of its own successful strategies for its clients. While we can't go into the details of those strategies, we can give a few ideas to help you get started on one of your own strategies.

Increase number of holdings by looking at recent classifications

The strategy described above buys and sells on a daily basis based on the day's classifications. On a single day, the number of holdings may be too low for your needs. Since our research shows that our classification's outperformance persists for a period of time beyond the 20 trading days documented here, you could choose to include stocks that matched your criteria anytime within the past X number of days.

Decrease turnover by rebalancing less frequently

The strategy described above buys and sells on a daily basis. Classifications can be volatile on a daily basis and can increase your turnover. If you require lower turnover, you could simply rebalance on a monthly (or longer) basis. Our research shows that our classifications'

outperformance persists for a period of time beyond the 20 trading days documented here. This means that even if you buy and hold stocks for a month, you would still experience outperformance.

Identify your own SI, SQ thresholds

We have proposed thresholds for SQ, but with some experimentation for the types of stocks you typically hold, you could develop your own thresholds.

Short stocks

Instead of choosing classifications that outperform, you could choose classifications that underperform and develop a shorting strategy.

Dynamically move between classifications

Based upon market conditions, select holdings based on the classifications that are predicted to do best during those market conditions.

Data Processing

Sources of Data

- SEC filings (10-K, 10-Q, 8-K) from Edgar Online
- Pricing data from Edgar Online
- 3-month treasury bill rates from the Federal Reserve

Coverage

Valspresso covers all publicly-traded U.S. companies. Fundamental classification requires at least 3 years of financial statements. Sentiment classification requires at least 3 years of trading history.

Frequency

Automated analysis is performed daily for all companies. Financials filed with SEC go through a manual quality review process by our SEC data provider, therefore, there can sometimes be a delay between the time a company files with the SEC and it shows up in our database. The data provider prioritizes based upon market cap and other investor relevance factors. For this reason, filings for large market companies usually appear the next day, but small-cap companies can sometimes take a few days. When developing a strategy and/or backtesting a strategy, this latency should be taken into account.

Point-in-Time Data

Our database, derived indicators, and backtesting system use point-in-time data. All data is aligned to ensure that the indicator calculation and backtesting system only consider data that was available on that particular trading day.

Dictionary

Price Date

The date for which the given security was traded on the stock market. The price date only includes days for which the stock market was open. Not all stocks are traded every day and not all markets share the same holidays.

Holding Date

The date that a given security is held in a portfolio. This portfolio might be a backtested simulation or a live portfolio. Valspresso defines different types of dates and enforces their application through its software to ensure there is no look-ahead bias, such as holding a security based on pricing information that is not yet available.

Price per Share (PPS)

The price of a security traded on the stock market. PPS represents a generic definition of stock price and is used as theoretical price used in formulas throughout this document. When those formulas are calculated, a specific price (e.g. prior day's closing price) is used.

Number of Outstanding Shares

The number of outstanding shares of stock used when PPS is reported when trading on a stock market. A company issues a number of shares to be traded on the stock market. Those number of shares can change over time because of splits, buybacks, and other corporate actions. This value is the implied denominator when reporting things such as PPS or EPS. The precise way that this implied denominator affects values such as PPS and EPS is a complex subject and outside the scope of this document.

Closing Price

The PPS of a stock reported by a stock exchange at the end of the trading day. The closing price is not adjusted for stock splits or other corporate actions. The closing price is the most commonly used PPS used to calculate indicators such as Sentiment Index.

Adjusted Closing Price

The Closing Price historical adjusted for stock splits and other corporate actions. Each time there is a new split, a stock's historical adjusted closing prices are recalculated. Adjusted closing price is not used to calculate sentiment index or its derivatives. Adjusted closing price is only used for assessing predictive capabilities of our indicators by calculating lagging or leading returns.

Ticker

A unique set of characters that represent a security traded on a stock exchange. Tickers can be 1 to 5 characters. Tickers are unique for any given day. In the event a security is delisted by an exchange, its ticker can be issued for a completely different security in the future. Therefore, a ticker can be assigned to different securities over time.

