Eagle Alpha

www.eaglealpha.com

Alternative Data
Use Cases

Edition 6

This report is sponsored by Standard Media Index (SMI).

SMI captures monthly advertising revenue results for internet names such as FB, Instagram, GOOG, YouTube, Twitter, SNAP, ROKU, Amazon, Pandora, Spotify and traditional media companies such as Fox, Disney, Sinclair, Trip Advisor and 300+ other public equities across the US, UK and AU markets. Clients couple SMI's data with its machine learning model for FB which has an establish track-record that correlates within 1% of a print typically. SMI releases its data monthly and distributes its FB forecast 6-weeks ahead of a print to its active roster of funds.

Table of Contents

List of Figures

Section 1: Executive Summary

Purpose Of This Report

The purpose of this report is to provide buyside firms with examples of how alternative data can be leveraged. This report is the sixth in a series of reports published by Eagle Alpha.

What's New?

The primary differences between Edition 5, published in early September 2018, and Edition 6, published in late November 2018 are:

1. Updated section on alternative data use cases (Section 4).
2. New sub-section with extracts of interviews with buyside professionals on altdata.tv. (Section 4.5).
3. New sub-section on quantitative approaches to online search data (Section 6).

Alternative Data

Definition

Alternative data is non-traditional data that can be used in the investment process. There are 24 categories of alternative data in Eagle Alpha's taxonomy including consumer transaction data, geo-location data and sentiment data. Our sophisticated clients have already identified 1,500 relevant datasets and several believe that our prediction of 5,000 datasets worldwide by 2020 is conservative.

Benefits

Alternative datasets are being integrated into the investment process primarily because they provide a greater volume of data and information compared to traditional datasets, provide unforeseen insight and are frequently available on a more timely basis than traditional data sources.

Drivers Of Adoption

The primary drivers of adoption include:

- Competitive dynamics / edge e.g. 78% of US hedge funds are already using alternative data.
- Growing evidence of alpha in alternative data e.g. 90% of managers get an ROI on their spend on alternative datasets.
- Providers of AUM expect to increasingly allocate to managers who are utilizing alternative data.
- Risk of being at a strategic disadvantage in the medium to long term.

Tipping Point

Alternative data is not new. Over fifty firms have been working with alternative data for over ten years. 2017 was a turning point for the alternative data space, however, because the broader asset management industry began integrating it into the investment process. In our opinion the alternative data space will 'cross the chasm' (the tipping point) by the end of 2018 / Q1 2019.

Datasets In Demand: Proprietary Analytics

Eagle Alpha provides clients with proprietary data that highlights which data categories and specific datasets are in demand.

Click-Through Data

Analysis of usage of our quantitative clients between Q3 2017 and October 2018 highlighted a few interesting trends. For example, datasets offering business insights and employment insights received a larger proportion of clicks as did consumer transaction data, which traditionally garners strong interest.

1-on-1 Meetings Data

Analysis of 1-on-1 meetings by data category shows business insights as being the most active, same as is seen in the click-through data. Section 3.4 provides profiles of 10 datasets that are seeing a high level of interest by asset managers.

Use Cases

Asset managers constantly ask us for examples of how alternative data can be used. Section 4 outlines use cases for each of the 24 categories of alternative data, by asset class and by manager type. We now also include extracts from a selection of interviews with buyside professionals on altdata.tv.

Asset Class

The majority of use cases apply to the equity asset class, both public and private. Equity use cases, however, are frequently relevant for corporate credit strategies. There are also many macro use cases, which could be used for discretionary or systematic strategies involving financial futures, derivatives, credit securities, commodities, etc. As an example, we outline how the Scientific Active Equity team at Blackrock uses alternative data for CMBS.

Manager Type

Several quantitative and discretionary managers have publicly disclosed some of their use cases. For example, WorldQuant discusses how it uses satellite data, social media data, sentiment data and trade data.

Fifty Case Studies

Asset managers regularly request case studies regarding how alternative data can be used. The majority of this report (pages 38 to 117) is dedicated to fifty case studies that are based on third party datasets as well as Eagle Alpha proprietary datasets and analysis. Many of the case studies have been backtested and/or verified by third parties such as JPMorgan, Citi Research, and academic researchers. The case studies are classified by category of alternative data, asset class and manager type start on page 38.

Online Search Data

Eagle Alpha has been utilizing quantitative methods to derive granular information about interest in specific company product categories and interest in a company stock. We provide an example regarding Home Depot (Section 6).

Eagle Alpha

Eagle Alpha provides alternative data solutions to the buyside.

Complimentary solutions are events, this use cases report and altdata.tv.

We execute a range of bespoke projects for clients. There are typically four types of projects: strategic investment integration projects, research questions, data assignments and alternative data driven models.

The core modules of Eagle Alpha are data sourcing, data analytics and data forum:

- Data Sourcing: this solution keeps clients on top of all the alternative datasets worldwide.
- Data Analytics: this solution enables discretionary managers to work with a variety of alternative datasets through a single dashboard platform.
- Data Forum: this is the leading alternative data industry group focused on legal & compliance, increasing efficiencies and driving best practices.

Eagle Alpha is a recognized leader in the alternative data space. For example, white papers by Citi, JPMorgan, BoA and Deloitte have all profiled the company.

Section 2: Alternative Data

Section 2.1: Definition And Benefits

We define alternative data as non-traditional data that can be used in the investment process. Because alternative data refers to a wide variety of data, it is frequently defined by what it isn't, rather than what it is. Traditional data includes market price data, financial data, and government and industry statistics. Eagle Alpha identifies 24 categories of alternative data, defined by either the collection method or the nature of the data (see Figure 1).

As of November 20, 2018, there were 861 datasets in our database that were spread across these 24 categories. It is interesting to note that some of our clients have already identified 1,500 datasets that are relevant to asset managers. We forecast there being 5,000 datasets by the end of 2020, which several Eagle Alpha clients believe is conservative.

Figure 1: Eagle Alpha's 24 Categories of Alternative Data

Source: Eagle Alpha

Alternative data sometimes qualifies as big data, and sometimes it is small enough to analyse in a spreadsheet. It may be structured or unstructured, or it may be structured data that's derived from unstructured data, such as a sentiment score for text. In our view the word "alternative" is currently in vogue but will fall away. The term the industry will use is simply "data", it will not matter if its traditional or non-traditional (alternative).

The consulting firm Quinlan & Associates[1] published a useful summary of the benefits of integrating alternative data into an investment process (see Figure 2). They identify five benefits:

1. Greater volume of data and information.
2. Unforeseen insight.
3. Competitive edge.
4. Fiduciary duty.
5. Efficiency.

[1] Source: 'Alternative Alpha: Unlocking Hidden Value in the Everyday', September 2017.

Figure 2: Benefits of Integrating Alternative Data into Investment Process

	Alternative Data	Traditional Data	Description
Greater Volume of Data and Information	• Frequently updated • Shorter history • Wide breadth	• Sparsely updated • Longer history • Narrow breadth	• Large amounts of data are being generated constantly, leading to more available information for analysis • Frequent updates mean managers can carry out analysis sooner, enhancing portfolio construction
Unforeseen Insight	• Wide breadth • Profound implications to multiple assets classes, or even industries.	• Narrow breadth • Only provides particular information, normally only directly relevant to few assets.	• New data provides information on more than just financial performance, leading to better trend and performance predictions • Discovery of hidden relationships can help devise investment strategies
Competitive Edge	• Requires investment and data capabilities	• Available to all • Analysis can be conducted relatively easily	• Talent and technologies are needed to gain value from alternative data, giving asset managers with the right resources higher alpha-generation
Fiduciary Duty	• Make use of available information and data	• Make use of financial data and few conversations only	• Talent and technologies are needed to gain value from alternative data, giving asset managers with the right resources higher alpha-generation
Efficiency	• Rapid and efficient research and analysis • Wide coverage	• Highly manual research and analysis • Narrow coverage	• Alternative data provides insights into multiple assets, and can replace certain aspects of the current research process. • Research analysts spend longer time on better modeling or investment strategies

Enhance alpha generation

Source: Quinlan & Associates, Eagle Alpha

Section 2.2: Drivers of Adoption

There are four key factors that are driving the adoption of alternative data:

1. Growing evidence of alpha in alternative data.
2. Increasing asset allocation to managers who are using alternative data.
3. Competitive industry dynamics.
4. Risk of being at a strategic disadvantage in the medium to long term.

1. Growing evidence of alpha in alternative data.

A 2017 Greenwich Associates survey of asset managers[2] highlighted that 90% of asset managers that are using alternative data are seeing a return on their investment.

Figure 3: ROI from Alternative Data

Source: Greenwich Associates Survey in association with Arcadia Data

Clearly measuring ROI on alternative data spend is not always straightforward. For example, some of our clients measure it as a multiple of spend whereas others measure it based on the number of interactions with a centralized 'Data Insights' team.

Other evidence that there is alpha in alternative data includes:

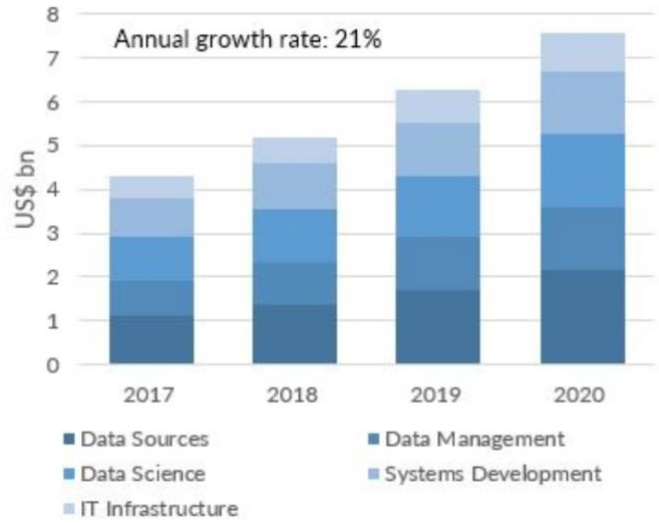
- Manager PR: some asset managers are being vocal about their returns in the press (e.g. NN IP[3]).
- Data vendors becoming funds: some alternative datasets are so successful that they became funds (e.g. CargoMetrics, Dataffirm).
- Case studies: Section 5 of this document provides 50 case studies that demonstrate alpha in alternative data.
- More renewals: as an aggregator of alternative datasets we see an increasing number of datasets obtaining renewals of licenses.
- Teams increasing: a headhunter that specialises in alternative data can show at least 50 asset managers that have steadily increased their alternative data headcount over the last 5 years.
- Spend increasing: Opimas forecast spend on alternative datasets to reach $2 bn by 2020 (see Figure 4).

The Opimas forecasts not only increased spend on datasets, but also on data science, IT infrastructure, data management and system development. More firms are taking advantage of advances made in machine learning (ML) and artificial intelligence (AI) as they leverage alternative datasets. These new analytical techniques facilitate the transformation of alternative datasets for analysis (cleaning and munging) and offer more approaches for extracting information. It's clear that the wave of innovation is adding fuel to the demand for more data.

[2] Source: 'Putting Alternative Data to Use in Financial Markets', September 2017.

[3] Source: 'NN IP Gains From Sentiment Analysis', January 2017.

Figure 4: Spending on Alternative Datasets to Reach US$ 2 billion by 2020



Source: Opimas

More recently a Greenwich Associates survey stated "the average investment firm spent about $900,000 yearly on alternative data…extrapolating this data across the whole market, we estimate that annual industry budgets for alternative data now stand at $300 million – almost double from one year ago."

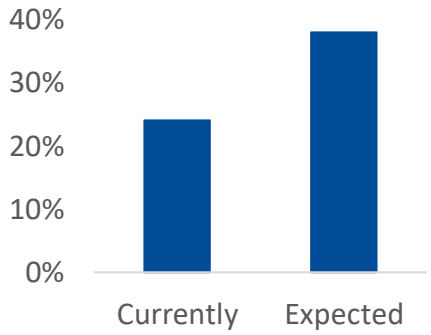
2. Increasing asset allocation to managers who are using alternative data.

EY published an asset management survey[4] in late 2017 that highlighted that institutional investors expect to increasingly allocate to managers who are utilizing alternative data (see Figure 5).

Figure 5: Responses to EY Survey of Asset Managers

What proportion of the hedge funds in which you invest use non-traditional or next generation data and "big data" analytics/artificial intelligence to support their investment process?

What do you expect that proportion to be in three years?



Source: EY, Eagle Alpha

The EY report stated:

- "Investors are continuously searching for the next strategy, asset class or investment opportunity that they believe is unique and can outperform. Given the developments in FinTech and excitement surrounding the technological capabilities to rapidly analyse different datasets, it is not surprising that investors are expecting an increased percentage of their hedge fund managers to be using non-traditional data and new analytics in their investment processes."

[4] Source: '2017 Global Hedge Fund and Investor Survey', November 2017.

- "Many investors view these advancements as an additional tool available to managers and those who are able to effectively harness the capabilities have a distinct advantage compared to those managers who are not deploying these capabilities within their investment process."

- "For years, these tools generally resided in the domain of quantitative managers. However, managers of all strategies have increasingly been innovative in developing methods to complement their investment strategy with these advancements."

3. Competitive industry dynamics.

Asset managers all have access to traditional data and information via traditional channels. Innovative asset managers are seeking an edge through non-traditional alternative data. In the US, 78% of hedge funds are either using or will start using alternative data in the pursuit of an edge (see Figure 6).

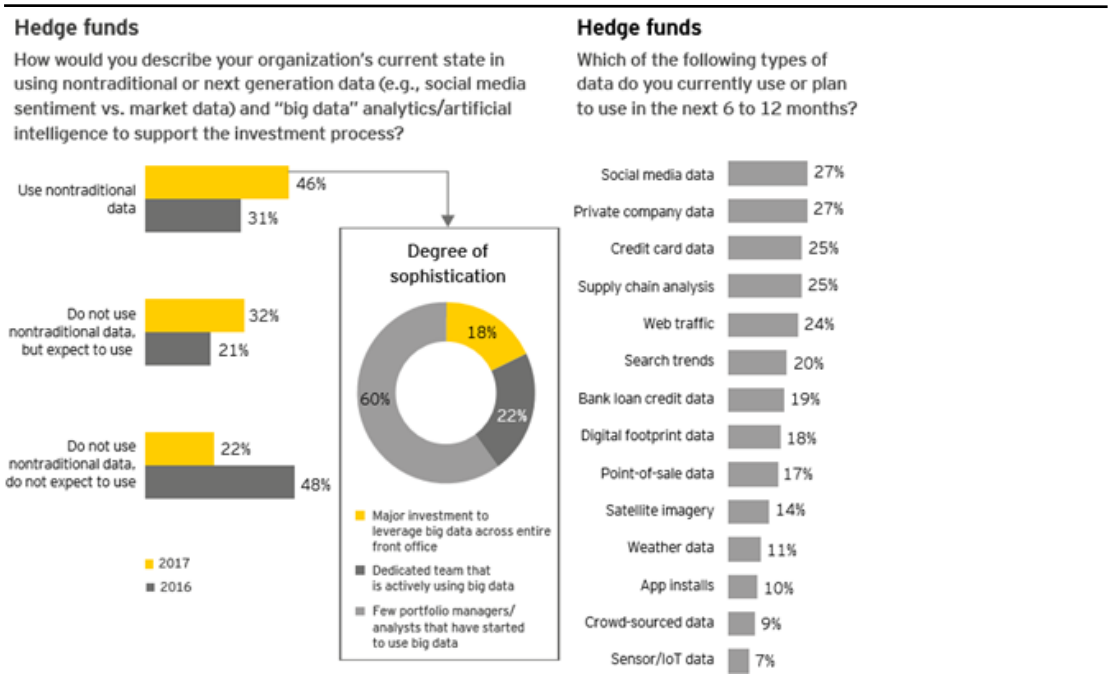
The EY report stated:

- "A larger quantity of managers see effective use of data and analytics as a key competitive advantage for the future. Smaller managers moved first, but managers of all sizes and strategies are now experimenting with big data analytics and AI."

- "What is striking is how quickly the landscape has changed. Last year (2016) almost half of managers said they did not use, and did not expect to use, non-traditional data in their investment process. However, in the current year (2017), 78% currently use or expect to use non-traditional data."

- "For 60% of managers "using non-traditional data and/or AI" means that they have a subset of their front-office teams experimenting with the tools. Fewer managers have fully dedicated teams or have made major infrastructure build-outs to support big data. As managers become more experienced and comfortable with these tools, we would expect to see continued evolution and sophistication around how managers incorporate big data into their investment strategy."

Figure 6: Excerpt from EY's Survey of Hedge Funds



Source: EY

In September 2015 Blackrock's SAE[5] group stated "we believe that in order to generate sustained alpha, investors should embrace acquiring, analyzing and understanding the fast-growing universe of data. Those that are unable to do so run the risk of falling behind in a rapidly changing investment landscape". Since then this view has gained acceptance amongst the wider asset management vertical. For example, at JPMorgan's annual quantitative conference[6] in May 2017 "there was widespread agreement that big data and machine learning is transforming the investment landscape".

4. Risk of being at a strategic disadvantage in the medium to long term.

In October 2017 Deloitte published[7] a report saying that firms that do not integrate alternative data into their investment process will be at risk of being at a strategic disadvantage.

The report stated:

- "Alternative data will likely transform active investment management over the next five years, from hedge fund management, to long-only mutual funds, and even private equity managers."

- "Those firms that do not update their investment processes within that time frame could face strategic risks and might very well be outmanoeuvred by competitors that effectively incorporate alternative data into their securities valuation and trading signal processes."

- "The risk impact and vulnerability for laggards may be much higher as compared to early adopters of alterative data."

As seen in Figure 7, the study noted three risks of late adoption:

1) Positioning risk.
2) Execution risk.
3) Consequence risk.

Figure 7: Risk Exposure Due to Late Adoption of Alternative Data

Source: Deloitte

[5] Source: 'The Evolution of Active Investing. Finding Big Alpha in Big Data', July 2015.

[6] Source: JPMorgan research report entitled '2017 NY Macro Quant and Derivatives Conference', 25 May 2017.

[7] Source: 'Alternative Data For Investment Decisions', October 2017.

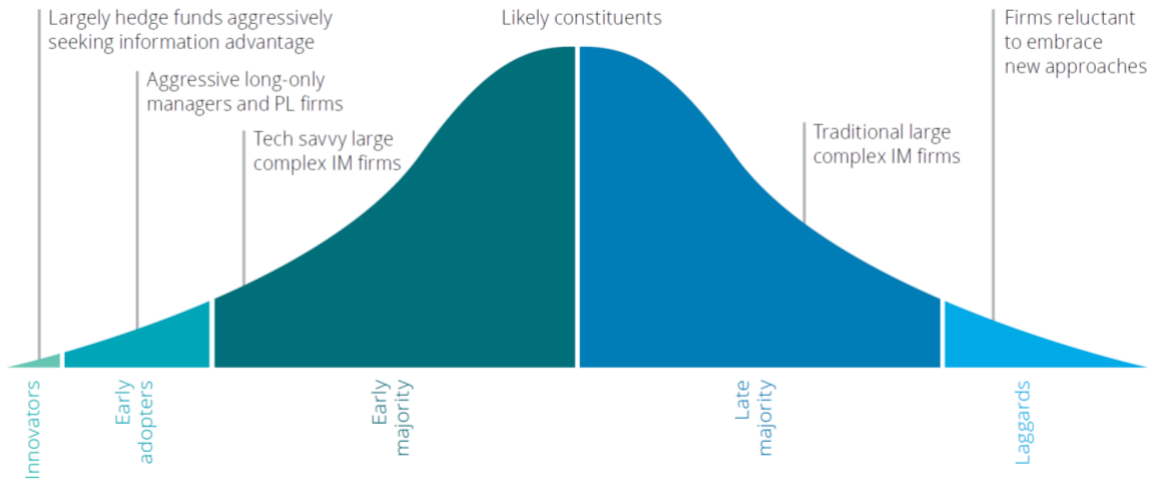
Section 2.3: The Tipping Point

Alternative data is not new given that 50+ innovative asset management firms have been working with alternative data for several years. These firms are primarily the larger quantitative and 'quantamental' firms. Jefferies, in its June 2017 paper entitled "Quantifying Intuition: Mapping the Data Science Landscape in the Hedge Fund Industry", stated that 20% of hedge funds with over $1bn in AUM already have a person dedicated to alternative data or a person is spending 50% of their time on alternative data.

In our opinion, we are moving towards the end of the early adoption phase – see Figure 8. Based on data from our CRM database there are 300 firms worldwide that are seriously working with alternative data today. Our definition of "seriously" is where an asset manager has at least one individual dedicated to alternative data. The majority of these firms are in the US. However, other regions are starting to catch-up. For example, we are aware already 40 discretionary firms in APAC that have at least one individual dedicated to alternative data.

Adoption, and awareness, increased rapidly in 2017. We anticipate the alternative data space will 'cross the chasm' (the tipping point) by the end of 2018 / Q1 2019. This view is supported by an EY survey that stated: "What is striking is how quickly the landscape has changed. Last year (2016) almost half of managers said they did not use and did not expect to use non-traditional data in their investment process. However, in the current year (2017), 78% currently use or expect to use non-traditional data."

Figure 8: Alternative Data Adoption



Source: Deloitte, Eagle Alpha

Between March and May 2018, Greenwich Associates interviewed 30 CIOs, portfolio managers, and investment analysts at investment management firms worldwide[8]. The respondents answered questions about how investment research will change in the next 5-10 years. Regarding whether or not they plan to use alternative data, 50% of investment managers stated that they currently incorporate it into their investment prices. Another 20% plan to in the next 12 months.

[8] Source: Greenwich Associates & Thomson Reuters, "Seismic Shifts; The Future of Investment Research", Q3 2018.

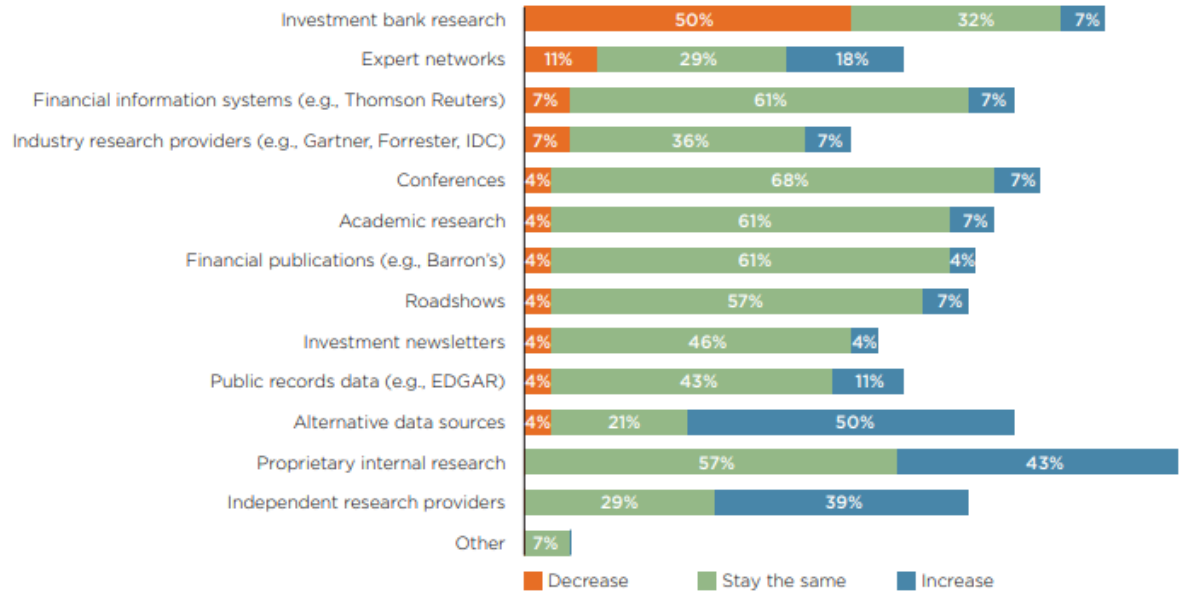
[9] Source: JPMorgan research report entitled '2017 NY Macro Quant and Derivatives Conference', 25 May 2017.

Figure 9: Excerpt from Greenwich Associates' Survey

Source: Greenwich Associates (30 respondents), Eagle Alpha

In response to a question regarding which sources of information managers expect to increase (decrease) their reliance on, the top source that managers cited was alternative data (see Figure 10).

Figure 10: Expected Change in Sources of Investment Research



Source: Greenwich Associates

Section 3: Datasets in Demand (Proprietary Analytics)

Section 3.1: Eagle Alpha Database Click-through Analytics

Eagle Alpha extracts insights from the aggregated click-through data of our global client base. User browsing behavior is used to identify data sourcing trends that reflect changes in buyside interest towards different alternative data categories and specific datasets. The product roadmap for our Data Sourcing solution includes the provision of this aggregated, anonymized data through our online interface. The insights and statistics will be presented for fundamental users, quantitative users and overall, to provide a holistic view of trends.

Eagle Alpha's Data Sourcing team analysed a subset of user's data between Q3 2017 and Q3 2018, identifying some interesting trends. The business insights, transaction and employment categories received the most interest (normalized versus overall category size) in Q3 2018. Open data, social media data and ESG data received the least.

We observed strong year-over-year (YoY) interest (+8.9%) in business insights in Q3 2018, with sequential increases between Q4 2017 and Q2 2018. This has been driven by several high-quality datasets, ranging from telecom portability to corporate flight tracking.

Despite strong interest in the consumer transaction category, we observed a decrease in click-through interest of 8.7% YoY in Q3 2018, following sequential declines between Q3 2017 and Q2 2018. We attribute this to a maturing of the consumer transaction data market, particularly in the United States which is well served by several high-quality providers.

We have observed a material increase in interest in the employment data category from both our quantitative and discretionary clients, with sequential increases in click-views between Q4 2017 and Q2 2018. The employment category includes datasets which harvest IRS and Department of Labor filings, listings of jobs from corporate websites and job "events" from CVs.

Conversely, we have observed a significant decrease in click-through interest in geo-location datasets, down over 2% YoY in Q3 2018. We accredit this to a realization that geo-location data has still some way to go before maturing as a category, with rapid panel growth and a lack of history representing major challenges to working with this type of data.

