This white paper explains the meaning of and methodology behind Prattle's Equities Analytics scores:

1. **What the scores mean.** The first section of this paper discusses how to understand our scores in the context of both financial information and linguistic patterns.
2. **The methodology behind the scores.** The second section of this paper explains how we build and update our lexicons utilizing machine learning.

What the Scores Mean: Patterns in Language

People speak in patterns. These patterns aren't random—if they were, communication would be unintelligible. Rather, linguistic patterns link specifically to the conscious and subconscious thoughts of the communicator.

Executives and political leaders often explicitly communicate the state of their company or country through linguistic patterns. For example, when former President George W. Bush said "This sucker could go down" in his 2008 speech at the brink of the financial crisis, his thoughts about the state of the U.S. economy were beyond doubt. When an executive speaks similarly about their company, it's time to sell.

But much of the valuable information contained in linguistic patterns is far more subtle. Executives at certain publicly traded casinos, for example, emphasize licensure and regulation when they are performing well. When performance is poor, however, their emphasis shifts to slots offerings, costs, and casino closings. Historically, only experienced analysts who have a history and rapport with the company's executives could detect these nuanced patterns. Now, technology has made it possible for algorithms to detect what once only above-average human analysis could.

The subtlest of patterns, however, are very difficult (or perhaps impossible) for human analysts to consciously assign meaning to. These patterns manifest in the language of a corporate executive like a poker player's tell. For instance, executives at a small-cap debt collection agency tend to shift focus to future product innovations when the company's performance slumps. Subtle shifts like these are quite challenging for human analysts to consciously notice...let alone concretely analyze and link to stock price.

Evan Schnidman / 617.304.6807 / evan@prattle.co

All of these linguistic patterns, from the obvious to the subtle, impact their audience. When it comes to finance, the relevant dimension of that impact is price. In other words, financial professionals are particularly concerned with how the language a company's executives use impacts the price of that company's stock.

What the Scores Mean: Algorithmically Mapping Language

To understand this relationship between linguistic patterns and stock price, Prattle's system algorithmically maps the language patterns corporate spokespeople and analysts use when describing a specific company to the historical performance of that company. This process produces a lexicon of scored expressions (words, phrases, sentences, etc.) valued in terms of their impact on stock price.[1] Then, Prattle's system uses this lexicon as a basis of evaluation for subsequent language by those executives and analysts.[2] This fully algorithmic analysis produces a score that is indicative of future company performance.

Prattle's scores capture price-relevant factors that are not captured by a company's fundamentals. By controlling for common fundamental quantitative factors like peer-company performance, Prattle scores represent the price movement that can be directly tied to the sentiment expressed in corporate communications.[3] While the scores will always be influenced by a company's financial trends, they are valuable because they capture and describe the price impact of previously unquantifiable information:

1. **Blatant price-relevant information.** For example, did the company get that drug approved or not, and how will that affect stock price?
2. **Subtle price-relevant information.** For example, is the CEO speaking in patterns historically consistent with poor performance, and how will that affect stock price?

As a result, the scores can contain information that is indicative of a company's future performance and often *uncorrelated* with other traditional factors used in forecasting performance. And, because Prattle builds each company's lexicon using the history of each company's language and stock price, each score represents a context-specific evaluation of the language in that specific company's communications.[4]

In short, the Prattle process produces a score that indicates a company's future performance, captures price-relevant information not included in a company's fundamentals, and is tailored to the specific relationship between a company's language and its stock price.

---

[1] Prattle's lexicons are databases of words, phrases, sentences, and sometimes whole paragraphs. Each word, phrase, sentence, etc. in these databases has been assigned a value based on its historical impact on stock price. Prattle has created a custom lexicon in this way for every company it covers, and that lexicon continually updates through machine learning.
[2] We will talk in more detail about Prattle's methodology later in this paper.
[3] If you'd like to learn more about how Prattle isolates for the impact of language on price, you can always schedule an appointment with our quant team. Reach out at info@prattle.co.
[4] Prattle's dataset goes back to 1999; it takes 16 quarters to fully train the algorithm on a company's language.

What the Scores Mean: Applications

A Prattle score of an earnings call captures how the change in a given company's tone will likely affect short term stock price performance. Scores center around zero, with positive numbers indicating a positive outlook and negative numbers indicating a negative outlook. For example, if an earnings call receives a Prattle score of 1, it indicates that the most recent earnings call is 1 percentage point more positive than the average of the 4 earnings calls preceding it.