CompanyId

An integer that represents a company. This number will not change even if the ticker changes.

Sentiment Index (SI)

SI is used to project expected earnings growth. For example, an SI value of 4 means that the company's stock price reflects an expectation that earnings should grow by a factor of 4.

Sentiment Quotient (SQ)

SQ is used to quantify the portion of a company's stock price that is attributable to future earnings growth. For example, an SQ of 0.8 implies that 80% of the company's stock price is attributable to expectation of future earnings growth and 20% to current fundamentals.

Fundamentals Classification (FC)

Possible Values: High, Medium, Low, Unknown

Valspresso's Automated Stock Analyst performs deep financial analysis of publicly-traded U.S. companies. This context is important to manage your strategy's risk. Let's say that one of your strategy's holdings is a stock with a high expected earnings growth. If that company's financial quality is poor, it is less likely to achieve or sustain that growth and therefore should be considered riskier. In that case, additional risk mitigation measures would need to be employed to achieve investment objectives.

Our system automatically assesses profitability, solvency, and management by inspecting each public company's financial statements and press releases filed with the Securities and Exchange Commission (SEC) and classifies the quality of companies' financials as High, Medium, or Low. If the system is unable to make a determination, the financials are classified as Unknown.

Sentiment Classification (SC)

Possible Values: Negative EPS, Pessimistic, Optimistic, Exuberant, Unknown

One of the unique attributes of Sentiment Index (SI) is that its value is on a scale that pivots around the fixed value of 0. This allows the automated analyst to objectively measure the amount of sentiment in a company's stock price and classify each company as Negative EPS, Pessimistic, Optimistic, or Exuberant. If the system is unable to make a determination, the company is classified as Unknown.

EPS Growth Classification (EC)

Possible Values: Growing, Flat, Declining, Unknown

The EPS Growth Classification (EC) field classifies each company at points in time as Growing, Flat, or Declining. If the system is unable to make a determination, the company is classified as Unknown.

Fundamental Grade (FG)

Possible Values: A, B, C, D, F, I

By combining the fundamental classification (FC) and the EPS Growth Classification (EC), the system assigns a Fundamental Grade (FG). Just like in school, grades range from "A" to "F", with "A" being the best. A company can also be given an "I" for incomplete. It means that there was incomplete or insufficient financial history for the system to grade that company.

Official Annual EPS

The Earnings Per Share (EPS) from a company's latest official (10-K) annual filing.

Official Quarterly EPS

The Earnings Per Share (EPS) from a company's latest official (10-Q) quarterly filing.

Preliminary Annual EPS

The Earnings Per Share (EPS) from a company's latest official (10-K) annual filing or latest preliminary annual (8-K) financial filing. In those calculations where "preliminary" values are allowed, the latest annual value is used regardless of whether it's official or preliminary.

Preliminary Quarterly EPS

The Earnings Per Share (EPS) from a company's latest official (10-Q) quarterly filing or latest preliminary quarterly (8-K) financial filing. In those calculations where "preliminary" values are allowed, the latest quarterly value is used regardless of whether it's official or preliminary.

Beta

36-month beta calculated on a daily basis. The benchmark used to calculate Beta is S&P 500 price return. Beta is calculated over the prior 36 months relative to the current price date being calculated. Therefore, beta changes daily for each stock.

Market Rate of Return

Market rate of return is calculated using the S&P 500 price return.

Risk-Free Premium

Three Month U.S. Treasury Bill Rate

Required Rate of Return

Required rate of return is calculated using beta, risk-free premium, and market rate of return.

GAAP

Generally accepted accounting principles (GAAP) refer to a common set of accepted accounting principles, standards, and procedures that companies and their accountants must follow when they prepare and present their financial statements. The Financial Accounting Standards Board (FASB) has the authority to establish and interpret GAAP in the United States.