ESG is a category that observed a material increase in interest from both our quantitative and discretionary clients between Q3 2017 and Q2 2018. The category is appealing to quantitative funds due to its breath of coverage, and low correlation with traditional smart beta factors (please refer to case studies #8 and #9). Despite this, interest in the category has tapered off in Q3 2018.

Interest in trade datasets has declined to roughly 2% from its peak in August of greater than 8%. Interest in social media and sentiment datasets declined in July and August but is surging lately. The two categories account for more than 17% of total views in October 2018. Another interesting trend is that Event Detection data is trending at 9% from less than 2% six months ago.

Clients of Eagle Alpha are provided data showing which specific datasets are obtaining the most, and least, click-views.

Figure 11: Click-Through Analytics Of Eagle Alpha's Database Users

Dataset Category	Q3 2017	Q4 2017	Q1 2018	Q2 2018	Q3 2018	Q3 YoY
Business Insights	6.9%	10.8%	15.1%	20.7%	15.8%	8.9%
Consumer Transactions	14.0%	8.0%	7.0%	5.1%	5.3%	-8.7%
Employment	2.1%	2.8%	5.5%	6.7%	3.8%	1.7%
Event Detection	2.3%	4.1%	3.1%	3.3%	3.3%	1.0%
Trade	4.6%	4.9%	8.6%	7.3%	4.5%	-0.1%
B2B Datasets	4.1%	2.2%	2.8%	3.0%	3.0%	-1.1%
Satellite & Weather	2.8%	7.9%	3.4%	2.5%	6.1%	3.4%
Data Aggregators	3.9%	7.3%	3.5%	4.6%	8.4%	4.5%
Geo-Location	6.0%	6.8%	5.5%	4.0%	3.9%	-2.1%
App Usage & Web Traffic	7.4%	4.1%	3.8%	2.3%	4.0%	-3.4%
Web Crawled Data	4.8%	3.8%	6.1%	8.7%	5.9%	1.1%
Advertising	3.2%	2.6%	2.6%	1.6%	2.1%	-1.1%
Sentiment	4.6%	3.7%	3.5%	5.7%	5.1%	0.5%
Store Locations	1.6%	0.7%	0.7%	0.3%	1.0%	-0.6%
Internet of Things (IoT)	0.9%	1.2%	0.9%	0.2%	0.9%	0.0%
Online Search	5.3%	2.3%	2.0%	0.8%	1.2%	-4.1%
Consumer Credit	1.6%	1.8%	1.3%	0.5%	0.8%	-0.8%
Reviews & Ratings	1.8%	2.4%	2.7%	2.0%	1.8%	-0.1%
Pricing	5.1%	6.6%	6.2%	5.1%	7.7%	2.7%
Public Sector	3.0%	2.7%	2.1%	1.5%	3.2%	0.2%
Expert Views	1.1%	1.5%	0.6%	0.7%	0.8%	-0.3%
Open Data	3.4%	3.7%	4.2%	2.6%	4.5%	1.0%
Social Media	7.6%	5.1%	5.6%	6.0%	5.2%	-2.4%
ESG	1.8%	3.2%	3.2%	4.8%	1.86%	0.02%
Grand Total	100%	100%	100%	100%	100%	

Source: Eagle Alpha

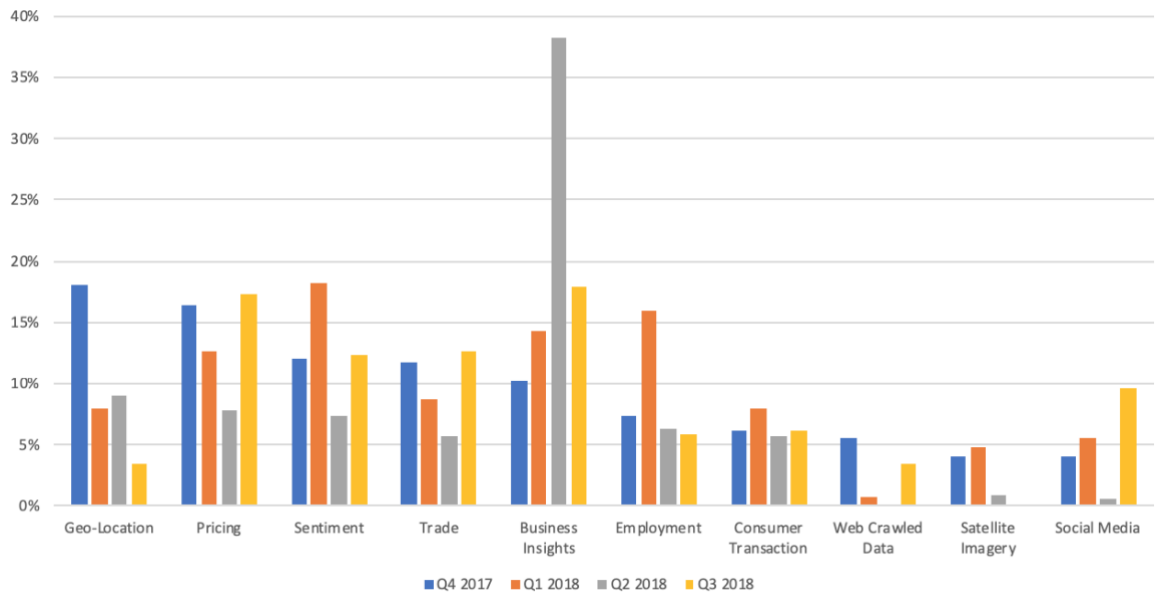
Section 3.2: 1-on-1 Meeting Analytics

In addition to hosting events (in New York, London and Hong Kong) where alternative data vendors pitch to buyside firms we also continuously organise roadshows and 1-on-1 meetings throughout the year. Analytics on these meetings are made available to clients.

In Figure 12, the breakdown of meetings by data category is shown on a quarterly basis. The sum of percentages represented by the bars of a same colour total 100%. The figures are similar to the click-through statistics, but figures reflect which data categories have vendors who are doing a lot of meetings, as well as what is in demand among Eagle Alpha clients.

Business insights is an active category in both click-through and meeting data due to new offerings. Similarly, employment data shows up in both sources as a category of interest in recent months due to high quality data providers and broad-based appeal to fund strategies. A number of new ESG providers caused a jump in meetings with those companies in the second quarter and into third quarter 2018.

Figure 12: Eagle Alpha 1:1 Meetings by Category (% of Total)



Source: Eagle Alpha

Section 3.3: Selected Profiles of Datasets in Demand

Below are summary profiles of 10 datasets that are in demand (based on our proprietary analytics).

Data Vendor #1

This dataset enables funds to identify future corporate transactions based on corporate flight activity. The output includes frequency of the aircraft and the location. The dataset also has data on the location of the suppliers of a particular company. It covers Russell 3000 companies and 150 private equity firms. It has data on 2,000 aircrafts; 405,000 flights; and 40,000 direct relationships between the companies.

Data Vendor #2

This vendor offers a rich and unique view of the global labour force at a company, industry, and global level. The vendor maps half a billion individuals, more than 10,000 global public companies, and millions of private and non-corporate entities (government, education, military, healthcare, etc.) in order to capture workforce dynamics. History: since 2007.

Data Vendor #3

This credit score data provider delivers predictive stress scores for credit, supply chain and financial professionals. Credit ratings are prepared for over 57,000 public companies worldwide and can be used to analyse equities with high bankruptcy risks. Aggregate crowd-sourced usage data from the vendor's subscribers, credit managers and supply chain professionals from Fortune 1000 companies, is incorporated during the construction of stress scores. History: since 2007.

Data Vendor #4

This provider delivers investment signals based on sentiment and emotions analysis of over 120,000 data sources, focusing on social trading websites and specialized press. Signals are available for more than 2,000 tickers as well as bonds, currencies and commodities. History: since 2014.

Data Vendor #5

This company is the global leader in telecoms ID data – number portability, number management, device IDs. They work with over 500 telecoms companies with coverage across US, EMEA and LATAM. The provider has access to central databases showing the churn of customers between service providers in real time. History: since 2014.

Data Vendor #6

This vendor leverages filings with the IRS and Department of Labor that pertain to employee benefit plans. The vendor tracks every company operating in the U.S. and their locations, accounting for over 95% of U.S. employment. The source filings are audited, cash based, traceable, and updated monthly. For approximately 4,000 tickers, the vendor provides point in time factors mapped to tickers relating to employment, cash contributions, benefits, growth, profitability, and credit. History: since 2010.

Data Vendor #7

This vendor's data tracks executives and alerts asset managers to changes in the C-suite's potential to deliver positive business outcomes and maintain a highly functional, stable, and performance driven executive management team. It reveals how company individuals and team as a whole navigate business cycle transitions and systemic volatility. History: since 2008.

Data Vendor #8

Completely unique in the industry, this job listing dataset only indexes jobs directly from employer websites. Updated daily with over 4 million jobs from more than 30,000 employers, the platform eliminates duplicate and expired job listings, as well as job pollution. From the core platform, the company has developed an array of products and services and achieved significant market traction in two primary business units: candidate sourcing and job market data and analytics. History: since 2007.

Data Vendor #9

Powered by advanced artificial intelligence-based mail platform, this vendor provides comprehensive and up-to-the moment ecommerce trend data. Clients use this data to understand the sales, revenue and spend trends across over 600 companies and 11,000 brands. The data is gathered from users who have agreed to share anonymized and aggregated data about their purchasing and travel behavior through analysis of their email accounts. Through maintaining the direct relationship with users, the vendor has clear, compliant, data access rights. History: since 2013.

Data Vendor #10

This vendor delivers high frequency and comprehensive South Korea export data. Preliminary export data includes total volume and value of all products, at all country destinations that are exported from South Korea. The dataset is created by aggregating and analysing customs declaration forms. History: since 2003.

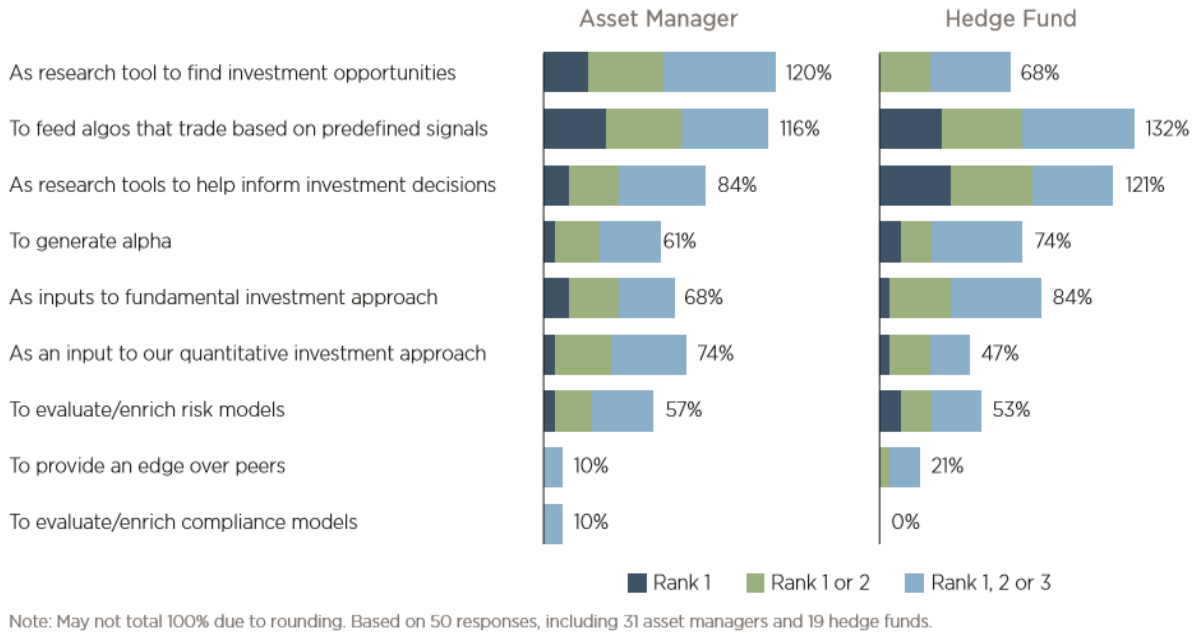
Section 4: Use Cases

Section 4.1: Overview

In this section we outline the different use cases of alternative data based on a Greenwich Associates survey in association with Arcadia Data. In addition, we summarize applications relevant to different categories of alternative data, asset classes (equity, macro and credit) and types of asset managers (quantitative funds, discretionary hedge funds and traditional fundamental asset managers). Finally, we provide transcripts from a selection of altdata.tv interviews.

Greenwich Associates survey of asset managers[9] in association with Arcadia Data, shows that use cases trended towards specificity, not generalities (Figure 13) and towards details (specific anomalies, micro level analysis) – see Figure 14.

Figure 13: Using Alternative Data

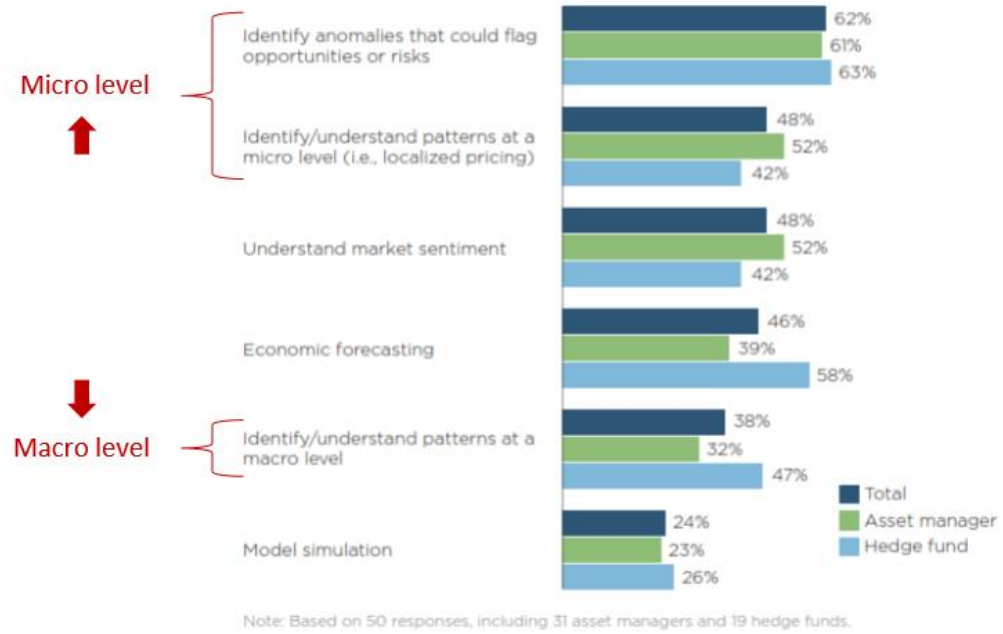


Source: Greenwich Associates Survey in association with Arcadia Data

In an August 2018 poll of Eagle Alpha clients, of which none ran a pure macro strategy, the number one reason stated for trialing a dataset was its potential to answer a specific research question. The next most common response was that it is from a data category that typically yields many insights. These responses are in line with the responses (seen in Figure 14) to the question, "How do you use alternative data?".

[9] Source: 'Putting Alternative Data to Use in Financial Markets', September 2017.

Figure 14: How Do You Use Alternative Data?



Source: Greenwich Associates Survey in association with Arcadia Data

Section 4.2: By Category of Alternative Data

Eagle Alpha created the first taxonomy for alternative data. This taxonomy has 24 categories and was developed to make the vast array of data types more understandable and approachable for applications.

Figure 15: Eagle Alpha's 24 Categories of Alternative Data

Source: Eagle Alpha

The most common categories, based on our dialogue and analytics, include consumer transaction, geo-location, satellite and sentiment. As at 20th November 2018, there were 861 datasets in our database.

Access to detailed profiles on these datasets and insights into the most (and least) common datasets, based on our dialogue and analytics, are available to clients of Eagle Alpha's Data Sourcing solution. Below we highlight a number of examples of applications for each of the 24 categories of alternative data. These examples are based on our discussion with asset managers since 2012 and Eagle Alpha's own research efforts. For each category we state the number of relevant datasets in our database (as at November 20th, 2018).

1. Advertising (26 datasets): data aggregators track corporate advertising spending on various platforms and by campaign. Advertising data exchanges have data on consumer interests over time based on their internet browsing habits. This data could be used to track category popularity e.g. luxury products and financial products such as mortgages, automobiles, cybersecurity etc. Very few investors have worked with this data and its predictive potential is largely unexplored. Another type of data in this category comes from data aggregators who monitor print, television and online media. The data is typically a sample or estimate rather than an accurate accounting. Nevertheless, it can be used to track corporate marketing messages.

2. App Usage & Web Traffic (51 datasets): web browsing traffic, both online and mobile, can be used to estimate company revenues, particularly if the web pages providing purchase confirmations are tracked. Mobile app usage data tracks the number of downloads and time spent using apps. It has been used to gauge the popularity of social media platforms, mobile games, media providers, e-commerce platforms, financial services, travel and lodging providers, software products and other consumer services and products. Trends in mobile app reviews can also help analysts evaluate product success. Country specific data can provide insights into product adoption internationally. Investors can also track services embedded in apps such as payment providers and advertising services. App usage and web traffic data is frequently volatile and in many cases a more accurate signal can be provided by consumer transaction data.

Please refer to the case study #4 on page 45 and #33 on page 92.

3. B2B (32 datasets): a variety of data aggregators offer datasets about enterprise B2B commerce, including supply chain analytics. Some of these datasets are relevant for a range of industries, such as ones that monitor enterprise level internet browsing activity and Alibaba's B2B trade index. Other datasets provide niche information, such as databases of industrial materials and databases of oil contracts and drilling concessions.

4. Business Insights (172 datasets): a heterogeneous group of datasets that provide unique insights into companies. One example is datasets that track intercorporate business connections. Other data providers track credit quality related business activity and/or apply machine learning techniques to large quantities of aggregated data in order to identify companies at risk of failing. Natural language processing algorithms applied to corporate communication text data also falls in this category.

Please refer to case studies #1 (page 42), #2 (page 43), and #14 (page 59).

5. Consumer Credit (15 datasets): marketplace lending data is frequently updated daily, showing amount of loan issuance, loan pricing, borrower credit quality and defaults levels. Other data providers track consumer credit quality overall in specific countries. This data, which is typically more timely than other sources, can be used for determining momentum and inflection points in the consumer credit cycle.

6. Consumer Transactions (36 datasets): this data can come from a variety of sources and can provide merchant level transaction data (e.g. retailer, airline, service provider), product level purchase data (e.g. food, beverages, electronics) and macro level data. Some data sources, such as credit card transaction data, represent a large user base. Other data sources involve smaller panels, such as 2% of consumers, yet still provide reliable signals.

Consumer transaction data is frequently used to estimate quarterly revenue growth as the data is available before quarterly corporate earnings are released. However, consumer transaction data can also be used by long-term investors to gain insights into consumer purchasing behaviour. Examples include rate of product adoption, trends in purchases of "premium" products, the effects of promotions and discounts, customer demographics and co-purchase behaviour. In

addition, payment processing data, such as usage of PayPal and Square, is frequently identifiable in consumer transaction data.

In April 2018 Eagle Alpha launched a consumer transaction dataset called RevCast based on data from a partner that is a prominent consumer transaction company and online search.

Please refer to case studies #6 (page 47), #12 (page 56), #13 (page 57), #15 (page 61), #16 (page 62), #17 (page 64), and #29 (page 84).

7. Data Aggregators (104 datasets): technological innovation has allowed aggregators to collect data from disparate sources and aggregate that data in a format that is helpful for asset managers. Aggregators may mine the deep web or carry out timely analysis of government filings and releases. Other aggregators operate exchanges, or platforms, where datasets may be purchased.

8. Employment (20 datasets): listings of job postings can be used to evaluate corporate strategy and direction, industry growth rates, and demand for specific skills. For example, is the demand for candidates with experience in Tableau, or Google AdWords, growing or plateauing? Another data provider tracks changes in corporate employees allowing analysts to identify companies with high employee turnover rates or companies with strong salesforce growth.

Please refer to case studies #5 (page 46), #10 (page 54), #22 (page 71), #27 (page 80), #43 (page 107) and #46 (page 110).

9. ESG (30 datasets): Alternative data sources can provide insights into the environmental, social and governance (ESG) standards at a company. Asset managers tend to use ESG characteristics for three purposes: 1) evaluate impact on risk/return of a portfolio; 2) early identification of risks; and 3) identify sustainable themes as alpha drivers, e.g. low carbon, clean energy and water, healthcare and education, sustainable supply chains, etc.

ESG criteria may be monitored via a variety of data categories, including social media, satellite, open and public data. Additionally, sources which monitor business complaints, business reputation, employee compensation and hiring trends can also be useful. Based on our conversations and analysis we believe it is difficult to establish a complete ESG framework without using alternative data.

Some vendors providing ESG data deliver frameworks that generate scoring. For example, one vendor scans tens of thousands of unstructured web sources and composes ESG scores for over 8,000 companies. Other vendors provide specific datasets that allow ESG analysts to focus on a certain factor within an ESG framework. We see growing demand for the latter as asset managers are starting to perform scoring in-house and creating their own internal databases of datasets.

Please refer to case studies #8 (page 50) and #9 (page 51).

Environmental Issues	Alternative Data Categories
Climate change and carbon emmissions	Event detection
Air and water pollution	Satellite & weather
Biodiversity	Trade
Deforestation	Internet of Things
Energy efficiency	B2B
Waste management	
Water scarcity	

Social Issues	Alternative Data Categories
Customer satisfaction	Reviews & ratings
Data protection and privacy	Employment data
Gender and diversity	Event detection
Employee engagement	Sentiment
Community relations	Social media
Human rights	
Labor standards	

Governance Issues	Alternative Data Categories
Board composition	Employment data
Audit committee structure	Web crawled
Bribery and corruption	Sentiment
Executive compensation	Expert views
Lobbying	
Political contributions	
Whistleblower schemes	

10. Event Detection (44 datasets): alerts to breaking news from major news wires or social media sources allow traders to react before news is fully discounted in asset prices. Other events monitored include government filings and weather.

11. Expert Views (15 datasets): topic and sentiment trends among experts in any industry, or field of expertise, can differ substantially from the trends observed in the general population and in news feeds. The volume of information shared via niche blogs and forums make it difficult for investors to synthesize all of the commentary. Natural language processing (NLP) tools can aid in summarizing sentiment and topics.

12. Geo-location (63 datasets): location data derived from mobile devices can yield timely information on visitation trends. Common industry applications include amusement parks, retailers, restaurants, hotels, travel, transportation and REITs. In addition to observing the levels in foot traffic, this data can be used to identify the impact of promotions and weather events. Cross brand loyalty and regional idiosyncrasies may be identifiable. Geo-location data providers receive location data from mobile app owners, bluetooth beacons and sensors.

Please refer to case studies #4 (page 45), #19 (page 68) and #24 (page 74).

Eagle Alpha's Search Signals dataset consists of a series of company revenue indicators constructed using Google search volumes for a company's product offering. Observing crossing points of various moving averages of an index provides an insight into search momentum for a company's products, and has demonstrated a relationship with reported company sales metrics.

We have found across our software coverage that the best moving average to use for our signal is a 12, 6 and 3 month moving average. This makes intuitive sense given the business model and the sales cycle of software companies.

13. Internet of Things (IoT) (15 datasets): consists of data derived from internet connected devices. Sensors provide data on traffic, which can be used to gauge local economic activity for real estate purposes or to track activity around warehouse distribution centers. Sensors can provide valuable information on agricultural crop health. Sensors can also track flow volumes in oil and gas pipelines. As more sensors types are deployed, Eagle Alpha expects a growth in the variety of data available from sensors.

14. Online Search (20 datasets): consists of data collected by search engines regarding the frequency of terms searched. Google search and Baidu are the largest providers of search data. Numerous academic studies have been published establishing that data regarding the volume of online searches can be used as an indicator of economic activity, as well as an indicator of consumer interest in a product or topic. Furthermore, these studies show that the best indicators are generally built with data from a basket of terms as opposed to a single term or a small number of terms.

Complex data science techniques are used to determine the most indicative search terms and the most effective model for combining those terms into an indicator. Online search data has more than 10 years of history and is available in a timely fashion. It is notably broad in its topic coverage.

Please refer to case studies #11 (page 55), #12 (page 56), #17 (page 64), #22 (page 71), #27 (page 80), #30 (page 86), #31 (page 88), #32 (page 90), #36 (page 96), and #42 (page 106),

15. Open Data (86 datasets): A tremendous amount of data is becoming available as open data. CKAN, Comprehensive Knowledge Archive Network, is a non-profit registry of open data. CKAN prepares data and provides access to data in ways that make that data more discoverable and usable. The CKAN data management platform is in use by numerous governments, organisations and communities around the world. Examples of open data that are relevant to investors include:

- Open Charge Map API that allows users to access data on locations of electric vehicle charging stations.
- The Wayback Machine provides a historical archive of internet pages which may be useful when backfilling data for a web crawling program.
- The GDELT Project provides a platform that continually records the world's news media from nearly every corner of every country in print, broadcast, and web formats, in over 100 languages, and provides a historical archive of news media content.

16. Pricing (109 datasets): aggregated pricing data of goods and services for both businesses and consumers is now more readily available than it has been in the past. This data can provide insights into corporate revenues and industry competition. Alternative measures of inflation have been developed using web crawled pricing data. This category also includes real estate sales, leases and rentals.

For case studies using pricing data, please refer to #18 (page 66), #25 (page 75), #26 (page 78), #28 (page 82), and #49 (page 115).

17. Public Sector (58 datasets): government agencies publish large datasets that can be used to gauge both social and economic activity as well as industry dynamics. Many of the datasets provide granular data from local governments as well as aggregations at the national level. This data is frequently not well indexed and can be so large that it is cumbersome to work with. Because of this, it is possible to find information that is not well discounted by the market.