Prattle scores are calculated using the difference between the most recent CAR score and the average of the previous 4 CAR scores for a given company. The CAR score for an earnings call corresponds to the expected 10-day cumulative abnormal return in the stock price—relative to the market benchmark—caused by the language of a given communication. For example, a score of 1.3 means that the expected 10-day CAR for that company is 1.3% higher than the benchmark. Because they are based on abnormal returns, the scores control for market movement and the risk-free rate.[5]

Prattle chose to base its scores on 10-day CAR for a variety of reasons. Prattle tested 3-, 30-, 60-, and 90-day windows as well, and the 10-day performed better for most companies without overfitting.[6] Beyond performance, extending the window out past 10 days increases the likelihood that non-idiosyncratic and systematic risk—price movement caused by factors other than the earnings call and related information—is included in the CAR, which decreases the ability of the algorithm to map language to identifiable price movement.[7]

While these scores have a variety of practical applications, we'd like to provide you with two key use-cases:

1. **Prattle scores as hot-takes.** Our scores allows users to quickly grasp the sentiment of the communication and what it could mean for a company's stock.

---

[5] Prattle's methodology borrows from the Capital Asset Pricing Model (CAPM). CAR is computed using an asset-specific single-factor market model. Market and asset returns are adjusted for the risk-free rate. Betas are estimated via OLS using a 170-day pre-event period. An intercept is included. It is assumed that the intercept (alpha) coefficient is not statistically distinguishable from zero.
[6] Signals decay at company-specific rates. Therefore, how long sentiment prevails for any company is idiosyncratic. Thus, our system is designed to first and foremost predict short term price moves, resulting from sentiment, following an earnings call. There is value in the signal post 10 days, but it decays and varies by how efficiently a company is covered. If you are trading to capture price relevant factors captured by the sentiment of the call, then the earnings call can be a catalyst to start a trade, add/reduce, or do nothing.
[7] However, Prattle has noticed that different windows may provide performance advantages conditional on market cap. For example, information dissemination for smaller companies is likely not as efficient as it is for larger companies; larger windows may be necessary to map the price movements relevant to the information introduced by a call for smaller companies, and vice versa for larger companies.

2. **Prattle scores as leading indicators.** When converted into time series data, the scores can serve as a leading indicator of company performance. This helps users get a broader view of the relationship between a company's language and its stock price and better anticipate price fluctuations.[8]

Prattle has chosen not to combine its primary source sentiment analytics with sentiment data from secondary sources such as press and social media to avoid introducing bias and diluting the value of the signals. By the time information has had time to be digested by a secondary party, others have also had the opportunity to react to the primary source, and the secondary source may include bias from the secondary entity. This results in shorter and weaker signals. Signals based on social media data, for example, often have aggressive decay rates of less than a few hours.[9]

The Methodology Behind the Scores: Identifying Reference Documents

The foundation of Prattle's algorithmic analysis of corporate communications is the unique lexicons we've built for each company we track. To construct a lexicon for each company, our algorithm is first trained on a series of historical corporate texts known as reference documents.[10] These documents are communications whose content has led to identifiable stock price movement.[11]

Let's look at some examples of communications that could serve as reference documents.

In late May of 2017, Movado (MOV) announced that it missed earnings expectations and revenue projections. Movado's stock price fell 4.5% through the end of the next day; Movado's subsequent 10-day abnormal return was -1.24% (z = -0.07). Prattle's residual score for the call was -0.19 (z = -0.07), accurately predicting Movado's performance.

The Phillips 66 (PSX) Q3 2017 earnings call tells another story. The oil, gas, and consumable fuels giant posted solid earnings and revenue numbers, beating projections. The stock price subsequently increased, but it underperformed the market, yielding an abnormal return of -0.47% (z = -0.21). Prattle's residual score for

---

[8] Additionally, when Prattle scores are used in concert with fundamental data, the synthesized approach can provide a more comprehensive, quantitative view of a company's current status and outlook than has ever been possible.

[9] In regard to performance differences, Prattle encourages prospects and clients to compare our signals to other signals in the marketplace. Our research finds that sentiment scores from non-algorithmic, preset positive/negative dictionaries generate no alpha.

[10] As of Q1 2018, these texts are all earnings calls transcripts. Some companies will do more than one earnings call per quarter. These companies, like Walmart, do a public earnings call with just prepared remarks and hold a separate call for Q&A from analysts. Other companies will do earnings calls for different markets or stakeholder groups. In an effort to be comprehensive, Prattle scores each of these calls. Since Prattle scores the natural language of the call, not the fundamentals, it does not matter if a company releases pre-open and the conference call is post-open.