Going Concern

Under generally accepted accounting principles (GAAP), the continuation of a reporting entity as a going concern is presumed as the basis for preparing financial statements unless and until the entity's liquidation becomes imminent. Preparation of financial statements under this presumption is referred to as the "going concern" basis of accounting.

Sentiment Indicators Data Feed

Product Overview

Valspresso's unique quantitative metrics are a result of 20 years of theoretical research that were further refined over 10 years of practical application. Valspresso's technology analyzes all publicly-traded companies on the U.S. major exchanges every day. This daily automated analysis includes a patented price sentiment analysis as well as deep fundamental analysis. Subscribers to Valspresso data feeds use them to develop their own alpha-generating trading strategies. The data feeds provided have been used internally for years at Valspresso to build trading strategies for its clients.

Version History

Version	FactSet Release Date	Notes
1	Sep 1, 2019	Original version used for FactSet evaluation
2	Pending	First version released to FactSet production. Changed SI and SQ thresholds for SC classifications. Added NegativeEPS and Exuberant SC classifications. Changed EC classification to be independent of SC. Renamed EC classifications. Minor changes to FC to reduce turnover. Added FG field.

File Structure

Format

The data feed is delivered as a comma-delimited file (CSV). Each column is separated by a comma. Each field's value is enclosed in double-quotes if it is a string (character). Each row represents a single ticker for a single trading day. The first (header) row contains the names of the columns and each column name is enclosed in double-quotes.

Example from January 2, 2004:

```
"VersionId","Date","Ticker","Exchange","CompanyId","Country","SI","SQ","FC","FG","SC","EC"
2,2004-01-02,"A","XNYS",8244,"US",-2.3068,1.0000,"Medium","D","NegativeEPS","Declining"
2,2004-01-02,"AA","XNYS",10803,"US",24.6598,0.9610,"High","A","Exuberant","Growing"
2,2004-01-02,"AAA","XNYS",379498,"US",,,,"Unknown","I","Unknown","Unknown"
2,2004-01-02,"AAC","XNAS",839429,"US",,,,"Unknown","I","Unknown","Unknown"
2,2004-01-02,"AAI","XNYS",4199,"US",8.7995,0.8980,"High","A","Exuberant","Growing"
2,2004-01-02,"AAME","XNAS",11445,"US",,,,"Medium","I","Unknown","Unknown"
2,2004-01-02,"AANB","XNAS",7319,"US",-0.4328,-0.7630,"Medium","D","Pessimistic","Declining"
2,2004-01-02,"AAON","XNAS",938,"US",-0.7779,-1.0000,"High","C","Pessimistic","Declining"
2,2004-01-02,"AAP","XNYS",90252,"US",,,,"High","I","Unknown","Unknown"
2,2004-01-02,"AAPL","XNAS",2035,"US",0.9417,0.4850,"Medium","D","Optimistic","Declining"
```

Values which are NULL, N/A, or unavailable will be represented as blank (i.e. open double quote immediately followed by closing double quote with nothing in between).

Columns

Column Name	Type	Format	Definition (see dictionary)
VersionId	Integer		Identifies the version of the classifications to allow incremental improvements without breaking historical backtesting.
Date	Date	yyyy-MM-dd	See Price Date
Ticker	String	1 to 5 characters	See Ticker
Exchange	String	4 characters	ISO 10383 - MIC Code
CompanyId	String	integer	See CompanyId
Country	String	2 characters	ISO 3166-1 – 2 character
SI	Decimal	4 decimal places	See Sentiment Index
SQ	Decimal	4 decimal places	See Sentiment Quotient
FC	String		See Fundamentals Classification
SC	String		See Sentiment Classification
EC	String		See EPS Growth Classification
FG	String		See Fundamental Grade

Coverage and Frequency

Valspresso covers all publicly-traded U.S. companies. Automated analysis is performed daily. Data feeds are delivered before the opening bell of U.S. markets (9:30 am ET).

Valspresso, Inc.
 11921 Freedom Drive, Suite 550
 Reston, VA 20190
 rnosegbe@valspresso.com
 (703) 675-7722