In H1 2018 a new supply chain shipping dataset launched. The dataset represents transactions which are subject to tracking or taxation by the US Customs and Border Patrol (CBP). In general, this covers shipments entering the United States via a maritime vessel through a federal port, along with a small number of exports and "pass-through" transactions tracked by the CBP. Unstructured data from the CBP is transformed into a normalized view with clear parent-subsidiary entity linkages, allowing investors to to expose, integrate, and consume the data in meaningful ways.

Contact us for an introduction to this data vendor.

18. Reviews & Ratings (29 datasets): product and service reviews posted online can be harvested and analysed for ratings trends and frequently mentioned topics. Numerous academic studies have shown that consumers place credence in online reviews and that favourable reviews generally lead to increased sales. At the same time, excessive negative reviews and complaints can be signs of poor management. App reviews can provide insights into consumer satisfaction with app services such as mobile banking. Other data providers track brand reputations by incorporating a variety of sources that gauge consumer and B2B opinions, including surveys.

Please refer to case studies #37 and #38 on pages 98 and 99.

19. Satellite (68 datasets): the interpretation of satellite images into data or intelligence is useful to asset managers on many fronts. It has been used as a data source for models which track industrial production, particularly in developing countries where there is scarce timely information. It can be used to track activity at mines, construction sites, plants, and retail locations. Satellite data is also used to estimate oil and gas inventories and production. It has been found to accurately predict quality of agricultural harvests. In addition to satellite, drone imagery is being utilized with increased frequency.

Please refer to case studies #20 (page 69), #21 (page 70), #40 (page 102), and #41 (page 104).

20. Sentiment (74 datasets): scoring of news feeds and social media posts by sentiment and novelty is a popular data source, especially for quantitative funds, due to its relatively longer history and columnar time series structure. Sentiment scoring may be applied to investor commentary, consumer attitudes toward products and brands, or mainstream news feeds. Sentiment data providers, in addition to mapping articles to entities such as government agencies and publicly traded companies, may provide additional scores relating to topic novelty, relevance, price impact estimate, and momentum. Data can be applied to factor models or used ad hoc in momentum and contrarian trading strategies.

Please refer to case studies #3 (page 44), #7 (page 49), #39 (page 101), and #47 (page 111).

21. Social Media (105 datasets): data from social media platforms can be used to analyse consumer trends, reception of product launches, brand popularity, customer satisfaction, product sales promotions, social and political movements, and corporate/customer engagement. Brands with a growing number of unique individuals engaging with that brand on social media have shown to have favourable sales momentum and brand strength is frequently a driver of stock prices.

Please refer to case studies #11 (page 55) and #34 (page 93).

22. Store Locations (16 datasets): tracking store locations can yield insights into corporate growth and strategy, particularly when store hours and promotions are also tracked. Store location data can also be used to evaluate addressable market size and market saturation.

23. Trade (42 datasets): macro firms leverage new alternative trade datasets for balance of payment estimates, insights into major commodity markets, indications of national competitive advantages and indications of consumer strength. Stock focused strategies use trade data to gauge sales of companies whose products can be linked to imports/exports of specific goods and to analyse supply chain activity. Trade data can also be used to gauge activity of transportation companies and publicly traded ports.

Please refer to case study #35 (page 95), #44 (page 108), #45 (page 109), and #48 (page 112).

24. Web Crawled (70 datasets): web crawling is a means of aggregating price, social media, ratings/reviews, employment and store location data via a computer program which requests information from public URLs. Web crawling is also employed to monitor corporate websites for changes such as buildout of website structure that reflects strategic initiatives, increased content in certain product lines, increased blog activity, promotional campaigns, and geographic expansion.

Web crawling can be used to monitor niche e-commerce sites and sites that offer specific services such as solar installations or software services. Information on government filings can sometimes be best accessed via web crawling. Data can be collected in-house or by companies that specialize in customized data collection. Datasets containing historical crawled data have been accumulated by web crawling companies and data aggregators.

Section 4.3: By Asset Class

In this section we outline examples of applications for the equity, macro and credit asset classes.

Equity

Equity investors are the biggest users of alternative data. Whilst the consumer and technology sectors are the most common there are alternative datasets available for every sector. Below we highlight a few examples of applications across 11 sectors and also disclose the number of datasets within each sector based on our current database (as of November 20th, 2018).

1. Consumer Discretionary (135 datasets):

For retailers, restaurants, and online travel companies, consumer transaction data can provide near real-time indications of consumer spending as well as longer-term insights into consumer trends.

For product and service companies online search data has proven valuable at identifying inflection points in consumer interest that ultimately lead to changes in revenue momentum.

Data crawled from e-commerce sites can also provide insights into product pricing dynamics as well as product traction as some sites disclose sales or sales rankings.

Social media data and sentiment scoring can aid in understanding customers' perceptions of brands and products.

2. Consumer Staples (85 datasets):

Consumer transaction data can provide valuable insight into performance of individual brands and retailers. ePOS data providers are particularly valuable for staples as they traditionally have strongest coverage for supermarkets and other consumer staples retailers.

Studies have shown that consumers who engage with consumer staples brands on social media tend to be active buyers of those brands. Social media data can provide indications of brand momentum, brand messages that resonate with consumers, and corporate strategy for customer acquisition.

3. Energy (55 datasets):

Trade data can provide insight into both demand and supply over the short and long-term. Not only does trade data capture current energy exports and imports, it also can be used to track upstream equipment imports. This data can be used for macro and micro economic applications.

Satellite imagery can be used to track storage tank utilization across the globe. There are data providers that provide comprehensive information on the US oil and gas industry offshore, onshore and fracking. There are similar data providers for the Canadian market. Satellite imagery can also be used to monitor the facilitation of oil fracking sites across the US.

Sensor data can be used to track the flow of oil and natural gas through pipelines.

Credit risk data can be used to evaluate the real-time credit quality of oil and gas service providers, which can come under stress due to cyclical factors. Similarly, employment data can be used to evaluate cyclical pressures or upswings.

4. Financials (62 datasets):

Growth in peer-to-peer lenders and other non-traditional lending channels has created a wealth of data to better understand the health of consumer credit and take the temperature of the lending market.
App data can give insight into the adoption of financial services especially amongst millennials.

Crawling data from consumer complaint databases can give an early indication of scandals involving financial institutions. In the U.S., regulators were alerted to the irregularities with Wells Fargo accounts from complaints to the Consumer Finance Protection Bureau database, which is open to investors use.

5. Healthcare (52 datasets):

Trade data tracks real-time imports and exports of medical devices, allowing investors to monitor product demand. Aggregators source data from healthcare customers to provide data on medical device / hospital supply procurement.
Other firms provide data on clinical trials and drug pricing which can be used in conjunction with social media commentary surrounding clinical trials.

6. Industrials (71 datasets):

Trade data enables analysts to track industrial company shipments such as specialty glass or auto parts. It also allows analysts to track the impact of trade tariffs and duties.

Geo-location data provides real-time information on passenger air traffic. GPS and trade data track the activity of ships and ports. Satellite data firms have been known to monitor large factories and key infrastructure projects which may add production efficiency or capacity.

Many industrial conglomerates are focused on niche markets that can be tracked by expert views and surveys of business customers. Due to technological changes, surveys have become significantly less expensive to administer.

7. Information Technology (56 datasets):

Hardware analysts can track the supply chain of major technology companies as well as the shipment of finished products by using trade data. Search interest for a particular software company's products and services can be monitored and compared to revenue growth. Traction of internet companies and online sales can be analysed by looking at mobile app data and email receipt data.

8. Materials (42 datasets):

Trade data can be used to analyse commodity flows on a global basis. This can be used for macroeconomic use cases as well as company fundamentals. Satellite imagery can be used to facilitate supply projections for agricultural commodities and to track mining and shipments of raw materials such as metals and ore.

9. Real Estate (40 datasets):

Using geo-location data, it is possible to track footfall to major shopping malls. This helps measure the health of mall sales in an environment where online commerce is taking market share.

Eagle Alpha has partnered with a UK housing data provider that offers data on UK house sales and rentals that covers 80% of the market.

Some data providers have information on inventory of commercial properties, building availability for lease or sale, existing tenant information and historical trends on demographics, occupancy and lease rates.

Real estate pricing trends in emerging markets countries can aid analysts in understanding the breadth of economic growth as well as inflationary pressures.

10. Telecommunication Services (22 datasets):

The market share figures of mobile telecommunications service providers is fairly easy to estimate due to signals emitted from mobile phones that are collected by a variety of intermediaries. Furthermore, network quality metrics such as latency, calls dropped, upload & download speeds, number of cell towers, etc are also readily available. Service contract pricing can be obtained for retail investors from web crawling.

11. Utilities (28 datasets):

Sensor data can be used to track the flow of natural gas through pipelines as well as the amount of electricity traveling over transmission lines. Power prices are tracked by a number of independent data providers.

Eagle Alpha's Web Queries tool can track discussions about regulatory issues such as permitting and pricing. Consumer complaints to government agencies and social media commentary can be used to track ESG issues for utilities.

Macro

Wells Fargo[10], in a paper in April 2017, gave a useful high level summary of applications related to macroeconomic analysis: "big data could help analysts solve many modern-day puzzles, such as productivity growth and its living standard relationship, micro-foundations of macroeconomic models, consumer/firm/investor behaviour and many more. In addition, big data would help include what is often the missing link of demographics in many economic/financial theories such as the consumption function. For instance, the current consumption function employed to analyse consumers' behavior estimates an average behavior that does not distinguish consumers' behavior by demographic or geographic region. Another potential utilization of big data would be to improve current methods to estimate the state of the overall economy as well as different sectors' / regions' performances. For the financial world, big data would increase opportunities for profits and help manage risk more efficiently by incorporating broader information in risk modelling".

Below we highlight examples of more specific applications across 8 macro categories and also disclose the number of datasets within each macro category (based on our database as of November 19, 2018).

1. Commodities (152 datasets): in 2016 Cargometrics (a shipping data vendor) was so successful that it turned itself into a hedge fund. In August 2017 Maersk invested in Cargometrics. There are several similar datasets that can be used by commodity investors to track supply and demand dynamics. Satellite data is also instrumental in forecasting agricultural commodity supply/demand as well as other raw materials such as iron ore.

2. Current Account (49 datasets): one of Eagle Alpha's data partners provides a trade nowcasting dataset that gives estimates of imports and exports for over 10 countries including China and Brazil. The timely and detailed breakdown of trade accounts allows analysts to better identify sustainable competitive advantages at the country level. Real time export data can be provided on a weekly basis at an aggregated level. For example, one can accurately track China export data by assessing South Korean export data which is captured on a weekly frequency.

3. Financial Instruments (115 datasets): data sources that detect breaking news via social media are highly relevant to financial asset prices. Similarly, sentiment analysis of news and blogs has been found to correlate strongly with interest rates and foreign exchange rates. Sentiment analysis of central bank communications is an example of a highly tailored

[10] Source: 'Big Data Applications in the Economics/Financial World Part I: Opportunities and Challenges', April 2017.

application that may have a better track record than more mainstream natural language processing of central bank press releases.

4. Housing & Real Estate (106 datasets): The Bank of England uses a dataset of online residential housing listings as this dataset provides a timelier and region-specific picture of real estate activity compared to traditional sources. It also provides a more real-time indication of residential real estate financing. Similar datasets, and ones that capture rental rates, are available for numerous countries.

5. Inflation (94 datasets): MIT's Billion Prices Project ("BPP") collects prices from hundreds of retailers to construct inflation indices. BPP suggests that its index is a good measure for predicting the U.S. inflation rate. Numerous data aggregators collect online pricing for industrial as well as consumer goods, providing valuable insights into sources of pricing pressures and price deflation.

6. Labor Market (74 datasets): one of the most popular datasets used by investors provides timely and granular data regarding the labor market, particularly the U.S. labor market. This dataset has 8 years of history, is mapped to thousands of tickers, is updated daily and can be analysed at various levels such as state, industry and job category.

7. Personal/Household Sector (126 datasets): consumer transaction data is used to forecast private consumption. Datasets with aggregated credit and debit card transactions are available at the regional and sector level for the US and China. These datasets are typically updated weekly, thus are timelier than government monthly figures. These datasets also offer a check on official government statistics.

8. Surveys/Cyclical Indicators (90 datasets): several asset managers have leveraged online search data to build indicators for topics such as housing and employment. For example, the Scientific Activity Equity team at Blackrock[11] published a paper that gave an overview regarding how it uses internet search data to predict U.S. retail sales.

9. Interest Rates: sentiment and Natural Language Processing can be used to analyse Central Bank releases. A hawkish or dovish assessment can be a strong predictor of yield curve shifts in the subsequent period following the release.

For detailed macro case studies please refer to case studies #39-48 on pages 101-114.

Credit

We have seen credit investors use consumer defaults data to gauge consumer credit, geo-location data to track distressed situations of retailers and analysis of news and social media for municipal bonds (e.g. local government bankruptcies). Below we detail specific examples regarding muni-bonds and commercial mortgage backed securities.

1. Municipal Bonds: in August 2017, IHS Market published a note[12] regarding how alternative data can be used for muni bond analysis. Below are the selected highlights:

- "Municipal bond investors are beginning to examine more "non-traditional" datasets for making investment decisions and surveillance given the potential wide gaps between reporting periods".
- "We believe that the inevitable rising rate environment will take away some of the cushion that has recently allowed municipalities to refinance into lower debt payments, which will increase the demand for new sources of data to more effectively price and tier municipal bond risk".
- "Using Puerto Rico's ports to gauge government income...the correlation is 54% and the R-squared is 0.2908 between the four-month lagged ship count and actual monthly tax collections from January 2014 to June 2017".

[11] Source: 'The Evolution of Active Investing. Finding Big Alpha in Big Data', July 2015.

[12] Source: 'Boats, Quotes, and Automobiles: Alternative Data for Municipal Bond Investors', August 2017.

- "Using auto registration data to determine demographic shifts. A proxy for population migration is the changes in the U.S. auto vehicle registrations…the migration of new luxury vehicles between states is one potential gauge for the movements of higher income and net worth individuals among states".

2. Commercial Mortgage Backed Securities (CMBS): The Scientific Active Equity team at Blackrock outlined[13] how alternative data can be used for CMBS:

- "On the surface, the universe of CMBS appears to be a rather unruly dataset – but beneath the surface are rich veins of data at the individual loan or pool level. By combining machine-learning techniques and portfolio manager expertise a deeper understanding of the underlying data can be gained".

- "In order to extract value from data on tens of thousands of CMBS loans, it is necessary to write programmes that can read important information that is generally contained in an annex to the prospectus supplement for each security. This is a complicated process with a high potential for error that requires strict quality control standards. In addition to the top line data that a well-designed algorithm can extract, CMBS can also contain subordinate financial details that are buried within sub paragraphs or footnotes of loan documentation. These can significantly alter the top line data".

- "The details of this kind of subordinate financial need to be assessed manually by experts with product specific knowledge in order to gain a more complete understanding of the risk profile of individual loans. If managers are able to combine the necessary human market expertise with a powerful computing platform, it is possible to build a model that can help to better predict loan default".

For a detailed case study regarding credit please refer to case studies #49 (page 115) and #50 (page 117).

3. Credit Default Models: Using private company data and combining it with social media sentiment and online search data, it is possible to build a more accurate public credit default model. There are 27 million companies in the U.S. which comprise 70% of U.S. GDP. Building models on these companies is more statistically robust and can be used on publicly traded companies and credits. Furthermore, the number of U.S. publicly traded companies has dropped by 50% in the last 20 years and the real economy is how B2B companies interact with each other.

4. Credit Scoring: Social physics is now allowing for more precise targeting of lending pools to qualified consumers. Mortgage lenders are now able to anticipate pre-pays more accurately.

Private Equity

Most alternative data sources produce information that is actionable in the short-term. Equity focused managers are the predominant users of alternative data because of the real time insights provided. The private equity approach to alternative data has a lot of similarities with that of public equity managers. In addition, we see overlap with the requirements of credit managers that are focused on single name credits and distressed debt.

Both fundamental public equity and private equity managers are often asking similar questions regardless of the fact that they typically have varied holding periods and influence or insight into a company. Namely "can alternative data help us with the due diligence of a company? can it support risk management of an existing fundamental process? can it provide us with insight that we don't get as a shareholder? can it provide us with insight that we don't get as non-executive directors? does alternative data assist me with my fiduciary duties as a director? The answer to these is almost always "yes".

Private equity firms are beginning to explore alternative data for longer duration investment decisions, picking inflection points in and monitoring of existing positions. Examples of data used by these firms are: marine data, demographic information, email receipts, online search, ETF flows, geocoded payment transactions, mobility data and social media data.

[13] Source: 'The Evolution of Active Investing. Finding Big Alpha in Big Data', July 2015.

The important starting point is knowing what the relevant research question is. Typically, analysts have a set of concerns, or pressure points, that are key for the success of their investment. These concerns are translated into KPIs and alternative data sources can be leveraged for more timely information or for unique perspectives and insights. Sometimes these questions come up as part of an acquisition process but often it relates to ongoing portfolio monitoring. Precise questions that our Bespoke Projects team can answer for private equity investors include:

- How is this consumer company products performing versus peers in terms of brand perception or pricing? Relevant data categories include: web crawled data, ratings, reviews and complaints and social media data.
- How has this specific manufacturing facility performing in terms of activity levels? Relevant data categories include: satellite and sensor data.
- Are the exports increasing for this company's products? Relevant data categories include: shipping data.
- What is employee morale within the business? Has employee turnover increased? Relevant data categories include: employment data.
- What are the demographics of the customer base and how are they evolving? Relevant data categories include: consumer transaction, social media and survey data.
- How is this large retail unit performing? Relevant data categories include: geo-location data.

Section 4.4: Applications by Type of Asset Management Firm

In this section we outline examples of applications for quantitative funds, discretionary hedge funds and traditional fundamental asset managers.

Quantitative Funds

Many of the early adopters of alternative data were the larger quantitative funds. The use cases are varied - at JPMorgan's 2017 quantitative conference 237 investors were asked "How do you plan to use Big Data and Machine Learning in investing?"

- 30% said to enhance existing risk premia / quant strategies.
- 20% said to build new risk premia / quant strategies.
- 25% said to enhance portfolio construction / risk management.
- 17% said to make discretionary calls / trades.
- 8% said do not plan to use.

In a July 2017 paper entitled "Discover the Hidden World of Alternative Data" WorldQuant stated: "for investors, especially those who play in the world of quantitative finance, increasing access to alternative data from emerging sources like IoT could give rise to a variety of new investment ideas and trading strategies". Below we outline specific applications, based on public information, of WorldQuant and Blackrock SAE.

WorldQuant[14]:

- Satellites "can keep track of the number of factories under construction in rural China, a possible indicator of the strength of that country's industrial production".

[14] Source: 'Discovering the Hidden World of Alternative Data', July 2017.

- "Social media has been a particularly rich source for analysis by investors who are trying to understand market sentiment and predict business performance".

- Sentiment "analysis is useful in determining new-product perception and brand reputation, assisting investors in forming predictions about the growth of a company or industry".

- "Real-time access to data on ship movements can prove useful in building a comprehensive picture of global shipping and an accurate understanding of competitor and market trends".

- "Acquiring diverse shipping data can be used to better understand the costs and health of companies' supply chains".

Blackrock SAE[15]:

- "One area of research that has proved rewarding is measuring the impact of employee sentiment on company profitability…today by using big data analysis and crowdsourcing research, we can not only measure employee sentiment on a much wider scale, but also in a more timely and accurate way. To do this we automated 'web-scraping' capabilities to look into job sites where employees provide feedback on thousands of employers around the world, as well as other sources of employee sentiment including social media, blogs and chat rooms. When all of the relevant data from these sources has been collected an overall sentiment score for a company is calculated and this score can be regularly updated, and changes monitored".

- "Rather than focusing on a 70-year-old monthly survey of 500 consumers (Michigan Consumer Sentiment Index), we think it is possible to achieve more accurate results with a more modern toolkit and a much larger sample size…Most consumers now precede big purchases with some research on the internet. By tracking search activity, around the world, or big-ticket items, we can gain a more complete picture of consumers' purchasing intentions, which also includes granular data on the types of items that are likely to be purchased in different geographic regions. The results of this type of analysis appear promising".

Discretionary Hedge Funds

According to the Barclays[16] June 2017 survey regarding alternative data, 24% of discretionary hedge funds use alternative data with the most popular categories being: 1) consumer credit card; 2) internet / social media; 3) weather; and 4) satellite.

Increasingly, we observe discretionary hedge funds using alternative data for stock due diligence and for testing negative or "short" based investment theses. For example, is the management's representation of the business accurate or if sentiment around the business brand and products is shifting. In some instances, this has been applied in regions where corporate disclosures and corporate governance is weak. This is applicable to both public and private markets.

Below we give three examples of hedge funds that have publicly disclosed that they use alternative data:

- **Third Point**[17]: in its 2016 year-end investor letter stated: "we have added data science to our toolkit for identifying interesting, uncorrelated opportunities".

- **Point 72**[18]: at a CB Insights conference in June 2016 the Chief Market Intelligence officer of Point72, Matthew Granade, stated that alternative data is useful for generating alpha. He said: "it is a real change from how investing used to

[15] Source: 'The Evolution of Active Investing. Finding Big Alpha in Big Data', July 2015.

[16] Source: 'Rise of the machines', June 2017.

[17] Source: 'Daniel Loeb's Third Point Hedge Fund 4th Quarter Commentary', February 2017.

[18] Source: 'The New Alpha: How Alternative Data is Going to Change Institutional Investing', June 2016.

work…if you want to understand what is going on with McDonald's, you are going to have to look at credit card transactions data, you are going to look at geo-location data, at app downloads and a handful of other things. And suddenly you are going to have a very robust picture of how McDonald's is doing, and you are not going to have to talk to McDonald's about that".

- **Citadel**[19]: the firm has built significant data capabilities. Ken Griffin, founder and CEO of Citadel, stated "our ability to leverage big data effectively in our investment process is critical to our success as a firm".

Traditional Fundamental Asset Managers

The Greenwich Q3 2018 survey noted in Section 2.3 asked investment managers how alternative data is applied to their research process. Of the 14 respondents to that question:

- 71% apply the data to their fundamental research approach.
- 43% use it to help identify new investment opportunities.
- 21% rely on it to aid in trading execution once an investment decision has been made.

Based on Eagle Alpha's dialogue with traditional fundamental asset managers, the five most common applications of alternative data are (in no particular order):

1. Identification of consumer trends and preferences: new products (Samsung S9), geographic expansion (e.g. Monster Beverages, Netflix), brand strength, and customer demographics.
2. Assess corporate quality from employee and customer reviews, social media commentary and government complaints.
3. Monitor industry competition: pricing, promotions and capital investments.
4. Evaluate corporate execution via website changes, store growth, employment data, and trade data.
5. Gauge pace of secular industry trends e.g. the electronification of automobiles and real-time payments.

Below we give three examples of traditional fundamental asset managers that have publicly disclosed that they use alternative data:

- **Schroders**[20]: in its 2015 annual report, Schroders stated that "analysis of 'big data' could become a key differentiator…this year we set up a Data Insights team, representing a significant new initiative for the Group. The team is focused on developments in data analytics for investment and research, to enhance and complement the existing skills of our fund managers and analysts". The report went on to say "the quantity of information available for investment research purposes is increasing at such a rate that traditional industry practices and skillsets are unable to absorb and process it. Global trends in digitalisation, social media, open data and technology are all creating vast streams of alternative data that are often highly unstructured and extremely obscure. However, they contain valuable and often unique insights".

- **State Street**: in an August 2016 Institutional Investor article[21] entitled "Unexpected risk meets unexpected data," State Street's Chairman and CEO highlighted a few ways investors can use different sources of information to enhance portfolio transparency and identify risk exposure ahead of potential black swan events e.g. "Online retail. When consumers order products, they may be helping investors better track inflation trends to help recalibrate investment strategies before, and after, an event. PriceStats, an inflation series built by State Street Global Markets on online data, uses technology to monitor price fluctuations on roughly 5 million items and tends to identify price shocks faster than

[19] Source: 'Uber's Laszlo Korsos Joins Citadel as Chief Data Officer', July 2017.

[20] Source: '2015 Markets in Review', March 2016.

[21] Source: 'Unexpected Risk Meets Unexpected Data', August 2016.

similar measures of offline prices, helping investors quickly understand potential shifts in inflation in more than 70 countries".

- **NN Investment Partners**[22]: In an interview in May 2017 a senior portfolio manager (Mark Robertson) told Fund Selector Asia that "opinions and emotions expressed in online content, from news articles, through blogs, forum posts, social media such as Facebook or LinkedIn, to tweets, can provide a sense of market sentiment that can reinforce or even anticipate fundamental indicators, thereby helping make investment decisions". Mark went on to say that "the big data we've incorporated is very good at picking out turning points and extremes". The article included a case study:
    - In early 2017, the team's scorecards were generally positive on commodity markets, despite some signals of credit tightening in China. As part of its sentiment analysis, the team monitored political risk and emotional sentiment indices around commodities.
    - "Around mid-April we saw a really sharp deterioration in both of these signals in our scorecards, at the same time the overall scorecard was still giving us a positive view".
    - The sentiment shift reinforced the misgivings the team already had, based on their fundamental analysis, and a decision was made to reduce commodity exposure. The Bloomberg Commodity Index subsequently fell from 86.3 on 13 April to 82.0 on 9 May 2017.
- **Neuberger Berman**[23]**:** In a January 2018 article, Neuberger Berman's Chief Data Scientist Michael Reece shared his thoughts on the adoption of alternative data sources by portfolio managers. In Mr. Reece's opinion, big data is useful for fundamental equity analysis as it provides "a new level of precision regarding what is actually happening on the ground to a business". He highlighted that credit card data and online pricing data can be aggregated, for example, and used by analysts to understand how a company is doing in between its earnings periods.
- **Goldman Sachs Asset Management**[24]**:** In the first edition of Quantinomics, it is highlighted that GSAM receives credit card data on a monthly basis with a six-day lag giving it a potential informational advantage compared to only using quarterly corporate earnings announcements. GSAM demonstrated its analysis of average credit card transaction sizes and found that growth in average sales is a good indication of future improvements in profitability. It was noted that "this data may help predict sales growth for up to a year in the future due to persistence in the growth of ticket size."
- **UBS Asset Management**[25]**:** In a November 2018 article, it is highlighted that "UBS is leveraging card payment information to monitor sales data against earnings estimates and potential share price impacts". It is mentioned that "alternative data signals help asset managers minimize risk while ensuring the delivery of superior investment performance on behalf of their clients."