[11] The machine generates expectations conditional only on observed language. It does not produce expectations based on language that was not used. Therefore, an effect may only be observed insofar as executives may choose not to fill that expected space with something (making other contributions more important within that document), or executives may introduce discussion of something else. While the machine does not explicitly generate predictions on unobserved language, we provide other tools to allow analysts to dig deeper into what was or was not said on the call. Our Core Comments feature algorithmically determines which portions of speech are important and unique for any given earnings call. Analysts can refer to these Core Comments to see if the language they would have expected to hear was present on the call.

the call was a strong -0.17 (z = -2.94), accurately predicting the company's performance. With quantitative fundamentals and the broader market ruled out, the language of the call is the obvious suspect for the cause of this price movement.

The Methodology Behind the Scores: Building Tailored Lexicons

The reference documents used to train our algorithm are mapped directly to varying degrees of stock price movement, making them a reliable, historically-rooted foundation for comparison. Using these reference documents, Prattle has mathematically linked specific words, phrases, sentences, and paragraphs—collectively referred to as "expressions"—to each company's stock price movement.

Prattle data is different from other sentiment products in that humans are not a part of the feature selection process. The algorithm selects features without any priors on what those features should be. As a result, the algorithm picks up on language patterns that might not otherwise be picked up by a person. When humans are involved, their bias is introduced into the process.

For Prattle, the language of corporate communications is understood in terms of what it indicates about future performance. Language patterns that indicate the company is optimistic about product performance, growth prospects, or the broader company outlook are assigned positive values. Language patterns that are historically indicative of a poor outlook are assigned negative values. These expressions and their corresponding weights make up the initial lexicons for each company. Because Prattle constructs a unique lexicon for each company, identical words and phrases could be positive for one company and negative for another.

While Prattle uses these lexicons as a basis of evaluation for subsequent communications, they are not static databases. Rather, they are constantly updated through machine learning. To explain this process, let's take a look at how the Intel (INTC) lexicon was built and is updated.

The Methodology Behind the Scores: Updating the Lexicons Through Machine Learning

Intel's lexicon was originally built using a set of reference documents mapped to company performance over the time period of 2003-2007. Since then, Intel's lexicon has been continually revised through the addition of new expressions and the alteration of the values of existing expressions based on their evolving use.

When our system first encounters a new expression (such as newly-coined words like "WannaCry", "Sandy Bridge", or "Skylake") in a corporate communication, it does not weight that expression because it does not have a basis of evaluation. The second time that expression appears, however, the system is able to weight it based on how it was used previously. In a similar way, the weights of existing expressions in the lexicon are

constantly updated based on the evolving way in which they are used in communications. Language referring to products may suggest positive performance at one point in time and negative performance at another time. The algorithm takes this into account over time. The first time a product mention results in poor stock performance, the algorithm will associate that term with negative sentiment. The latent representation of product-related language will then become more negative, and the next time it sees the language, it will mark it as such.[12]

In short, the weights of expressions within the initial lexicon (the training set) are determined by their association to stock price movement. The weighting of new expressions is then determined by how they are used relative to the language in the training set. Thus, the weighting of the expressions in the lexicon evolves as the algorithm detects changes in their use over time.

Conclusion

This paper has explained the meaning of and methodology behind Prattle's Equities Analytics scores.

What our scores mean

1. Prattle Equities Analytics scores capture the expected change in stock performance caused by the language of an earnings call.
2. These scores can serve as hot-takes on a communication's sentiment as well as leading indicators of company performance.

The methodology behind the scores

1. For each company Prattle evaluates, Prattle selects a set of reference communications whose sentiment can be directly linked to movements in the company's stock price.
2. Prattle then algorithmically maps the expressions in each company's reference documents to their impact on the company's stock price, values the expressions based on their impact, and stores each company's expressions based on its own unique lexicon.

[12] The Prattle system does not weight by speaker or forum. The system is designed to capture differentiation in natural language found in calls as well as differentiation in natural language found at the speaker level as it relates to underlying stock price movement. When a new earnings transcript is scored, our algorithms score the language from the call according to the current lexicon of expressions and historical market reaction to the language used, but once the actual 10-day CAR is realized, the machine learning aspect accounts for the difference in the expected CAR and realized CAR to adjust weightings of current expressions and introduce any pertinent new expressions into the lexicon for future communications.

3. Using these lexicons, Prattle evaluates subsequent communications by each company. Prattle then updates these lexicons through machine learning based on the evolving use of these expressions and the introduction and use of new expressions.

If you have more questions regarding our data or methodology, please feel free to reach out to our team.

Evan Schnidman
617.304.6807
evan@prattle.co

William MacMillan
734.358.5794
bill@prattle.co

prattle.co