[22] Source: 'NN IP Gains From Sentiment Analysis', May 2017.

[23] Source: 'Asset Matters: Big Data – The Future Is Here Already', January 2018.

[24] Source: 'Credit Card Data', April 2018.

[25] Source: 'UBS Asset Management Taps Alternative Data to Increase Alpha', November 2018.

Section 4.5: Buyside Use Cases Mentioned on Altdata.tv

On September 2018, Eagle Alpha launched altdata.tv which gives viewers access to the views of thought leaders in the alternative data space. See https://altdata.tv/

Some of the alternative data use cases mentioned during the interviews are highlighted below.

Mani Mahjouri, CEO/CIO at Blueshift Asset Management

- "We live in an age where we have new data sources. Lots and lots of new data sources. They enable us to kind of think about the world and measure things that we were never able to measure before. That gets coupled with, you know, advances in computing power, so we can ask more questions that are more difficult from the computational perspective. And advances in computational techniques as well, so we can mimic more familiar patterns of learning through Artificial Intelligence. That's really leading to sort of a real renaissance in quantitative thinking and our ability to measure effects from the environment, simulate those into signals that help us understand how information can be transmitted and effectively make its way into securities prices."

- "Let's say we want to be very comprehensive, so we can build a system where we monitor some key Walmart stores across the nation, measure growth and consumer appetite. If we want to be as precise as possible, how accurate can we make that measurement? Walmart has about 4,177 stores. Let's say you put somebody in front of each one of those stores and let's say we charge them $7.25 an hour, which I think is roughly the national average minimum wage. Let's say that we put them in front of stores eight hours a day and we do that for 365 days a year. So, you multiply basically every number I just said and get about $88 million. So, for 88 million dollars a year we can have a very finely tuned measurement of how many people are going into and out of a store like Walmart."

- "A store like Walmart generates a large portion of revenue from in-store purchases and this can very valuable information if, for instance, they're not doing so well relative to market expectations and we are able to pick this up in sort of our unique data measurement format. On the flip side, we could spend this money and find out that the market's actually accurate and there's not much of a surprise to be had. So, you kind of need that combination of deeply grained measurement with a sort of a mismatch between expectations and what's actually happening. Fortunately, that does happen a lot in today's world, but that's $88 million and it might not pay off."

- "I was looking into headlines the other day and we are down to about $5 million to put a payload into space and that payload could include numerous satellites. I'm not an expert in how much it costs to put things into space, but I know that with those satellites in space orbits can be tuned to measure cars in parking lots all across the world. So, for the fraction of the cost we can monitor the entire economy from space. That's alternative data.

Where does computing and artificial intelligence come in? Well, we use computer vision to study those parking lots and computers are pretty good at counting the number of cars in the parking lot."

- "This has been developing through time and it's an interesting process to follow. We are hitting (going back to that Walmart example) that inflection point where it actually is feasible to put this together. You have a confluence of consumers of this data, folks on the buyside recognizing its value, with the right entrepreneurs looking to develop the products. The VC community seized the value proposition."

- "Ultimately, the sort of mega trend is this curve technology that's enabling us to measure information more effectively and more comprehensively. That's something that we put together with sort of just timeliness, we're in an economy where there are investments in new businesses and a new way of thinking. So, this kind of confluence that makes this moment in time fairly special."

- "There is interaction between these datasets. It's all about building the model that describes how the world works. Some of these datasets are disparate and sparse. We don't have continuous and constant sets of data. Everything that's like that, let's put it this way, has already been studied for quite a while. So, a lot of the stuff that's new and coming out is sort of more bespoke and could land itself much more readily to discretionary analysis."

Olga Kokareva, Data Sourcing and Strategy at Quantstellation

- "It's important to understand that usage of alternative data by fundamental hedge fund managers and by quants are two very different processes. Fundamental hedge fund managers normally use alternative data to reinforce their investment thesis that they derived from their regular research process. For example, a manager can hold a long position in a retailer and they are thinking about closing it, but they are not sure. So, instead of waiting for the next quarterly report they can start looking at foot traffic data or credit card data. If the sales numbers are indeed going down, they might consider and close this position earlier."

- "With quants it's a different story. Quants derive their investment thesis by analyzing data and they've been doing it for years. So that was a natural extension of the quantitative investment process – to start analyzing alternative datasets in exact same manner that they've been analyzing traditional financial data. For that reason quants became early adopters of alternative datasets."

- "I think at this early stage it is fair to say that fundamental hedge funds are behind because it's taking them some time to implement data-driven research into their investment process. But with all these new tools like dashboards and ready-to-use statistcs that vendors create speifically for fundamental hedge funds, they are catching up and they will get there pretty soon."

- "As alternative data space evolves, both buyside and vendors are trying to come up with more creative ways to extract alpha from the data. One of the ways to do so is combining different datasets to create some sort of hybrid products. For example, you can overlay news sentiment with supply chain data. As the industry moves towards consolidation and we see a lot of partnership deals, I think these vendors are very well positioned to create this type of hybrid products."

- "Another interesting development that we see when it comes to use cases is derivative or second order data. The usage of alternative data by investment participants actually by itsels creates another layer of data. A good example of what I'm talking about is EDGAR Log File data which tracks internet search of the SEC database."

- "We wouldn't be surprised to see in that in the foreseeable future all the commercial vendors will start creating derivative datasets at scale. It can actually help them to refresh their data offering, especially if they're concerned that their data is getting commoditized."

Dmitri Pekker, ex Managing Director and Head of Alternative Data at Och-Ziff

- "I think that quant funds, especially some of the prominent ones, have been using alternative data for a long time. Probably before it was known as alternative data. Fundamental funds are coming to terms with that as well and starting to use it."

- "Individual analysts have been using it in kind of simple ways. I recently saw a survey that said that Excel is the main tool to analyze alternative data and it's not the best tool. So, fundamental funds are coming onboard, but it's still slow."

- "I think that fundamental funds have a fundamentally different case for using alternative data. They're not trying to generate a signal. They look at how things are in the larger picture, so what alternative data gives them is better inputs into their models. Any kind of fundamental analysis involves a prediction. If you improve the quality of input data or you can replace some of the assumptions that they make with the actual data, you improve the quality of the model and of the output."

- "I think there's room for quantitative strategies and for fundamental strategies. They might come close together. You can start translating some of the insights that fundamental analysts have into quantitative data and input it in quantitative models and vice versa. There might be some mashing of the two, but there's definitely room for both."

- "Another argument that I've heard is that alternative data gives you some short-term prediction rather than long-term prediction. If it's a short-term prediction and you start building a very large position (like fundamental funds typically do). By the time you build, it can be too late. But I think there's still value if you know something better, or you know something earlier."

Michael Beal, CEO at Data Capital Management

- "One thing that I'm very excited about is the consolidation of some of the alternative data players. There has been a lot of the scattering and now it started to be conslidated in some of the larger distribution channels. I'm seeing more and more capital coming into this space."

- "I want every single piece of information that exists, that a human can have, in a machine readable format. So I'm excited about this consolidation that will lead to bigger capex budgets that are coming around and new types of datasets that are being explored."

- "The second thing that I'm seeing here is that probably in about a year we are going to start to see a shake-out which will be the very first time. Right now, anybody who says they're doing something alternative data related can get some funding, can get some clients despite the fact they're adding nothing to the general market. That's really important cause we need the prioritization of the capital. The capital that is going to those who are not adding value should go to exploring new data sources."

- "The other thing that I'm really excited about that we are going to be moving over the next year from alternative data largerly being focused on the US market to the global focus. That's very exciting because a lot of the information exists in the US and now we're making it in a machine readable format, but there are a lot of things that you just cannot know about India or China or Africa. Outside of having McKinsey or BCG to go on the ground and do the diligence, we can't have it."

- "This all started for me because I was thinking as an investor working in China. One of the parts of the thesis was around the average height of buildings. I had to go to Tier 1 cities, then Tier 2 and Tier 3, so I started thinking wouldn't it be great if there was some satellite provider that would tell me the height of all these buildings in real time. So that's actually going to really open up a lot of the different markets and it's going to lead to some of the volatility in the emerging markets get dampened."

- "As we continue to move from just the US-focused to a real global focus, you will see the barriers to interacting with new markets really coming down. And that's really exciting."

Geoffrey Burger, CEO & Founder at Culture Capital LLC

- "Financial statement data is limited in its ability to disclose sustainable data. There's been a long history of gap data and structures around that. There's limited history with using frameworks like SASB (Sustainable Accounting Standards Board) or IR (Integrated Reporting). As these frameworks become more prevalent, you will get more of that information in financial statements. Right now, it actually lives outside the financial statements and therefore you really need other techniques for capturing this information."

- "The areas related to the S, the social aspect of ESG is very important. We are focused on culture and we think that looking at areas related to social aspects are very valuable. A lot of this information is vague, and its description has dual meaning, so we think it's very important that alternative data can really capture what's happening in that space."

- An example of this would be thinking about gender diversity. A CEO of a large global bank, I think, said it really great when asked about why gender diversity is important to his strategy. He said, 'why would you engage only half the talent pool to find the best talent and why would you only engage half your customers to build the best product for them'. So, I think when you have someone at that high level saying something like that, that's a very powerful message. It's a very robust information that lives outside the financial statements."

- "A lot of companies are building large data teams to really understand their own messaging around communication, so they can better engage their customers, their employees, the government agencies. Companies have found out the hard way that if you're if you're not actively messaging, then someone else will take up that space in the vacuum. It's very costly to rebuild a brand. If you have a great story, tell it. If you don't work on it. There's a lot of that information and alt data can help capture that information."

Section 5: Fifty Case Studies

In this section we present fifty case studies that are based on Eagle Alpha proprietary datasets and tools as well as third party datasets. Our goal is to show case studies relevant to a variety of asset classes and investment approaches and which use a variety of alternative data types.

The case studies are in the following order: equity, macro and credit. Within each asset class the case studies are in the following order: quantitative use cases, short-term discretionary use cases and long-term discretionary use cases. For many of the case studies, information is provided directly from the data provider and Eagle Alpha may not have backtested the data or fully audited the backtesting results provided.

Edition 5 includes ten new case studies (#1, 2, 3, 4, 11, 12, 13, 14, 38, 39).

12	Equity (Hertz)	Discretionary (short-term)	Consumer Transaction, Online Search	Eagle Alpha correctly predicted Hertz quarterly revenue using its RevCast data.	56
13	Equity (Bilibili)	Discretionary (short-term)	Consumer Transaction	Chinese consumer bank card spending used to track spending on specific videogames sold by Bilibili, allowed investors to anticipate better than expected second quarter 2018 revenues and earnings.	57
14	Equity (AMZN, WFM)	Discretionary (short-term)	Business Insights	JetTrack used its corporate aviation dataset to uncover meetings between Amazon and Whole Foods that occurred before their $13.7 billion deal was announced.	59
15	Equity	Discretionary (short-term)	Consumer Transaction	An accuracy rate of 80% was observed across ten companies when Eagle Alpha estimated quarterly revenues based on email receipt data.	61
16	Equity (JE)	Discretionary (short-term)	Consumer Transaction	Using consumer transaction data, Eagle Alpha's predictive model for Just Eat correctly pointed to stronger than expected revenues in H2 2015 and H1 2016.	62
17	Equity (SBUX)	Discretionary (short-term)	Consumer Transaction, Online Search	Eagle Alpha's predictive model for Starbucks using email receipt data accurately predicted a revenue beat in Q2 2017.	64
18	Equity (601633 CH)	Discretionary (short-term)	Pricing	In Q4 2014, The CAI (China Auto Insight) data was more accurate than consensus estimates at predicting an important inflection in revenue growth.	66
19	Equity (AMZN)	Discretionary (short-term)	Geo-location	Foot traffic data to Whole Foods was used to track the impact of price reductions.	68
20	Equity (DKS)	Discretionary (short-term)	Satellite Imagery	Satellite data analysis showed falling traffic at Dick's Sporting Goods since the start of 2017. On August 15th 2017, Dick's Sporting Goods reported worse-than-expected results.	69
21	Equity (CMG)	Discretionary (short-term)	Satellite Imagery	Satellite data analysis showed falling traffic at Chipotle Q4 2014. This analysis was ahead of the street as sell side analysts only began revising Chipotle down 3 quarters later.	70
22	Equity (CMG)	Discretionary (short-term)	Employment, Online Search	Negative growth in active job listings and Google search data were used to correctly call worsening momentum for Chipotle.	71
23	Equity (AR)	Discretionary (short-term)	Sensor	A sensor data provider observed production increases at Antero Resources and correctly anticipated management would raise company guidance.	73
24	Equity (RSH)	Discretionary (short-term)	Geo-location	Geo-location data was used to anticipate disappointing sales at RadioShack stores.	74
25	Equity (GPRO)	Discretionary (short-term)	Pricing	Online pricing data pointed to negative fundamentals for GoPro, which were reflected in subsequent results and stock price.	75
26	Equity (FIT)	Discretionary (long-term)	Pricing	Online pricing data showed improving sell-through trends for Fitbit in the first half of 2017. On August 2nd 2017, Fitbit reported better than expected results with adjusted revenue of $353.3m vs consensus estimate of $339.2m.	78

27	Equity (HUBS)	Discretionary (long-term)	Employment Online Search	Accelerating growth in active job listings and Google search data were used to correctly call improving momentum for HubSpot.	80
28	Equity (EXPE)	Discretionary (long-term)	Pricing	A web data provider accurately predicted that EXPE would miss 3Q17 room-night growth expectations, based on decelerating trends in reservation growth it started flagging in August.	82
29	Equity (SQ)	Discretionary (long-term)	Consumer Transaction	Analysis of the email receipt data for Square indicated that the growth in number of sellers has been in decline since the first quarter of 2010.	84
30	Equity (FINL)	Discretionary (long-term)	Online Search	The search indicator correctly predicted weakness in Finish Line sales ahead of earnings in December 2016.	86
31	Equity (BRBY)	Discretionary (long-term)	Online Search	Citi concluded that the short-term 1-month YoY observation crossing over the 3-month moving average YoY indicates major inflection points of same store sales growth for Burberry.	88
32	Equity (FINL, FL, NKE, UAA, ADS)	Discretionary (long-term)	Online Search	Online search data provided early indicator of weakness in sportswear sector. The industry analysis supported our case for fundamental weakness for FINL and FL and pointed to longer term fundamental issues for both companies.	90
33	Equity (TWX)	Discretionary (long-term)	Mobile App	App data showed an early indicator of a positive inflection in revenue growth for HBO, one of Time Warner's largest divisions.	92
34	Equity (ATVI)	Discretionary (long-term)	Social Media	Using social media data, we correctly highlighted that the Overwatch game was well positioned to set a new sales record for Activision Blizzard.	93
35	Equity (042670 KS, 013570 KS, 036890 KS)	Discretionary (long-term)	Trade	The South Korea real-time export data accurately tracked revenue of construction machinery companies throughout the Q1'12 – Q2'17 period.	95
36	Equity (LULU)	Discretionary (long-term)	Online Search, Social Media, Pricing	Eagle Alpha's analysis proved correct i.e. Lululemon reported sales growth of 13% YoY in Q3 2016 which was in line with our expectations.	96
37	Equity	Discretionary (long-term)	Online Reviews	Companies that get better reviews from employees post better share price performance.	98
38	Equity	Discretionary (long-term)	Online Reviews	The assessment of service and product quality of auto lenders has proven to be indicative of deeper operational issues which are then reflected in stock prices.	99
39	Macro	Quantitative	Sentiment	Analysis and scoring of central bank communications yields more accurate predictions of central bank policy versus market consensus expectations.	101
40	Macro	Discretionary (short and long-term)	Satellite Imagery	Satellite imagery is used to estimate industrial activity in China.	102

41	Macro	Discretionary (short-term)	Satellite Imagery	A relationship was identified between copper prices and estimates of copper inventories using satellite imagery data.	104
42	Macro	Discretionary (long-term)	Online Search	Eagle Alpa's US Unemployment index has a 5-year correlation of 0.9 with the US Unemployment Rate, with an out-of-sample prediction improvement of 14% compared with a baseline ARIMA model.	106
43	Macro	Discretionary (long-term)	Employment	Employment data enables more granular analysis of the labor market by sector.	107
44	Macro	Discretionary (long-term)	Trade	Statistical backtesting has proven the indicators are frequently a better predictor than street-mean estimates, while also having the advantage of a substantial lead time.	108
45	Macro	Discretionary (long-term)	Trade	The South Korea real-time export data accurately tracked overall China exports.	109
46	Macro	Discretionary (long-term)	Employment	Using payroll processor microdata improves forecast accuracy of an aggregate labor market activity measure.	110
47	Macro	Discretionary (long-term)	Sentiment	Quantitative metrics of emotional content in market narratives may complement other indicators and analysis in helping to gauge systemic risk.	111
48	Macro	Discretionary (long-term)	Trade	Trade data was used to track the surprising macro-economic turnaround for Brazil in 2016.	112
49	Credit	Discretionary (long-term)	Pricing	Online property listings data was used to research buy-to-rent investors which led to the subsequent clamp down on mortgage lending.	115
50	Credit	Discretionary (long-term)	Credit Risk	Testing showed that the probability of bankruptcy within 12 months ranges from 10% to 50% when a company gets a stress score of "1".	117

1. Equity > Quantitative > Business Insights

Key Takeaways

An analysis of a trade credit dataset shows significant informational value (alpha) as the amount of corporate trade credit outstanding, and amounts past due, correlated well with subsequent stock price performance.

Dataset

This trade credit dataset includes the monthly records of accounts receivables for firms selling to publicly traded companies in the U.S. In addition to absolute dollar amounts, it includes data on the amount of trade credit that is past due. The data is insightful as trade credit makes up roughly a quarter of corporate financing and suppliers are considered to have insight into a company's business strength. Failure to pay a supplier is a warning sign for corporate health because when a firm delays supplier payment, it risks not being able to transact with that party in the future.

Case Study

A recent study[26] examined correlations between 1) the amount of trade credit a company takes on relative to its assets and 2) the amount of trade credit past due; and 3) stock price performance. They found that companies with large amounts of trade credit and those that pay back their credit balances on time tend to have stock outperformance in the near term. To carry out the study, more than 5,700 publicly traded companies with trade credit balances were examined over the 2001-2017 time period. Monthly, the companies were split into quintiles based on their amount of trade credit relative to their assets. They were also split into terciles based on the proportion of trade credit past due (high, medium, low). The subsequent one-month stock price performance was observed for each company.

- Results show that a portfolio that purchased companies in the top quintile of trade credit to total assets, and shorted those in the bottom quintile ranking, outperformed relevant market indices by about 6-7% on an annualized basis (significant at 1% level, assuming monthly rebalancing).

- A portfolio that purchased companies with relatively larger trade credit balances and a lower proportion of past due credits, while selling companies with relatively larger trade credit balances and a higher proportion of past due credits, outperformed the market indices marginally better, roughly 6.2-7.3% annually (significant at 1% level, assuming monthly rebalancing).

- For companies with relatively low trade credit balances as a percentage of assets, timeliness of payments was a not a statistically significant factor in their stock's price performance.

- A portfolio which purchased companies with low levels of past due credits and sold those with high levels, outperformed the market by roughly 3-4%.

The researchers' study of this dataset is ongoing. They are finding that the results are much more robust when looking at the percentage past due of the major suppliers and the change in total trade credit of the major suppliers.

[26] Working paper authored by Li (yifanli@sfdu.edu) from College of Business, San Francisco State University, Lourie (blourie@exchange.uci.edu) from Paul Merage School of Business, University of California, Irvine, Ruchti (ruchti@andrew.cmu.edu) from the Tepper School of Business, Carnegie Mellon University.

2. Equity > Quantitative > Business Insights

Key Takeaways

Information in trade credit dataset, and user activity of the dataset, show alpha generating potential for U.S equities.

Dataset

This trade credit dataset includes the monthly records of accounts receivables for firms selling to publicly traded companies in the U.S. In addition to absolute dollar amounts, it

- includes data on the amount of trade credit that is past due;
- provides scores that reflect the click patterns of the users researching each company in the database;
- provides a proprietary credit score that reflects the possibility of bankruptcy in the next twelve months.

The data is insightful as trade credit makes up roughly a quarter of corporate financing and suppliers are considered to have insight into a company's business strength. Failure to pay a supplier is a warning sign for corporate health because when a firm delays supplier payment, it risks not being able to transact with that party in the future.

Case Study

The strong correlation was observed between 1) proprietary click pattern scores, 2) proprietary credit scores, and 3) stock price returns was during the 2001 to 2017 time period. The monthly click pattern score takes into account the number of dataset queries about a company, and the depth of research about that company, by the users/contributors of this dataset (suppliers of trade credit).

A large group of publicly traded companies was selected and divided into quintiles based on their percentage change in monthly click pattern score. Within each quintile, companies were categorized by their credit score; high, medium, low. It may seem counterintuitive that the best performing stocks were those of companies with increased click pattern scores and the lowest credit scores. However, increased querying and research is a sign of interest in extending credit rather than concern about credit quality.

A factor neutral portfolio which purchased stocks of low credit scoring companies with strong growth in their click pattern score and shorted low credit scoring stocks with low click pattern score growth, would have substantially outperformed the relevant market indices by roughly 12% annually (significant at the 1% level). The return is achieved by monthly resorting and rebalancing of the portfolio. The poorest performing stocks were those of companies with low credit quality and low growth in click pattern scores.

3. Equity ($RUI) > Quantitative > Sentiment Data

Key Takeaway

A systematic strategy that buys and sells stocks based on signals from sentiment scores of corporate earnings calls outperforms the Russell 1000 Index ($RUI).

Dataset

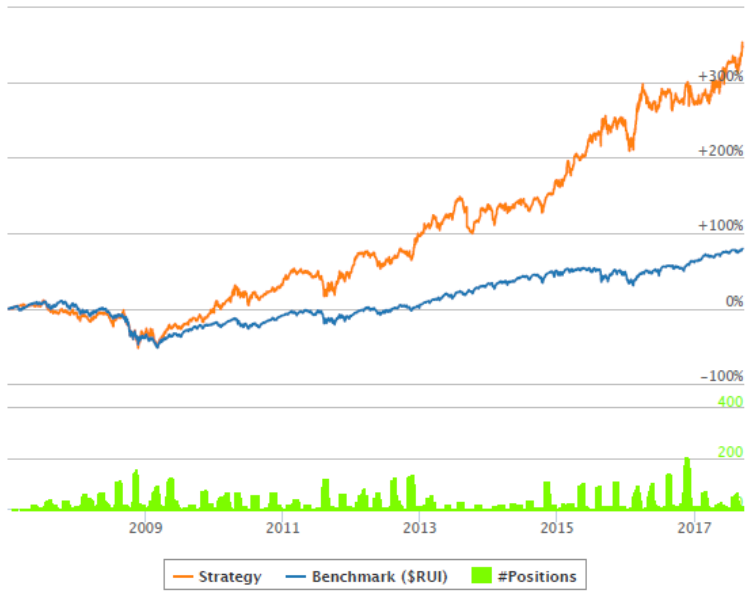
Sentiment scores are created via a proprietary machine learning algorithm that analyzes data on earnings calls from each company including the initial presentation and Q&A session with equity analysts. The score for an earnings call corresponds to the expected cumulative abnormal return in the company's stock price following the earnings call.

The long equity strategy involves purchasing stocks in the Russell 1000 Index when their earnings call score is greater than 1. Assets purchases are sold after 21 days.

Case Study

Backtest results from 1/2007-9/2017 period demonstrates robust performance against $RUI.

Figure 16: Results of Systematic Strategy Based on Sentiment Scores of Corporate Earnings Calls



Monthly Performance Metrics:

Months of Outperformance	57.36%
Most Consecutive Months of Outperformance	13
Most Consecutive Months of Underperformance	4

Benchmark Relative Metrics:

Beta	1.04
Information Ratio	0.58
R^2	0.60

Performance Metrics:

Portfolio	Return	Sharpe	Sortino	Alpha	Standard Dev	Max DrawDown
Fund	347.33%	0.65	0.86	2.64	1.73%	-57.44%
Benchmark	79.99%	0.37	0.45	0.00	1.29%	-56.88%

Source: Sentiment Data Provider, Eagle Alpha

4. Equity > Quantitative, Discretionary > Mobile App Data, Geo-location Data

Key Takeaway

Analysis of Chinese mobile phone data can be used to develop predictive series which produce stock market buy/sell signals.

Dataset

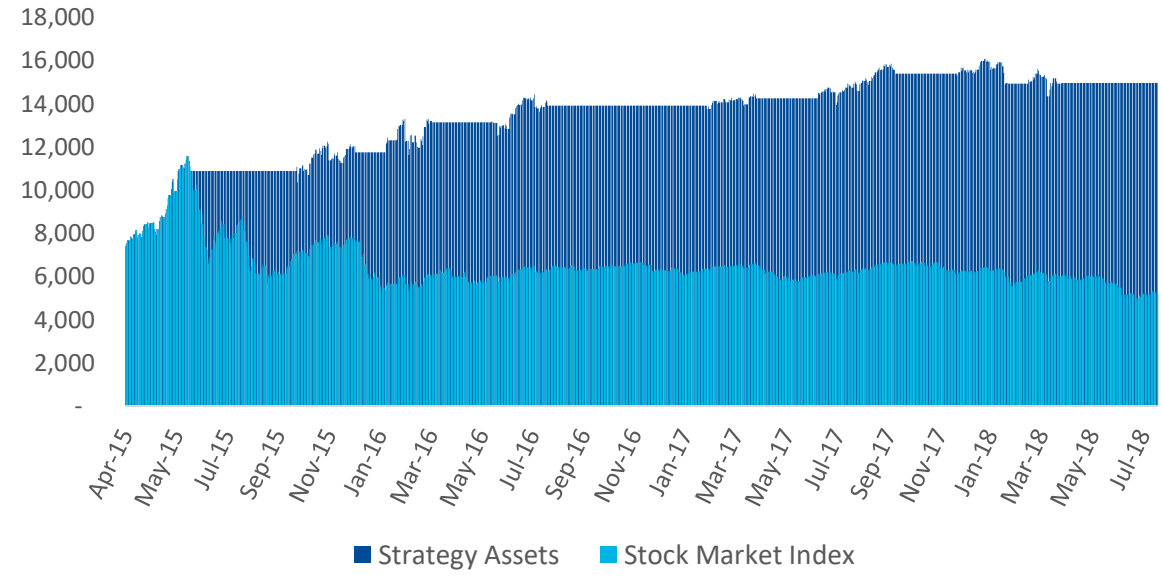
The data is derived from more than 900 million mobile phone users, mainly in China but also in the US and South East Asia. Data collected includes apps installed, frequency and duration of app usage, and geo-location data. The data has been used to track specific Chinese industries, such as steel production and consumption, as well as consumer trends.

Case Study

An asset allocation indicator based on stock market investors' mobile phone behavior produced “Risk On” and “Risk Off” signals. A strategy that was invested in a Chinese equity index ETF during “Risk On” periods and invested in money market funds from January 1, 2015 till the end of 2017 would have yielded an annual return of 20.7% with a maximum drawdown of 8.9%. The corresponding sharp ratio was 1.39.

Seen below is the growth in asset value of funds invested in the Chinese Stock Index ZZ500 and assets invested according to the signals provided (excludes income from money market and transaction costs):

Figure 17: Assets Invested in Strategy vs Stock Market Index



Source: Mobile App and Geo-location Data Provider

5. Equity > Quantitative > Employment Data

Key Takeaway

Testing shows that there is alpha in the data, with the "Jobs Active" variable producing the highest and most consistent returns (6-8% yearly).

Dataset

Completely unique in the industry, the job listing dataset only indexes jobs directly from employer websites. Updated daily with over 4 million jobs from more than 30,000 employers, the platform eliminates duplicate and expired job listings, as well as job pollution.

Backtesting/Significance

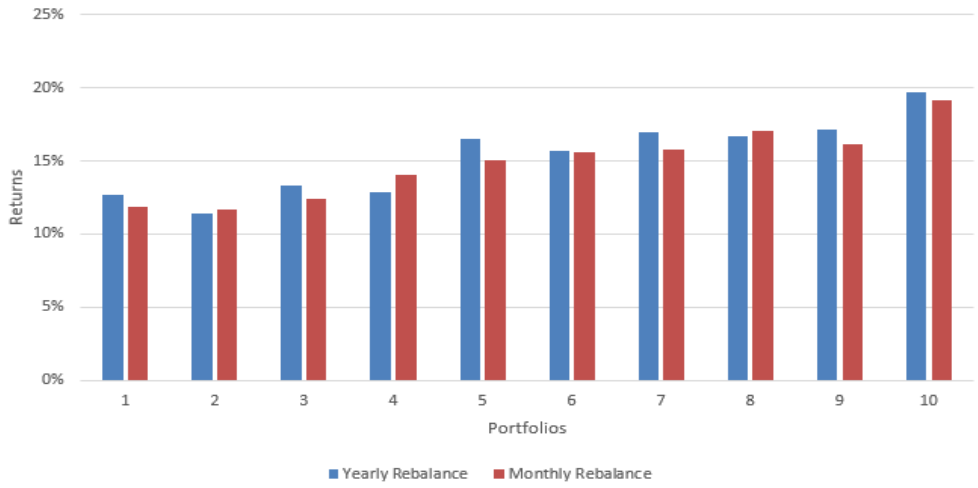
We used two variables: "Jobs Created" (number of job posting created by companies) and "Jobs Active" (number of job postings that are currently active and open). We also normalized these variables and transformed them using different change specifications. Then we deflated all variables by the firm market value of equity at the end of the year. In total, we had 10 variables.

We formed monthly and annual portfolios by dividing the sample of firms based on the 10 variables into both deciles and quintiles. The top portfolio (for deciles - portfolio 10) is the portfolio of 10% of the firms where the variable examined was the highest for a given month/year, and the bottom portfolio (portfolio 1) is the 10% of the firms where the variable is the lowest.

We then calculated each portfolio future returns after the formation period using both monthly and yearly return horizons. Finally, we calculated the hedge returns, i.e. the top portfolio average return minus the bottom portfolio average return (portfolio 10 minus portfolio 1 in the case of deciles).

The results suggest that there is alpha in the variables, with "Jobs Active" producing the highest and most consistent returns (Figure 18). Yearly hedge returns were between 6-8%. In the case of "Jobs Created", yearly hedge returns were between 2-4%.

Figure 18: Jobs Active Portfolio Returns



Source: Employment Data, Eagle Alpha

6. Equity > Quantitative > Consumer Transaction Data

Key Takeaway

Cross-sectional analysis of our partner dataset, performed by J.P. Morgan, generated annualised returns of 16.2% and a Sharpe ratio of 1.13.

Dataset

The dataset is delivered through a partnership with a provider that collects anonymized purchase data from around 2 million active shoppers, scanned from email purchase receipts. It covers over 600 merchants from more than 25 industries.

The provider transforms multiform unstructured email receipt data into a normalized digestible consumer transaction dataset. The dataset is also granular as it includes item and SKU-level transaction data, which is filtered into 53 product categories.

Geography	Coverage	Mapped to Tickers	History	Collection Frequency	Delivery Frequency	Lag Time	Delivery Method
USA	600 merchants	No	Since 2013	Real-Time	Weekly	7 days	Excel, TSV

Below we include an extract from J.P. Morgan's report "Big Data and AI Strategies. Machine Learning and Alternative Data Approach to Investing." J.P. Morgan analyzed a dataset of email receipts for 97 companies. 36 of these were private companies, and 61 public, 31 of which were S&P 500 constituents.

Taking liquidity into consideration, J.P. Morgan decided to test trading signals for the S&P 500 companies only.

Backtesting/Significance

We analyzed three time series: the dollar spend, number of orders and number of buyers. While number of orders and number of buyers are highly correlated (~99%), dollar spend is not highly correlated with number of buyers/orders (~25%).

We aggregated the daily spend/order/buyer data into a weekly score and calculated week-over-week percentage change for each. After winsorizing to 5th-95th percentile, we tested both the level and z-score as signals. Based on a cross-sectional comparison, we went long the top 5 stocks and short the bottom 5 stocks. The portfolio was rebalanced weekly.

Figure 19: Sharpe ratios of various strategies using Dollar Spend, Buyer Count and Order Count

Dollar Spend Data	Top 6/ Bottom 6
Level	0.29
Z-score 4 week	1.13
Z-score 5 week	0.72
Z-score 6 week	0.67

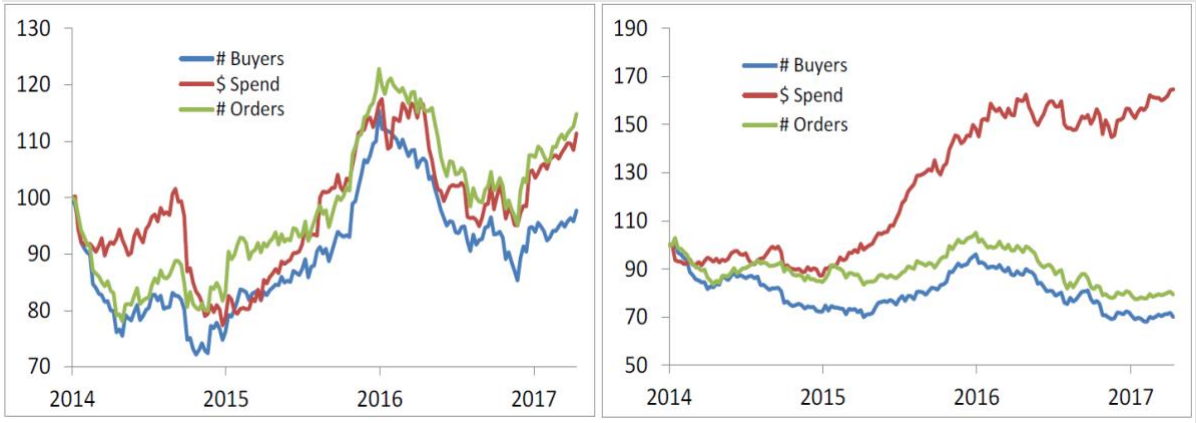
Buyer Count Data	Top 6/ Bottom 6
Level	0.02
Z-score 4 week	-0.71
Z-score 5 week	-0.49
Z-score 6 week	0.04

Order Count Data	Top 6/ Bottom 6
Level	0.36
Z-score 4 week	-0.49
Z-score 5 week	-0.14
Z-score 6 week	0.11

Source: J.P. Morgan Macro QDS, Eagle Alpha

We also plot cumulative returns using the level (i.e. percentage of aggregated figure) and the 4-week z-score for all 3 datasets.

Figure 20: Performance of level (left) and time-series z-score of changes (right) as trading signals



Source: J.P. Morgan Macro QDS, Eagle Alpha

The 4-week z-score on the Dollar Spend metric displayed an impressive sharpe ratio of 1.13 (Figure 19). Annualised returns for this same portfolio also impressed at 16.2%. This can be seen in the right-hand chart in Figure 20 above.

7. Equity > Quantitative > Sentiment Data

Key Takeaway

Both Long and Short legs of Long/Short strategy created using sentiment data from this provider contributed to the global outperformance of the portfolio. The latter also exhibits a much lower maximum drawdown than the market due to the capacity of our explaining variable (emotional agitation) to anticipate market correction.

Dataset

This provider delivers investment signals based on sentiment and emotions analysis of over 120,000 data sources, focusing on social trading websites and specialized press. Signals are available for more than 2,000 tickers as well as bonds, currencies and commodities. History: since 2014.

Backtesting/Significance

For a weekly rebalancing frequency (λ=5), the provider simulated for each week an investment in a basket of n stocks (e.g., 30 for the DAX) where the position (Long or Short). A summary of the results for the DAX stocks is presented below. The grey line represents the DAX performance, the blue line shows net returns while the orange line shows gross returns.

Figure 21: Backtesting Results for DAX Stocks

Strategies (Daily Net Return)	Return Ann.	Volatility Ann.	Sharpe	Max DD
Long/Short	12.6%	18.7%	0.67	-18.2%
Long signal only	10.0%	10.2%	0.98	-13.8%
Short signal only	1.5%	15.1%	0.10	-18.2%
DAX	10.1%	19.6%	0.51	-28.3%

Source: Sentiment Data Provider

8. Equity > Quantitative > ESG Data

Key Takeaway

Backtesting shows that a portfolio optimized using ESG data from this data vendor outperforms both the benchmark index and a portfolio optimized using traditional factors.

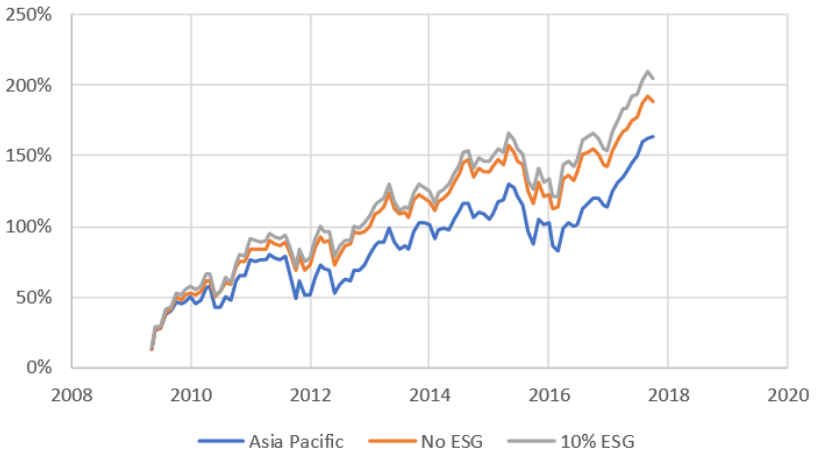
Dataset

This vendor provides a robust ESG (environmental, social, governance) dataset to institutional investors. They rate over 20,000 companies across a spectrum of ESG key performance indicators. Each KPI is represented by a score of 0-100 and can be used in various ways across investment strategies. Ratings begin as of 2009.

Case Study

Backtesting shows that a portfolio optimized using ESG data from this data vendor outperforms both the benchmark index and a portfolio optimized using traditional factors. This is true for both realized returns and shape ratios. The max drawdown for the ESG optimized portfolio is also less than the benchmark portfolios.

Figure 22: Asia Pacific Cumulative Returns



	Asia Pacific	No ESG	10% ESG
Realized Return (% Ann)	12.09%	13.30%	14.02%
Realized Risk (% Ann)	14.98%	13.38%	14.04%
Sharpe Ratio	0.81	0.99	1.00
Realized Active Return (% Ann)		1.21%	1.93%
Realized Active Risk (% Ann)		4.49%	3.97%
Information Ratio		0.27	0.49
Realized Beta	1.00	0.85	0.90
Max DrawDown	20.7%	17.6%	16.8%

Source: ESG Data Provider

9. Equity > Quantitative > ESG Data

Key Takeaway

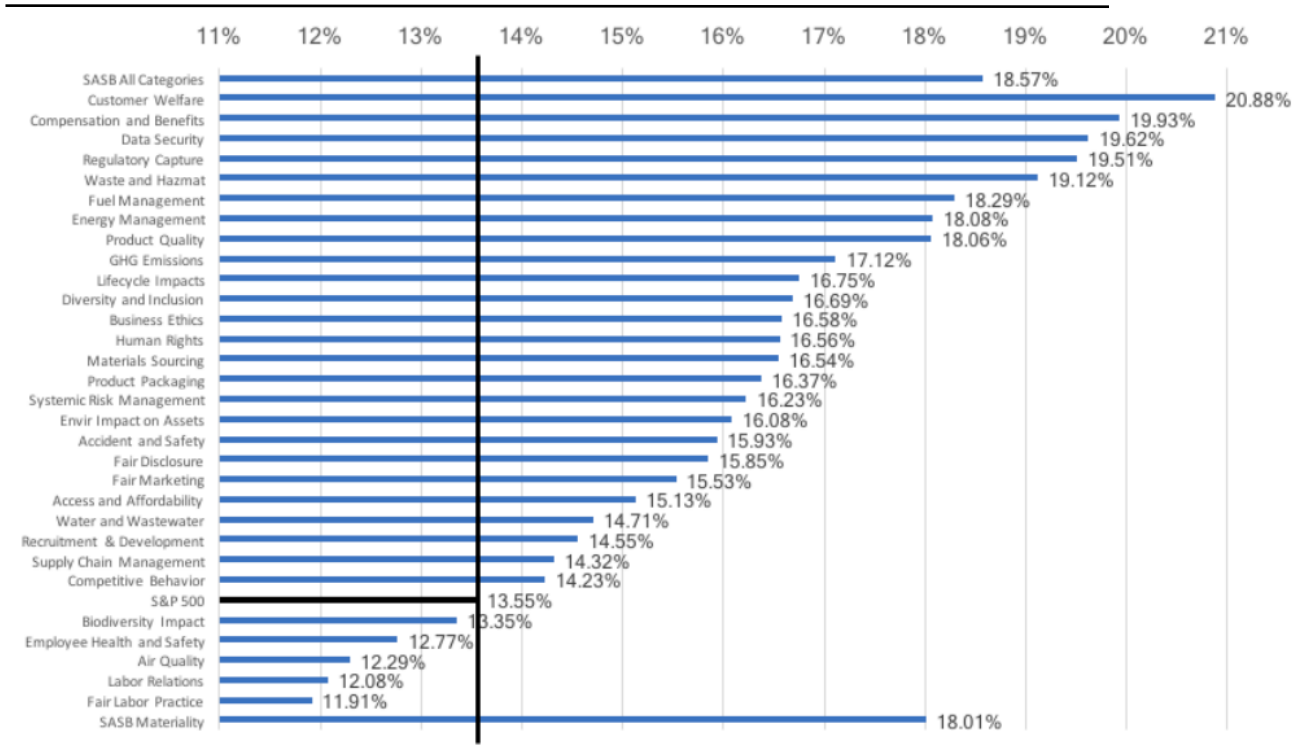
Backtesting shows that the provider's factors perform better than many traditional quant factors and are additive to a multi-factor investment approach.

Dataset

This vendor applies artificial intelligence (AI) to uncover timely ESG data on a variety of asset classes. The company delivers timely and investable insights by revealing intangible value and risk factors from unstructured text. The dataset history is September 2016.

The company adopted the framework designed by the Sustainability Accounting Standards Board (SASB). SASB has done extensive research on the sustainability categories that have material impact on companies in an industry. See Figure 23 below for more details.

Figure 23: Companies that Perform Well on the SASB Categories Outperform



Source: ESG Data Vendor

The SASB framework differentiates this vendor from other data providers. This vendor also performs daily company scores as events occur.

Case Study

The company scans tens of thousands of unstructured web sources and composes ESG scores for over 8,000 companies. Historical simulated backtesting results presented in Figure 24 show the total return from a portfolio of stocks scored by the provider. These results do not include fees, transaction costs, commissions, taxes, or any other market frictions.

The provider's insight score is a moving average of the faster-moving event-based score while the momentum score is the relative ranking of the change in the insight score over the past 12 months.

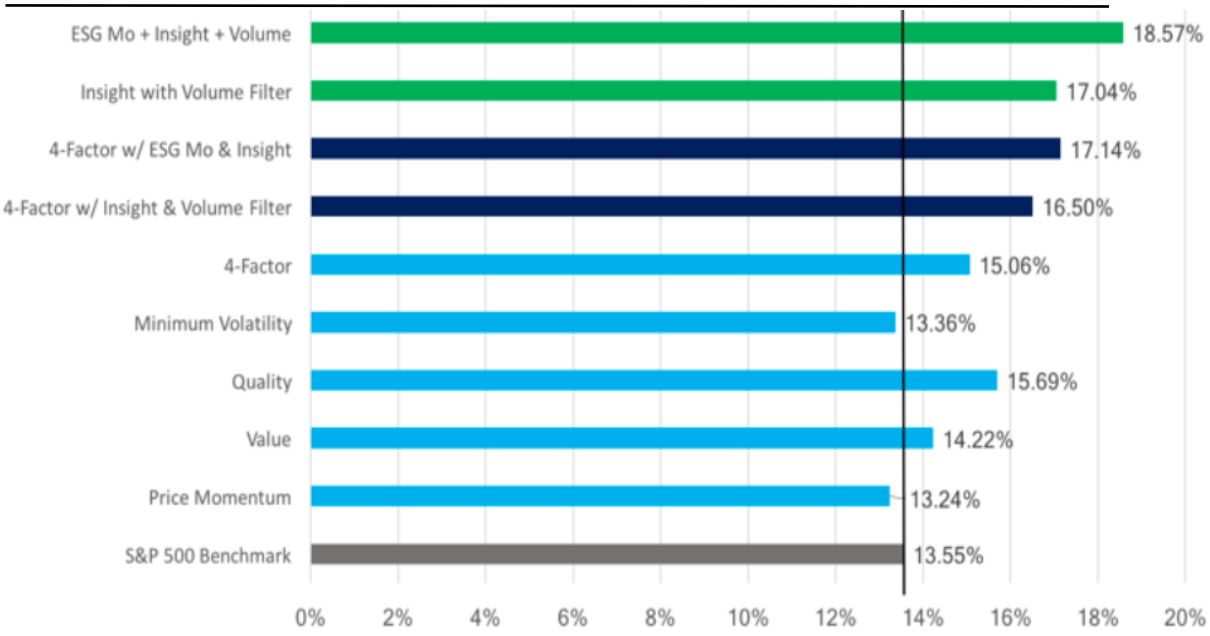
Figure 24: Cumulative Returns

		Cumulative Returns		Excess Returns	
Year	Benchmark	Insight	Momentum	Insight	Momentum
2013	27.2%	33.4%	32.6%	6.2%	5.4%
2014	14.0%	16.7%	20.1%	2.8%	6.1%
2015	-2.7%	3.2%	3.5%	5.9%	6.2%
2016	14.4%	12.3%	14.7%	-2.1%	0.3%
2017 YTD	9.4%	12.1%	13.6%	2.7%	4.2%
Since Inception	77.1%	103.0%	115.2%	25.9%	38.1%
Annualized	13.5%	17.0%	18.6%	3.5%	5.0%
Sharpe Ratio	1.00	1.29	1.36	0.29	0.36

Source: ESG Data Vendor

These scores provide a way to quantify many factors missed by traditional quantitative analyses. Backtesting shows that the provider's factors perform better than many traditional quant factors and are additive to a multi-factor investment approach.

Figure 25: Annualized Returns



Source: ESG Data Vendor

10. Equity > Quantitative > Employment Data

Key Takeaway

Testing showed that firms with lower employee turnover systematically outperformed those with higher turnover rates.

Dataset

This vendor offers a rich and unique view of the global labour force at a company, industry, and global level. The vendor maps half a billion individuals, more than 10,000 global public companies, and millions of private and non-corporate entities (government, education, military, healthcare, etc.) in order to capture workforce dynamics. History: since January 2007.

Case Study

Education and employment histories from an aggregator of public profile information for CRM system enrichment was used to gather information on more than 1.4 billion job change events (there are around 344 million profiles with basic details and 262 million profiles with employment records).

Testing showed that firms with lower employee turnover systematically outperformed those with higher turnover rates. In order to demonstrate this, the vendor computed monthly abnormal turnover rates (see the formula below), sorted firms by this characteristic and formed a long/short portfolio from the top and bottom quintiles.

$$AbnTurnover_{i,t} = Turnover_{i,t} - PredTurnover_{i,t},$$

Using a three-month lag, the long portfolio earned, on average, 1.23% monthly return, while the short portfolio earned a substantially lower return of 0.10% per month.

11. Equity (Samsung) > Discretionary (Short-Term) > Social Media, Online Search

Key Takeaway

In March 2018 Eagle Alpha highlighted lack of consumer interest in the launch of Samsung's flagship Galaxy S9 handset using data from its Web Queries tool and Google Trends.

Datasets

Eagle Alpha Web Queries – A query-based tool that enables Eagle Alpha research analysts and clients to search multiple web sources e.g. review sites, blogs, forums, videos and Facebook. Analysis of web and social media text data has proved insightful in tracking consumer attitude towards major events and long-term trends.

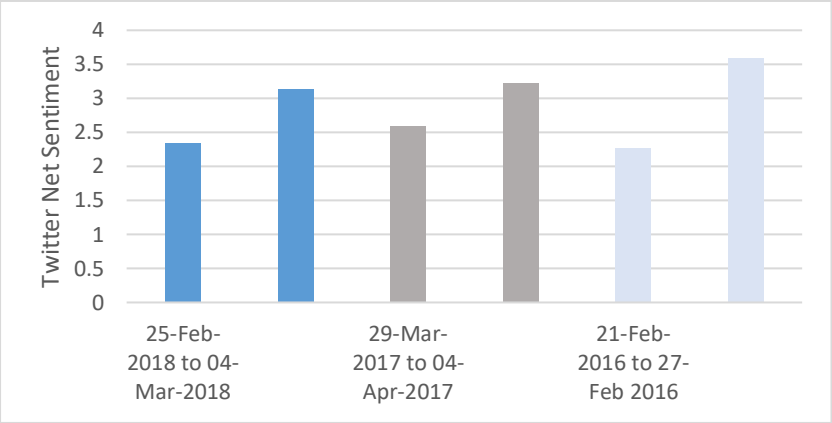
Google Trends – A public web facility based on Google Search that shows how often a particular search-term is entered relative to the total search-volume over time across various regions of the world.

Case Study

Eagle Alpha's Data Analytics team published a research note in March 2018 titled "Samsung: Early Demand for S9 lags S8". We used Eagle Alpha Web Queries tool to analyse Twitter sentiment. This pointed to weaker consumer interest for the Samsung S9 than its predecessor the S8. The lack of material upgrades to many features seemed to be contributing to this weak performance.

Over 50% of Twitter conversations focused on the camera and performance features, compared to the excitement regarding the new display around the launch of the S8. Furthermore, search interest in the S9 was closer to that of the S7 than the S8. This signalled a lower level of consumer demand for the product.

Figure 26: Device Net Sentiment



Source: Eagle Alpha Web Queries

On July 6th 2018, Samsung announced early results for the June quarter with weak profits. The weak profit outlook was put down to 'lacklustre performance of its flagship Galaxy S9 smartphone' (Financial Times). The FT article also highlighted that the S9 sales are below prior models and the lowest since the S3 in 2012.

12. Equity (HTZ) > Quantitative > Consumer Transaction, Online Search

Key Takeaway

Eagle Alpha correctly predicted Hertz 2Q 2018 revenues using its RevCast data.

Dataset

RevCast incorporates multiple datasets, as well as Eagle Alpha's extensive experience working with alternative data:

- Consumer Transaction Data[27] – transactional dataset from a prominent consumer transaction data company that is highly representative of real spend.
- Online Search Data – indices of search volumes for company specific keywords, sourced from Google Trends.
- Historic Financials – reported sales for each company.

RevCast characteristics:

- Geographic coverage: US.
- Coverage: 60 US Equities.
- History: 3-6 years.
- Collection Frequency: Weekly to Monthly.
- Delivery Frequency: Quarterly.
- Lag: 7-14 days after quarter-end.

Case Study

On July 6th 2018, Eagle Alpha's RevCast model produced a forecast number for Hertz revenues which was set at USD 2.45 billion. The consensus estimate on that day was set at USD 2.3 billion. On August 6th 2018, Hertz reported revenues of USD 2.4 billion, a 7% increase YoY.

Kathryn Marinello, CEO and president of Hertz commented: "In the U.S., our turnaround initiatives are bearing fruit as a result of effective strategies, experienced leaders, and critical investments in fleet, marketing, and our retail operations. At the same time, we're developing and testing new technology platforms with the future in mind. The successful launch of those systems in the second half of 2019 will further support our strategy to sustainably grow revenue, improve productivity and drive innovation over the long term."

[27] Data leverages a representative sample of spending across consumer payment vehicles including anonymous, aggregate transactions that are controlled for biases in the data to represent overall payments. The data is refreshed on a weekly basis and leverages tens of millions of payment transactions. The resulting data input provides aggregate sector trends for groups of large, national retailers.

13. Equity (Bilibili) > Discretionary (Short-Term) > Consumer Transaction Data

Key Takeaway

Chinese consumer transaction spending was used to track spending on Fate/Grand Order videogames sold by Bilibili, allowing investors to gauge revenue growth and understand its drivers for the newly listed company.

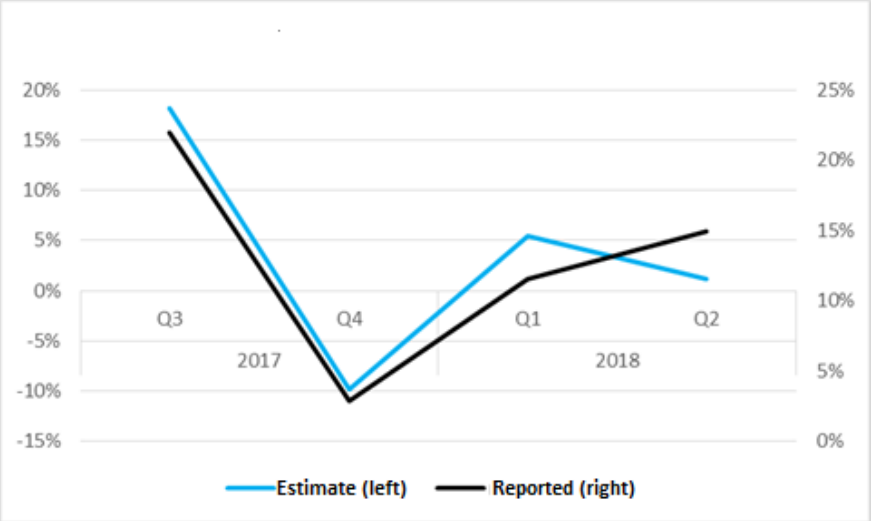
Dataset

China consumer spending data from the largest provider of Chinese credit cards.

Case Study

While consumer bank card data is well known for its usage in estimating corporate revenues, it typically does not yield insights on item level purchases. In this case, the data vendor tracks spending on a specific videogame sold by Bilibili; the anime inspired game Fate/Grand Order. The data provider can get app transaction data via iTunes if the consumer uses a bank card to pay for it. Additionally, transaction data can be filtered for the specific price of Fate/Grand Order purchases to identify some frequently-used bank cards. This data is aggregated and used to estimate game specific revenues, which can be important for analysts who want a better understanding of revenue drivers. In Figure 27 below, the black line shows quarter-over-quarter growth in Bilibili reported revenues that are attributable to the game Fate/Grand Order while the blue line represents revenues estimated by the data provider.

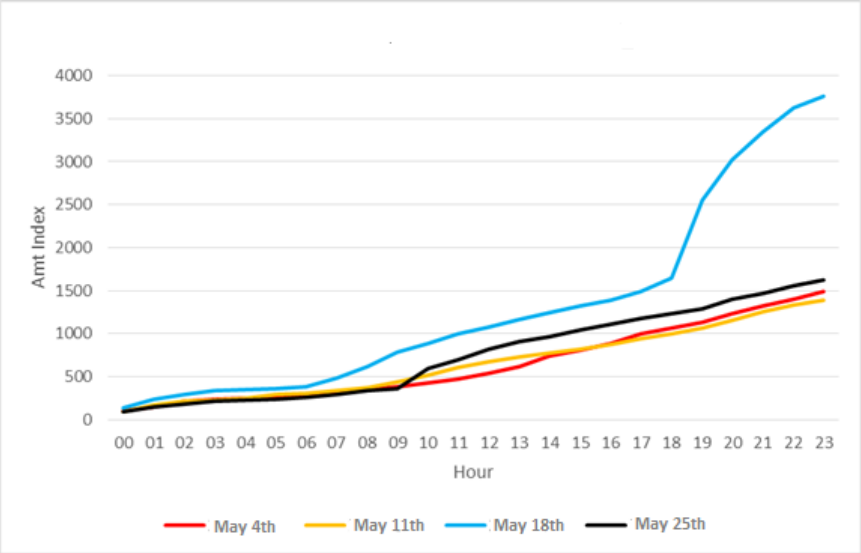
Figure 27: Estimated vs. Reported Fate/Grand Order Revenues; Quarter-Over-Quarter Growth



Source: Bank Card Data Provider

The time stamping of the data can be used to quantify the revenue impact of game events, such as launches of features and gaming competitions. Figure 28 below shows the results of tracking Fate/Grand Order in late May 2018 when there was a lot of publicity around the game.

Figure 28: Estimated Fate/Grand Order Revenues by Hour



Source: Bank Card Data Provider

Bilibili's reported revenue in Q2 2018 was roughly in line with analysts' estimates and earnings were higher than expected. The stock rallied after earnings were released, partly because of the scrutiny it is under as a newly listed company. With Chinese regulators clamping down on apps sold in Chinese app stores, this data allows investors the ability to estimate the impact on quarterly sales and evaluate whether the market is over or under-reacting to the news.

14. Equity (Amazon & Whole Foods) > Discretionary (Short-Term) > Business Insights

Key Takeaway

JetTrack used its corporate aviation dataset to uncover meetings between Amazon and Whole Foods that occurred before their $13.7 billion deal was announced.

Dataset

This dataset is provided by Eagle Alpha's exclusive partner JetTrack, the leading corporate aviation data vendor. JetTrack allows investors to identify future corporate transactions based on corporate flight activity. The output includes the frequency of the aircraft and the location. The company also has data on the location of the suppliers of a particular company. Coverage includes Russell 3000, 150 private equity firms and data on 2000 air crafts, 405,000 flights and 40,000 direct relationships between the companies.

Case Study

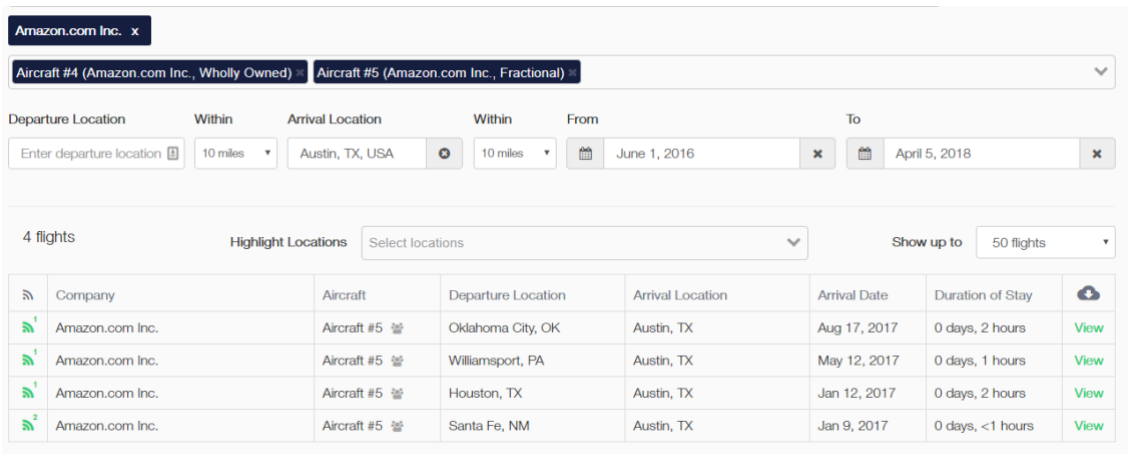
As Whole Foods' stock price plummeted in mid-April, some investors may have sought to cut their losses and leave their positions in the company. But, if savvy investors had been using JetTrack during this time, they might have seen strong indications that something big was about to happen.

If you had been an analyst using JetTrack in April 2017, you could have searched year-to-date flights by Amazon to Austin, TX — which is where Whole Foods is headquartered. That search would have shown two flights to that point:

- Monday, Jan. 9: Santa Fe, NM to Austin, TX.
- Thursday, Jan. 12: Houston, TX to Austin, TX.

If you were then to go back to June 1, 2016, you would see that these were the only flights to Austin, TX. Amazon had explored the grocery space, and activist investor Jana Partners had taken its stake in Whole Foods. There was clear reason to track both companies, and then even more reason to track the two companies after Amazon made a return visit to Austin on Friday, May 12 (see Figure 29).

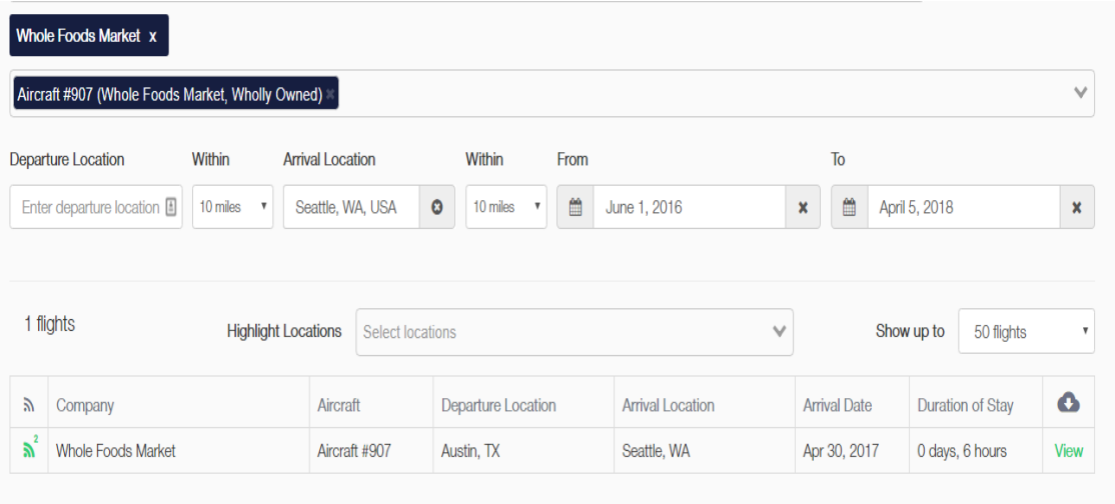
Figure 29: Amazon Flying to Austin, TX



Source: JetTrack

Using JetTrack, you could have also searched for a reciprocal flight by Whole Foods to Seattle, where Amazon is headquartered. And, when searching for flights by Whole Foods to Seattle, you would have seen there was only one flight between June 1, 2016 and the closing of the acquisition on June 16, 2017. That flight by Whole Foods to Seattle took place on April 30, 2017 — just two weeks after news of Jana Partners' stake in Whole Foods entered the mainstream.

Figure 30: Whole Foods Flying to Seattle



Source: JetTrack

15. Equity > Discretionary (Short-Term) > Consumer Transaction Data

Key Takeaway

We have observed an accuracy of 80% across ten reports incorporating email receipt data published by Eagle Alpha's Data Analytics team. The signals also captured material stock moves on the day of earnings announcement.

Dataset

The dataset is delivered through a partnership with a provider that collects anonymized purchase data from around 2 million active shoppers, scanned from email purchase receipts. It covers over 600 merchants from more than 25 industries.

The provider transforms multi-form unstructured email receipt data into a normalized and digestible consumer transaction dataset. The dataset is also granular as it includes item and SKU-level transaction data, which is filtered into 53 product categories.

Geography	Coverage	Mapped to Tickers	History	Collection Frequency	Delivery Frequency	Lag Time	Delivery Method
USA	600 merchants	No	Since 2013	Real-Time	Weekly	7 days	Excel, TSV

Backtesting/Significance

Since the beginning of July 2017 Eagle Alpha's Data Analytics team published ten quarterly predictive reports utilizing email receipt data. In each instance Eagle Alpha built stock level predictive models using the email receipt data and then published reports highlighting whether our predictions suggested current consensus numbers appeared too high, too low or in-line. Below is an overview of the results.

Figure 31: Email Receipt Data Backtesting Results

Company	Report Date	Historic MAPE*	Eagle Alpha Prediction	Result	Stock Move on Earnings
Chipotle Mexican Grill	13-Jul-17	3.7%	Miss	Miss	-2.3%
PayPal	13-Jul-17	1.6%	Beat	Beat	2.3%
Square	13-Jul-17	1.4%	In-Line	Beat	-4.7%
Papa John's International	14-Jul-17	1.4%	Beat	Miss	9.6%
Starbucks	14-Jul-17	5.8%	Miss	Miss	-9.2%
GrubHub	16-Oct-17	2.6%	Beat	Beat	11.2%
Mindbody	17-Oct-17	1.1%	Beat	Beat	2.2%
Papa John's International	18-Oct-17	2.0%	In-Line	In-Line	-8.5%
PayPal	19-Oct-17	1.4%	Beat	Beat	5.5%
Square	19-Oct-17	2.4%	Beat	Beat	3.4%

* MAPE = Mean Absolute Percentage Error. It is the average of quarterly prediction errors for each predictive model.

Source: Eagle Alpha Analysis, Email Receipt Data

Of the ten reports, eight proved accurate, indicating a hit rate of 80%. The table above also shows there was significant alpha in the signals, as indicated by the stock move on the day of the earnings.

16. Equity (Just Eat) > Discretionary (Short-Term) > Consumer Transaction Data

Key Takeaway

Using consumer transaction data, Eagle Alpha's predictive model for Just Eat correctly pointed to stronger than expected revenues in H2 2015 and H1 2016.

Dataset

Consumer transaction data is provided by Eagle Alpha's partner and consists of anonymized panel data drawn from users of the UK's largest personal financial assistant, an app/web platform. Data is collected in near real-time, directly from bank and credit card records.

The panel consists of 230 million transactions across over 500,000 accounts and grows by an average of 5 million transactions per month. With each panel member holding an average of 4 financial accounts, across multiple providers, this consumer transaction dataset represents the UK's largest complete collection of consumer financial activity.

Geography	Coverage	Mapped to Tickers	History	Collection Frequency	Delivery Frequency	Lag Time	Delivery Method
UK	U.K. National, with regional breakdowns	Yes	Since 2012	Daily	Daily, Weekly, Fortnightly & Monthly	3 days	AWS S3, FTP

Backtesting/Significance

Eagle Alpha's Just Eat model based on U.K. consumer transaction data has shown a MAPE of 1.8% in backtesting, a large improvement on consensus which has an MAPE of 3.7%.

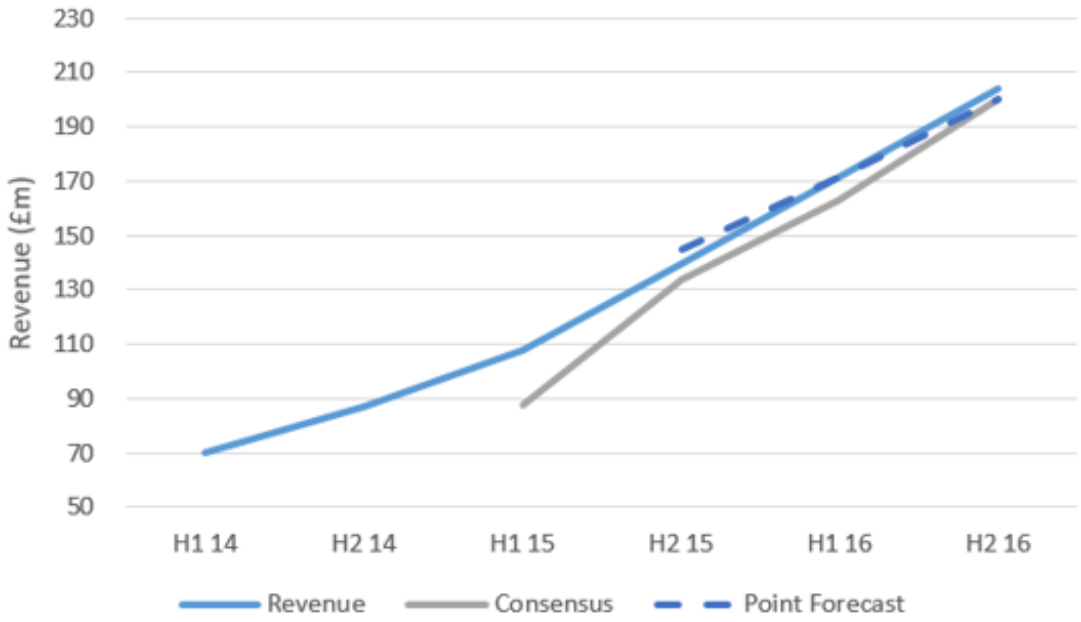
Note: testing of this dataset across 41 UK retailers found that it improved the predictive power of a baseline ARIMA model in 88% of cases.

Case Study

Using consumer transaction data, Eagle Alpha's predictive model for Just Eat correctly pointed to stronger than expected revenues in H2 2015 and H1 2016.

Our ARIMAX model forecasted H1 2016 revenues at 171.6m compared to a consensus estimate of 163m. In H2 2015 our model predicted revenues of 144.8m compared to a consensus estimate of 134m.

Figure 32: Eagle Alpha's Model Predicted Stronger Than Expected Results



Source: Eagle Alpha Analysis, Consumer Transaction Data, Bloomberg

Actual revenues for the second half of 2015 came in at 140m beating consensus, in line with our prediction. Following the H1 2016 beat the Just Eat share price rose by 4.3% on the day and 17.6% on the month. Similarly, the stock rose by 3% after the H2 2015 beat.

17. Equity (Starbucks SBUX) > Discretionary (Short-Term) > Consumer Transaction

Key Takeaway

Eagle Alpha's predictive model for Starbucks using email receipt data accurately predicted a revenue beat in Q2 2017.

Datasets

The dataset is delivered through a partnership with a provider that collects anonymized purchase data from around 2 million active shoppers, scanned from email purchase receipts. It covers over 600 merchants from more than 25 industries.

The provider transforms multi-form unstructured email receipt data into a normalized and digestible consumer transaction dataset. The dataset is also granular as it includes item and SKU-level transaction data, which is filtered into 53 product categories.

Geography	Coverage	Mapped to Tickers	History	Collection Frequency	Delivery Frequency	Lag Time	Delivery Method
USA	600 merchants	No	Since 2013	Real-Time	Weekly	7 days	Excel, TSV

Google Trends is a public web facility based on Google Search that shows how often a particular search term is entered relative to the total search-volume over time across various regions of the world. Using Google Trends, Eagle Alpha has built company specific indices based on search terms that are related to a given retailer's product offering. This involves an exhaustive process for identifying search terms related to a company's revenues using both internal and third party tools.

Backtesting/Significance

Email Receipt Data

Our Starbucks model, based on email receipt data, has a MAPE of 5.8% and the standard deviation of our error is 6.9%. Note: broader testing of the email receipt dataset across 66 US retailers found that it improved the predictive power of a baseline ARIMA model in 60% of cases.

Over the last 19 quarters our Starbucks search index based on Google search volumes has demonstrated a two-quarter hit rate of 68% in in-sample testing, and a three quarter hit rate of 53%. This means that 68% of the time the three-month moving average has moved in the same direction as same store sales over a two-quarter period.

Online Search Data

Note: Eagle Alpha's Data Analytics team has published 49 quarterly reports for consumer companies incorporating Google search data. 37 of these indicators proved accurate equating to a hit rate of 76%.

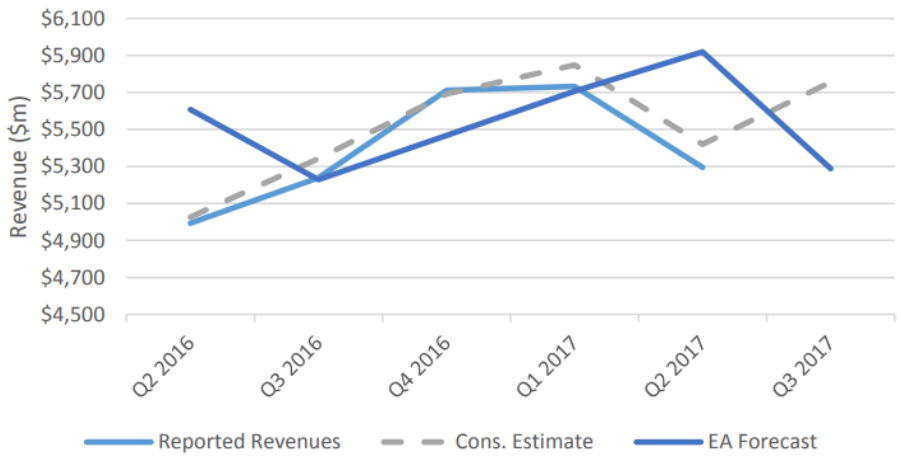
Case Study

Using email receipt and Google search data, Eagle Alpha published a research report on July 14th 2017 predicting a weaker June quarter for Starbucks versus consensus.

Figure 33 below shows that according to Eagle Alpha's predictive model Bloomberg estimates for Starbucks FYQ3 were too high. Consensus revenue estimates were at $5.75bn as analysts were predicting a YoY revenue growth of 9.9%.

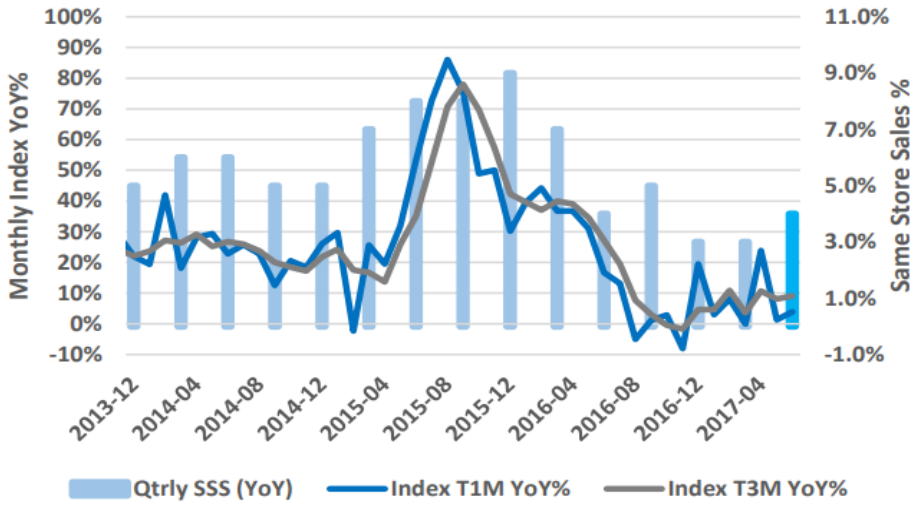
Our Starbucks search index (Figure 34) was in line with the previous two quarters, calling into question consensus expectations for SSS to accelerate later in 2017.

Figure 33: Eagle Alpha Model Predicted June Quarter Downside for Starbucks



Source: Eagle Alpha Analysis, Email Receipt Data, Bloomberg

Figure 34: Search Data Suggested Stagnant Quarterly Growth



Source: Eagle Alpha Analysis, Email Receipt Data, Bloomberg

On July 28th 2017, Starbucks reported weaker-than-expected results with revenues of $5.66bn versus consensus estimates of $5.75bn. The company lowered its full year 2017 forecast giving weak same store sales guidance. Starbucks also decided to close all 379 of its Teavana stores.

18. Equity (Great Wall Motors 601633 CH) > Discretionary (Short-Term) > Pricing Data

Key Takeaway

In Q4 2014, The CAI (China Auto Insight) data was more accurate than consensus estimates at predicting an important inflection in revenue growth.

Dataset

Eagle Alpha's CAI dataset is provided through a partnership agreement with a leading Chinese financial automotive consultant. The dataset is collected using a large panel of dealerships throughout China, combining other data sources such as web data and more traditional data sets to create a large and well-structured database.

Geography	Coverage	Mapped to Tickers	History	Collection Frequency	Delivery Frequency	Lag Time	Delivery Method
China	National, with regional breakdowns	No	Since 2012	Mixed – Month/ Bi-Monthly	Mixed – Month/ Bi-Monthly	Between 5 & 20 days	API, CSV

Backtesting/Significance

The CAI data shows a 99% correlation with revenues for domestic Chinese auto manufacturer Great Wall Motors (601633 CH) over a 5-year period, and a 95% correlation with YoY revenue growth. The calculated dealership revenue also correctly projected the directional movement of reported revenues for Great Wall in fourteen of the fifteen quarters between Q1 2013 and Q3 2016.

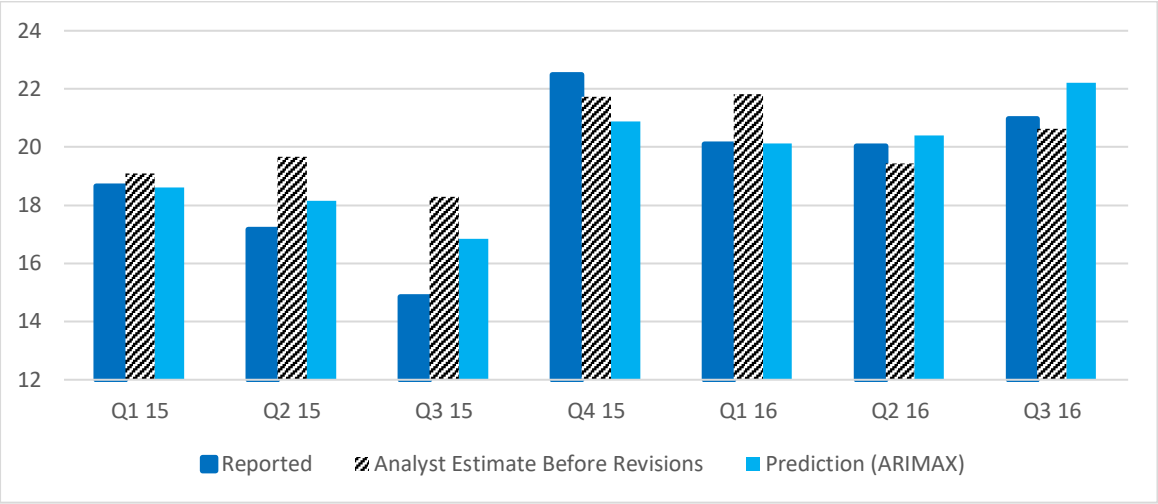
An ARIMAX[28] model for predicting Great Wall Motors (601633 CH) revenue demonstrated an out-of-sample MAPE of 4.9%. This compares to a MAPE of 10.3% on a baseline ARIMA model using just historic revenues. The error rate for market consensus estimates was 8.1% over the same period. Directional accuracy is also markedly improved over the baseline model, increasing from 57% to 86%.

Case Study

The 4.9% error rate discussed above is better than the error rate for market consensus estimates of 8.1% over the same period. This is shown in Figure 35 below. The analyst estimates in this calculation were taken one month after the end of the quarter, as the CAI data is published between 5 and 20 days after month-end. It's clear that the predictive model using CAI data provides investors with a more accurate revenue estimate for Great Wall Motors.

[28] Autoregressive integrated moving average with exogenous variables

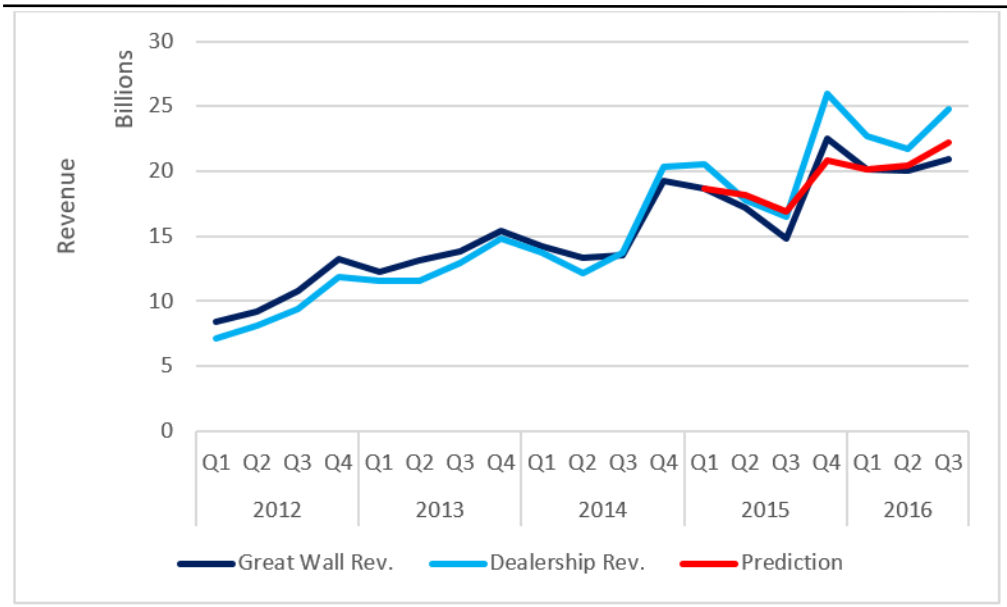
Figure 35: Reported Numbers vs Estimates vs Eagle Alpha Prediction



Source: Eagle Alpha Analysis, Bloomberg, CAI Data

In Q4 2014, the CAI data was more accurate than consensus estimates at predicting an important inflection in revenue growth. Analyst estimates were for QoQ revenue growth of 17% for Great Wall Motors, compared to reported QoQ growth of 42%. Estimated QoQ growth based on the CAI raw data was much closer at 48%.

Figure 36: Great Wall Revenue Prediction



Source: Eagle Alpha Analysis, Bloomberg, CAI Data

19. Equity (AMZN) > Discretionary (Short-Term) > Location Data

Key Takeaway

Foot traffic data to Whole Foods was used to track price reductions.

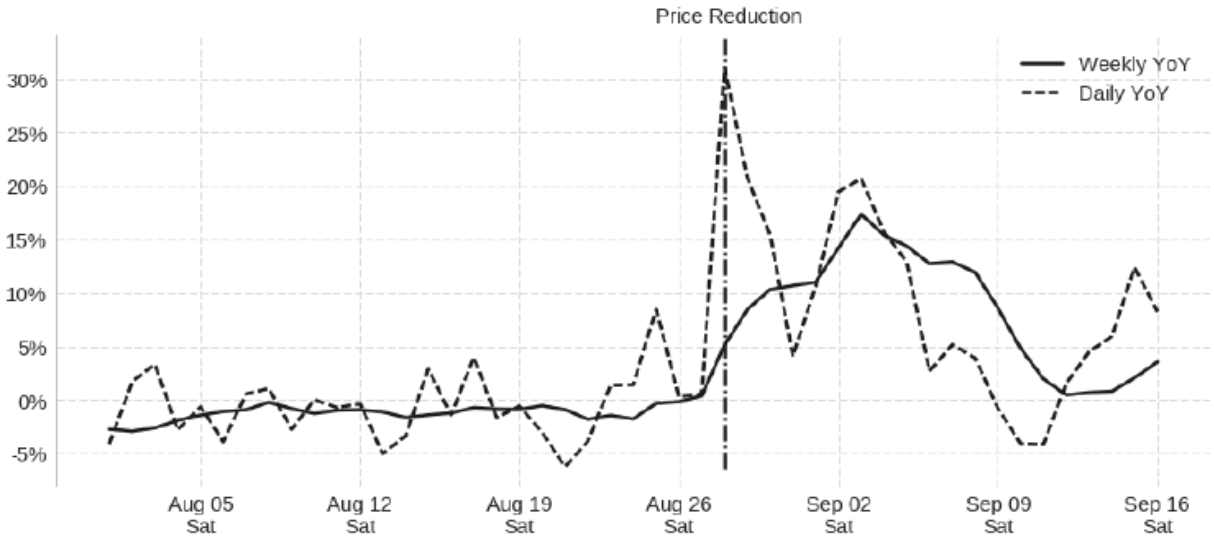
Dataset

The provider transforms real-time locations from mobile phones across the globe into objective and actionable insights on the performance of businesses, markets, and economies.

Case Study

In order to measure the change in foot traffic to Whole Foods in response to the price reduction on August 28, while also controlling for the impact of seasonality and Labor Day weekend, The figure below shows the year-over-year ("YoY") change for daily and weekly rolling windows of foot traffic. For the week ending Sunday, September 3, the weekly YoY change in foot traffic to Whole Foods peaked at 17%. On August 28, the daily YoY change in foot traffic peaked at 31%.

Figure 37: YoY Change in Foot Traffic for Whole Foods' Customers



Note: The weekly rolling window is defined as a seven-day period of time that increments or rolls forward by adding one new day and simultaneously dropping the oldest day. The rolling window is plotted in terms of the YoY change relative to the comparable seven-day period in 2016.

Source: Location Data Provider

20. Equity (Dick's Sporting Goods DKS) > Discretionary (Short-Term) > Satellite Imagery

Key Takeaway

Satellite data analysis showed falling traffic at Dick's Sporting Goods since the start of 2017. On August 15th 2017, Dick's Sporting Goods reported worse-than-expected results.

Dataset

The case study presented below was prepared by a satellite imagery data analytics company that catalogued and backtested more than one million parking lot images that accounted for 1.5 billion cars over seven years. The traffic data obtained from satellite images can then be used to analyze cumulative YoY car count growth rates and compare it to revenue growth and share price dynamics.

Case Study

Car counts at Dick's Sporting Goods have been falling since the start of 2017. There was a 7.2% YoY drop in Q1 2017 and then a further 8.4% YoY drop in Q2 2017 (Figure 38). The downward trend continued in July and August 2017 proving to be an early indicator of Q2 2017 results.

On August 15th 2017, Dick's Sporting Goods reported worse-than-expected results with same store sales rising by 0.1% which was much lower than the company's forecast of 2-3% and the consensus estimate of 1.7%.

Figure 38: DKS Cumulative YoY% Change in Car Counts

Source: Satellite Data Analytics Provider

21. Equity (Chipotle CMG) > Discretionary (Short-Term) > Satellite Imagery

Key Takeaway

Satellite data analysis showed falling traffic at Chipotle Q4 2014. This analysis was ahead of the street as sell side analysts only began revising Chipotle down 3 quarters later.

Dataset

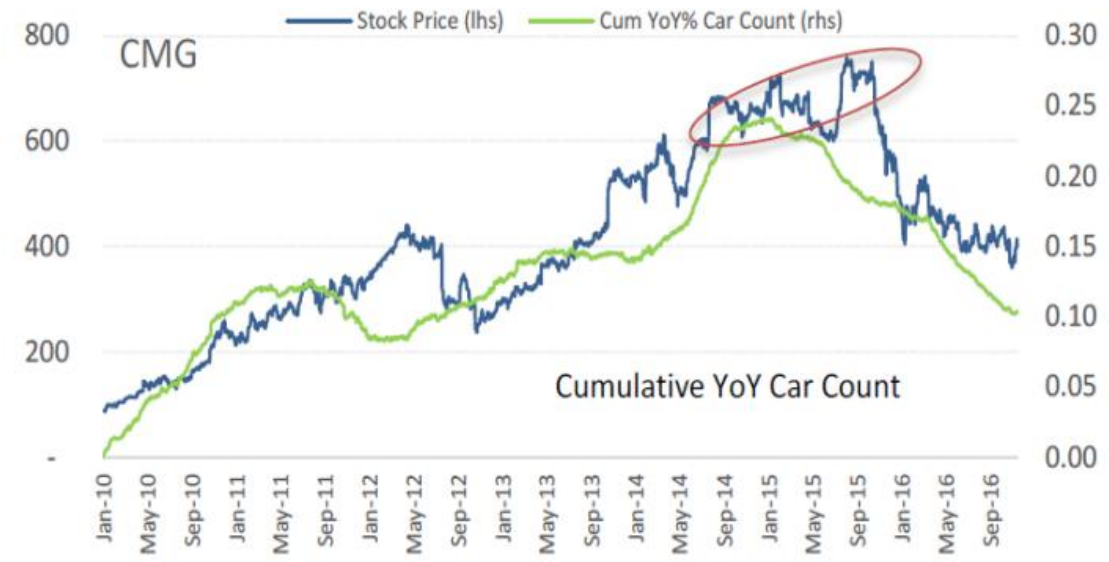
The case study presented below was prepared by a satellite imagery data analytics company that catalogued and backtested more than one million parking lot images that accounted for 1.5 billion cars over seven years. The traffic data obtained from satellite images can then be used to analyze cumulative YoY car count growth rates and compare it to revenue growth and share price dynamics.

Case Study

Figure 39 below shows how satellite imagery data was used to study Chipotle's structural inflection points over the 2014-2016 period. The car count traffic peaked in late 2014 before recording a quick fall from then onwards. This analysis was ahead of the street as sell side analysts only began revising Chipotle down 3 quarters later.

The E. coli outbreak in October 2015 had a big impact on the company's share price and the following outbreak of norovirus in March 2016 caught investors' attention as well. Most recently, Chipotle reported lower than expected Q4 2016 results on 2 February 2017.

Figure 39: CMG Cumulative Car Count vs Stock Price



Source: Satellite Data Analytics Provider

22. Equity (CMG) > Discretionary (Short-Term) > Employment, Online Search

Key Takeaway

Negative growth in active job listings and Google search data were used to correctly call worsening momentum for Chipotle.

Datasets

Job Listings Data

Completely unique in the industry, the job listing dataset only indexes jobs directly from employer websites. Updated daily with over 4 million jobs from more than 30,000 employers, the platform eliminates duplicate and expired job listings, as well as job pollution.

Google Trends Data

Google Trends is a public web facility based on Google Search that shows how often a particular search-term is entered relative to the total search-volume over time across various regions of the world.

Backtesting/Significance

Job Listings Data

Backtesting of the dataset demonstrated alpha in the dataset, with “Jobs Active” producing the highest and most consistent returns. Yearly hedge returns were between 6-8%.

Google Trends Data

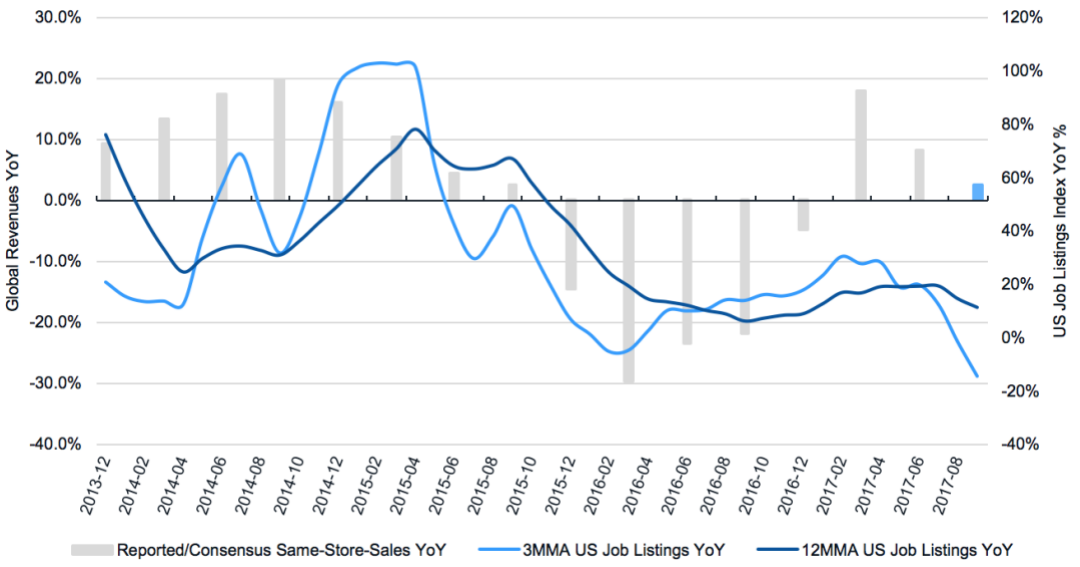
Over the last 20 quarters the Chipotle indicator has demonstrated a three-quarter hit rate of 90% in in-sample testing. This means that 90% of the time the three-month moving average has moved in the same direction as same store sales over a three-quarter period. One-quarter hit rate is 82%.

Case Study

On 16th October 2017, Eagle Alpha published a Data Analytics report titled “Chipotle: Job Listings and Online Search Point to Further Slowdown”. Figure 40 below shows 3 and 12 month moving averages of company job listings data against quarterly same store sales (SSS).

Growth in active job listings has been a strong indicator of inflections in revenue growth in the past (e.g. SSS slowdown in 2015 and bottom out in mid-2016). Job listings growth inflected down in Q2 2017 and the data was showing us that the longer-term momentum has stalled.

Figure 40: Chipotle Job Listings Index



Source: Job Listings Data

Figure 41 below shows that our Search Signal indicator, built using Chipotle specific terms from Google Trends, flattened in the June quarter and then turned negative as the 1 month index crossed below the 3 MMA. In our report, we noted that "historically there has been about a one quarter lead time before search terms get reflected in reported numbers".

Figure 41: Chipotle Search Signal

Source: Job Listings Data

On October 25th 2017, Chipotle reported weak Q3 numbers and lowered guidance for the rest of the year.

23. Equity (AR) > Discretionary (Short-Term) > Sensor Data

Key Takeaway

The sensor data provider observed production increases at Antero Resources and correctly anticipated management would raise company guidance.

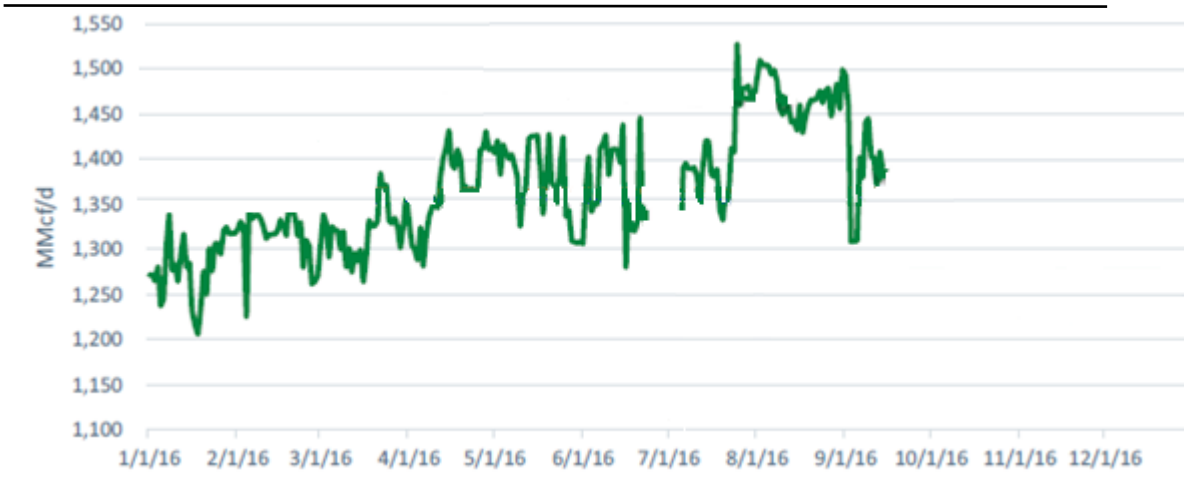
Dataset

The vendor provides energy market data and intelligence across the oil, power, natural gas and LNG, agriculture, petrochemical and NGLs. The company measures market fundamentals using thousands of patented and proprietary land, sea, and satellite monitors strategically deployed worldwide, delivering exceptional insight and intelligence to clients. History: since January 1999.

Case Study

In 2016, the vendor monitored Antero Resources' seven rigs: six in Marcellus, NY and one in Utica, NY. The sensor data provider observed production increases and anticipated management to raise the company guidance.

Figure 42: Antero Resources Daily Productions in 2016



Source: Sensor Data Provider

In September 2016, Antero Resources announced: "The increase in production guidance… is primarily a function of the improved recoveries and drilling efficiencies Antero has achieved throughout the year."

24. Equity (RSH) > Discretionary (Short-Term) > Geo-location

Key Takeaway

Geo-location data was used to anticipate disappointing sales at RadioShack stores.

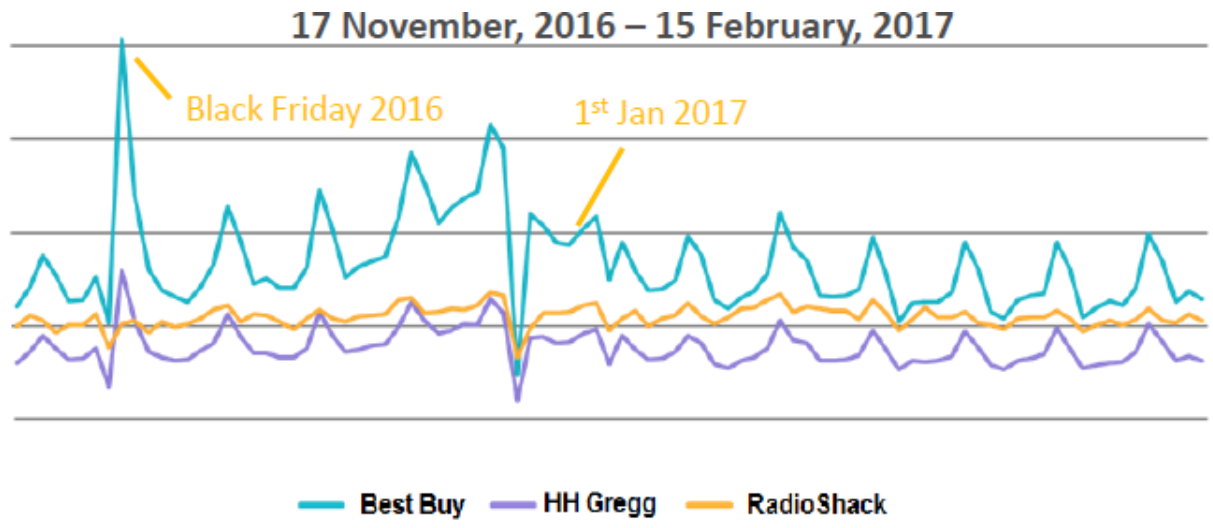
Dataset

The vendor receives raw location signals from 100s of app partners and also collects data from its own apps. In total it processes 1.7 billion visits per month from 130 million devices globally. History: since April 2016.

Case Study

The vendor analysed real-time tracking data in the 2016 holiday period. The figure below shows that RadioShack's sales remained more or less stable throughout the period, unlike major competitor Best Buy.

Figure 43: Visitors Per Store



Source: Geo-Location Data Provider

On March 9th, 2017, RadioShack announced bankruptcy and the closure of more than 200 stores citing decreased holiday sales as a contributing factor.

25. Equity (GoPro GPRO) > Discretionary (Short-Term) > Pricing Data

Key Takeaway

Online pricing data pointed to negative GoPro (GPRO) fundamentals, which were reflected in GoPro's subsequent results and, ultimately, GoPro's stock price.

Dataset

Pricing data is crawled from the websites of large retailers. This can provide an insight into long term trends, as well as the most recent trading performance, and is based on metrics such as average selling price and share of bestselling products in a category. Eagle Alpha owns this dataset.

The dataset currently supports over 100 brands/companies from categories including Tech Hardware (e.g. Apple, Cisco, Canon), Consumer Electronics (e.g. Sony, LG Electronics, Harman), Household Appliances (e.g. Whirlpool, Electrolux) and Leisure Products (e.g. Mattel, Hasbro) and the list is growing all the time.

Geography	Coverage	Mapped to Tickers	History	Collection Frequency	Delivery Frequency	Lag Time	Delivery Method
U.S	U.S Consumer Products Market	Yes	Since 2013	Daily, Weekly and Monthly	Daily, Weekly and Monthly	1 day	Excel/CSV

Backtesting/Significance

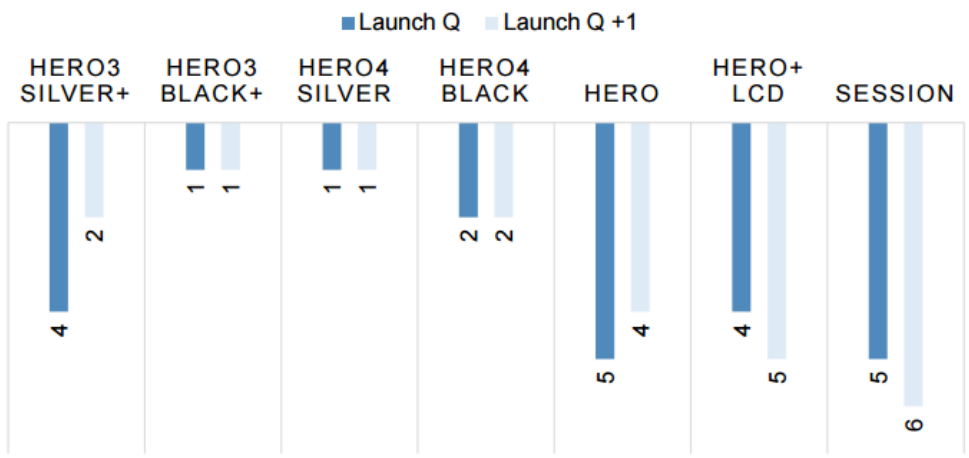
Pricing data has proved valuable at providing a directional indicator for the sales of consumer products, particularly around major product launches.

Note: Eagle Alpha has published 21 quarterly reports on consumer stocks incorporating crawled pricing data. 13 of these indicators proved accurate equating to a hit rate of 62%.

Case Study

Eagle Alpha first published a Data Analytics report on action camera manufacturer GoPro (GPRO US) on October 21st 2015. The data from US electronics websites pointed to potential weakness in GoPro revenue for the third quarter of that year. The crawled data was showing weak demand for GoPro's products, and a negative mix shift to lower end products that was likely to impact average selling prices (ASP). The report also highlighted weakness in the ranking of bestselling cameras, including the Session product which had recently been released (Figure 44).

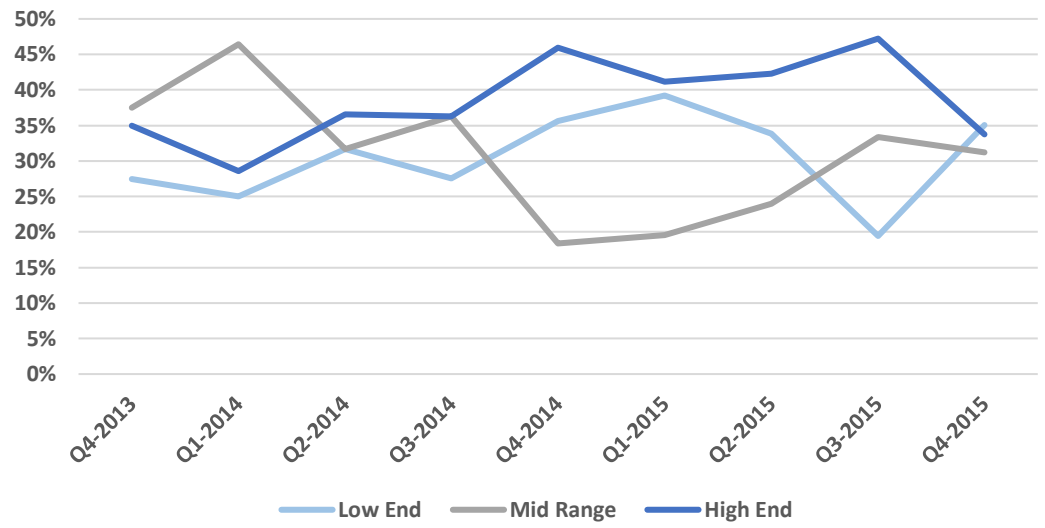
Figure 44: Weakness in the Ranking of Bestselling Cameras



Source: Eagle Alpha Analysis, E-Commerce Data

Subsequent Eagle Alpha reports in 2016 continued to point out that demand for GoPro was diminishing and that ASP trends continued to remain under pressure. A report on February 9th 2016 titled "Focus on Entry Level Products Ignores Issues Elsewhere" showed that GoPro was seeing price pressure across all price points and that growth in lower end cameras was unlikely to compensate for price and share compression of mid-range and higher end products (Figure 45).

Figure 45: Split of Bestsellers by Price Segment

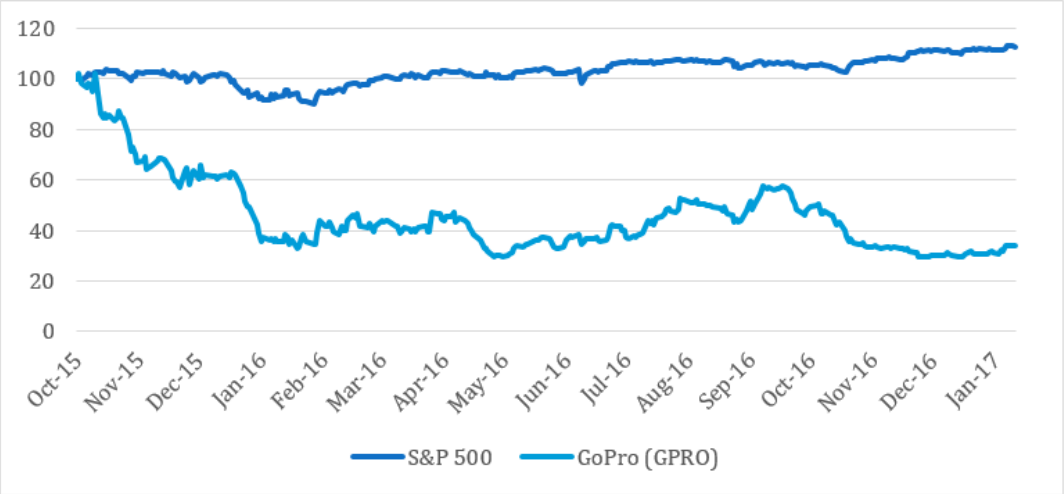


Source: Eagle Alpha Analysis, E-Commerce Data

By the September quarter of 2016 Eagle Alpha analysis was showing that "key elements are stabilizing for the action camera brand … but YoY pressure persists".

These negative fundamentals were reflected in GoPro's stock price over the period of Eagle Alpha's coverage. The relative performance of GoPro compared to the S&P 500 over the period can be seen in the following chart (Figure 46).

Figure 46: Indexed Share Price Performance



Source: Eagle Alpha Analysis, Bloomberg

26. Equity (Fitbit FIT) > Discretionary (Long-Term) > Pricing Data

Key Takeaway

Online retail data showed improving sell-through trends for Fitbit in the first half of 2017. On August 2nd 2017, Fitbit reported better than expected results with adjusted revenue of $353.3m vs consensus estimate of $339.2m.

Dataset

Pricing data is crawled from the websites of large retailers. This can provide an insight into long term trends, as well as the most recent trading performance, and is based on metrics such as average selling price and share of bestselling products in a category. Eagle Alpha owns this dataset.

The dataset currently supports over 100 brands/companies from categories including Tech Hardware (e.g. Apple, Cisco, Canon), Consumer Electronics (e.g. Sony, LG Electronics, Harman), Household Appliances (e.g. Whirlpool, Electrolux) and Leisure Products (e.g. Mattel, Hasbro) and the list is growing all the time.

Geography	Coverage	Mapped to Tickers	History	Collection Frequency	Delivery Frequency	Lag Time	Delivery Method
U.S	U.S Consumer Products Market	Yes	Since 2013	Daily, Weekly and Monthly	Daily, Weekly and Monthly	1 day	Excel/CSV

Backtesting/Significance

Eagle Alpha has published 21 quarterly reports on consumer stocks incorporating crawled pricing data. 13 of these indicators proved accurate equating to a hit rate of 62%.

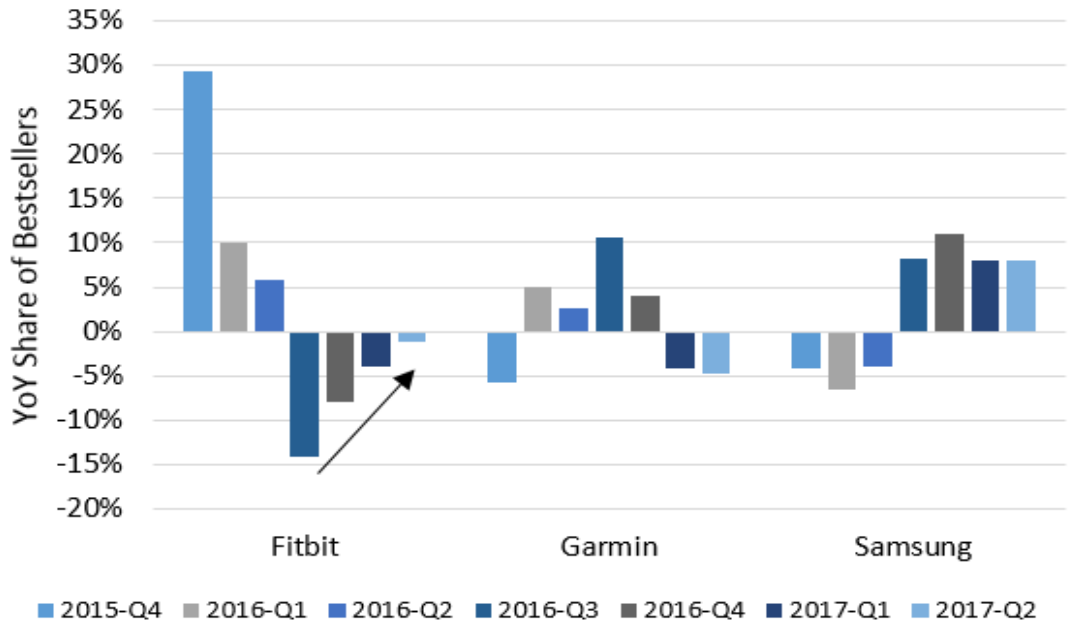
Case Study

On July 4th 2017, Eagle Alpha published a research note analyzing Fitbit's Q2 2017 with the use of online retail pricing data. Our analysis showed improving sell-through trends in the first half of the year. We noted the company's previous comments on improving inventory and concluded that Fitbit would report stronger than expected revenue.

Figure 47 below shows that Fitbit's share of bestsellers stabilized in Q2 as it further consolidated its position as the number one ranked fitness watch (Figure 48).

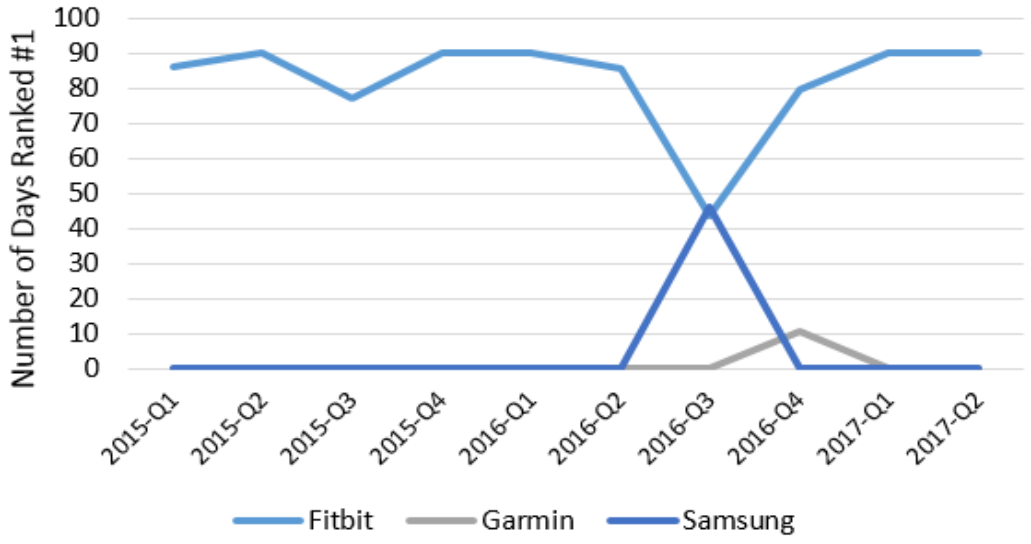
On August 2nd 2017, Fitbit reported better than expected results with adjusted revenue of $353.3m vs consensus estimate of $339.2m. Inventory was again mentioned by CEO James Park following these results. James stated "consumer demand in the second quarter was better than anticipated, enabling Fitbit to reduce channel inventory and generate better sales," which was in line with our analysis.

Figure 47: Online Data Showed Fitbit's Improving Share Trends



Source: Eagle Alpha Analysis, Online Retail Data

Figure 48: Fitbit Consolidated Top Ranking in Fitness Watches in Q2 2017



Source: Eagle Alpha Analysis, Online Retail Data

27. Equity (HubSpot HUBS) > Discretionary (Long-Term) > Pricing Data

Key Takeaway

Accelerating growth in active job listings and Google search data were used to correctly call improving momentum for HubSpot.

Datasets

Job Listings Data

Completely unique in the industry, the job listing dataset only indexes jobs directly from employer websites. Updated daily with over 4 million jobs from more than 30,000 employers, the platform eliminates duplicate and expired job listings, as well as job pollution. From the core platform, the company has developed an array of products and services and achieved significant market traction in two primary business units: Candidate Sourcing and Job Market Data and Analytics.

Google Trends Data

Google Trends is a public web facility based on Google Search that shows how often a particular search-term is entered relative to the total search-volume over time across various regions of the world.

Backtesting/Significance

Job Listings Data

Backtesting of the dataset demonstrated alpha in the dataset, with "Jobs Active" producing the highest and most consistent returns. Yearly hedge returns were between 6-8%.

Google Trends Data

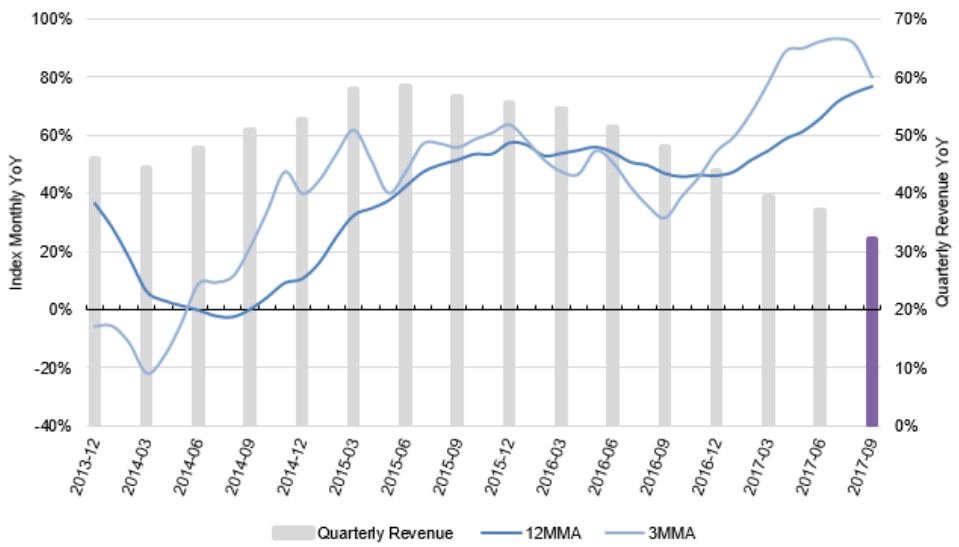
Eagle Alpha's Data Analytics team has published 36 quarterly reports for consumer companies incorporating Google search data. 25 of these indicators proved accurate equating to a hit rate of 69%.

Case Study

On 11th October 2017, Eagle Alpha published a Data Analytics report titled "HubSpot: Alternative Data Signals Possible Revenue Inflection". Figure 49 below shows 3 and 12 month moving averages (MMA) of company job listings data against quarterly revenue numbers.

The correlation between these two time series diverged in late 2016 and we concluded: "With job listings accelerating to the upside in 2017, management views on business prospects appear to have turned bullish. We view this positively for revenue growth in coming quarters."

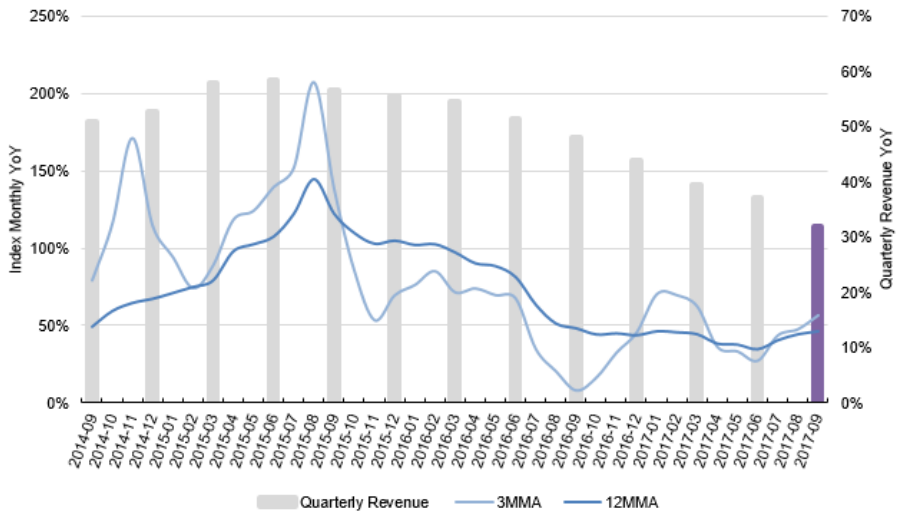
Figure 49: HubSpot Job Listings Index



Source: Eagle Alpha Analysis, Online Retail Data

Figure 50 below shows that our Search Signal indicator, built using HubSpot specific terms from Google Trends, experienced a similar trend as it bottomed in late 2016 and the 3 MMA crossed over the 12 MMA. We view this as a positive inflection point, but it can take several quarters to be reflected in company fundamentals.

Figure 50: HubSpot Search Signal



Source: Eagle Alpha Analysis, Online Retail Data

On November 3rd 2017, HubSpot reported better-than-expected Q3 results with revenue growing by 38% which was driven by 40% subscription revenue growth. The company also raised full year guidance ($370M to $371M vs consensus of $367.94M).

28. Equity (Expedia EXPE) > Discretionary (Long-Term) > Online Pricing

Key Takeaway

A web data provider accurately predicted that EXPE would miss 3Q17 room-night growth expectations, based on decelerating trends in reservation growth it started flagging in August. Its room-night estimate was within 1pp of reported.

Dataset

For EXPE, the provider analyzes comprehensive gross reservations, active properties, and bookable rooms on the Expedia.com and Hotels.com websites. Its systems are designed to visit the sites each day and find all active properties, then observe key facts for each property including location, property type, and booking messages, e.g. “10 people booked this property in the last 48 hours”. The data includes reservations from all EXPE brands (Expedia.com, Hotels.com, Travelocity, Orbitz, etc.) and is granular down to the property level.

Backtesting/Significance

The provider has been analyzing EXPE’s gross reservation data since mid-2015. Gross reservations refer to distinct accommodation transactions that are pre-cancellation and do not incorporate length of stay. Because gross reservations are captured at the time of booking, versus EXPE reporting at the time of stay, the data is a leading indicator of quarterly trends. The provider assumes a typical 1-month lag between time of book and time of stay.

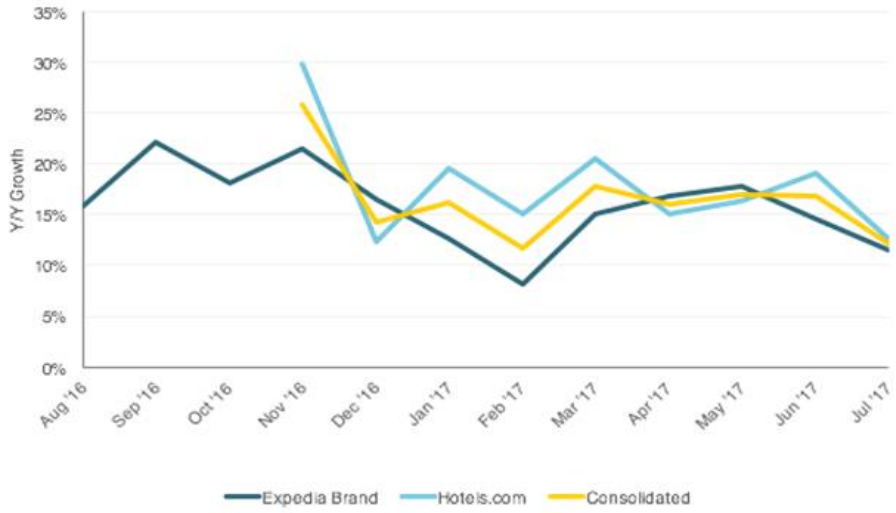
Gross reservation data has accurately predicted room-night growth within 2pp over the past three quarters.

Case Study

In August 2017, the provider reported that based on June and July data, EXPE’s core OTA business (i.e. ex-HomeAway) was seeing worsening trends in room-night growth. Both Hotels.com and what they refer to as Expedia brand (Expedia.com + regional brands) saw a slowdown in July, while Expedia brand had been deteriorating since May. See Figure 51.

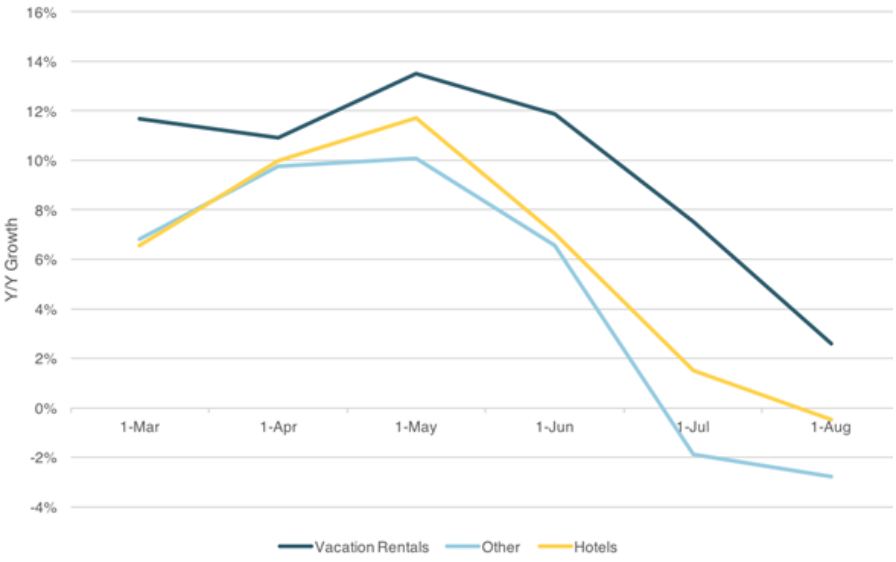
In its September update, the provider indicated that growth remained soft, with core OTA room-nights to come in at 13-15% Y/Y in 3Q17. The major driver of the pressure continued to be Expedia brands, as Hotels.com strengthened, with US destination reservations (the region represents roughly half of room-nights) dropping to flat Y/Y. See Figure 52.

Figure 51: Expedia Brand and Hotels.com Y/Y Growth in Gross Reservations Decelerated in July



Source: Online Pricing Data Provider

Figure 52: Expedia's Y/Y Reservation Growth Rate by Property Type, US



Source: Online Pricing Data Provider

The final 3Q17 preview indicated that core OTA trends remained light in September, suggesting a slow start to 4Q17 room-nights as well. The data provider also flagged in its preview that metasearch rebalancing efforts appeared to be a headwind in 3Q and that HomeAway would likely be the one bright spot due to continued monetization success. Both trends were consistent with company 3Q17 conference call commentary.

29. Equity (Square SQ) > Discretionary (Long-Term) > Consumer Transaction Data

Key Takeaway

Analysis of the email receipt data for Square indicated that the growth in number of sellers has been in decline since the first quarter of 2016.

Dataset

The dataset is delivered through a partnership with a provider that collects anonymized purchase data from around 2 million active shoppers, scanned from email purchase receipts. It covers over 600 merchants from more than 25 industries.

The provider transforms multi-form unstructured email receipt data into a normalized and digestible consumer transaction dataset. The dataset is also granular as it includes item and SKU-level transaction data, which is filtered into 53 product categories.

Geography	Coverage	Mapped to Tickers	History	Collection Frequency	Delivery Frequency	Lag Time	Delivery Method
USA	600 merchants	No	Since 2013	Real-Time	Weekly	7 days	Excel, TSV

Backtesting/Significance

Eagle Alpha's predictive model for Square has a mean absolute percentage error (MAPE) of 1.4%, compared to a consensus error of 4.7%. The standard deviation of our error is 1.6%.

Eagle Alpha's Data Analytics team has published 10 quarterly reports incorporating predictive indicators using the US Email Receipt Data. 8 of these indicators proved accurate equating to a hit rate of 80%.

Case Study

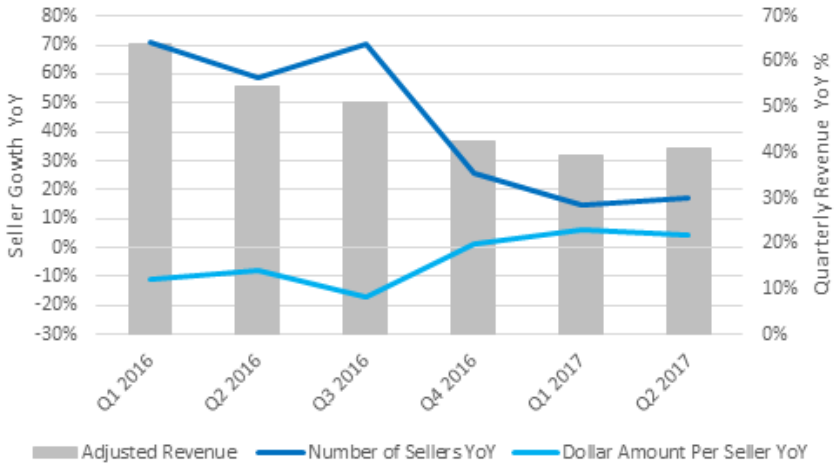
Square held an analyst day in May 2017 where the company gave details on its growth strategies. Two key parts for future growth are to move the company up market to get larger merchants, or sellers, to use Square devices and secondly to grow and retain existing merchants.

Larger sellers are users that have greater than $125k in annualized gross payment volume (GPV). In the June quarter of 2017, Square reported stronger than expected revenue. Adjusted revenue ticked up sequentially to 41% YoY growth, from 39% in the March quarter.

In our Data Analytics report, titled "Email Receipt Data Reveals Improving Seller Fundamentals", Eagle Alpha examined historical trends for Square sellers and dollar spend at sellers using an email receipt dataset. Square does not disclose seller metrics in annual or quarterly reports.

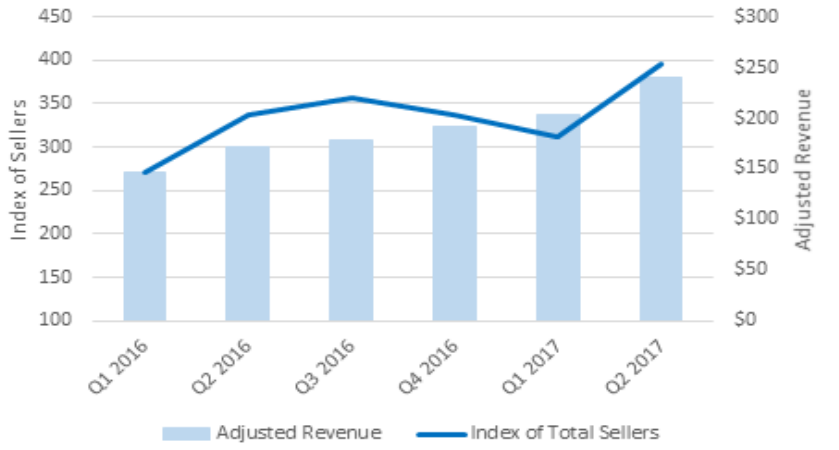
Analysis of the email receipt data indicated that the growth in number of sellers has been in decline since the first quarter of 2016. This has been the primary driver of slowing revenue growth for Square over this period (Figure 53). However, in the June quarter the number of sellers using Square's platform increased sequentially from 15% in March to 17%. This can also be seen in the unique seller index, where sellers increased to 396 from an average of 325 for all of 2016 (Figure 54).

Figure 53: Metrics for Sellers Improving



Source: Eagle Alpha Analysis

Figure 54: Unique Seller Index



Source: Eagle Alpha Analysis

From the email receipt data, we can also observe that average spend per seller was declining on a YoY basis during 2016. However, growth in spend per seller turned positive in Q4 2016 and this improvement has been helping revenue growth over the first half of 2017.

Data on growth in total sellers, and the amount sold per seller, may be an indication that growth strategies outlined by the company are starting to gain traction and are a positive sign for longer term fundamentals of the company.

On November 8th 2017, the company reported strong third quarter results and raised guidance for the year. Full year adjusted revenue was increased to $963m-$966m from $925m-$935m.

30. Equity (Finish Line FINL) > Discretionary (Long-Term) > Online Search Data

Key Takeaway

The search indicator correctly predicted weakness in SSS ahead of earnings in December 2016. The company's stock reacted to the change in fundamentals and is down significantly since then.

Dataset

Google Trends is a public web tool based on Google Search that shows how often a particular search term is entered relative to the total search-volume over time across various regions of the world. Using Google Trends, Eagle Alpha has built company specific indices based on search terms that are related to a given retailer's product offering. This involves an exhaustive process for identifying search terms related to a company's revenues using both internal and third party tools.

Geography	Coverage	Mapped to Tickers	History	Collection Frequency	Delivery Frequency	Lag Time	Delivery Method
Global	B2B, Retail, Luxury, Restaurants	Yes	Since 2006	Daily, Weekly and Monthly	Monthly	1-3 days	JPG, PNG, SVG, PDF, CSV

Backtesting/Significance

Observing crossing points of three-month and one-month moving averages of one of these search indices has proved predictive of inflection points in revenue growth for companies across a broad range of sectors including retail, luxury goods and restaurants.

Over the last 18 quarters the Finish Line indicator demonstrated a three-quarter hit rate of 78% in in-sample testing. This means that 78% of the time the three-month moving average has moved in the same direction as same store sales over a three-quarter period.

Note: Eagle Alpha's Data Analytics team has published 49 quarterly reports for consumer companies incorporating Google search data. 37 of these indicators proved accurate equating to a hit rate of 76%.

Case Study

On December 5th 2016, Eagle Alpha published a report on US sports retailer Finish Line using Google Trends data. Figure 55 below shows the signal pointing to a strong uptrend prior to the August 2016 quarter when the company reported better than expected results.

However, the directional change of the index in the November 2016 quarter indicated that Finish Line's SSS (same-store sales) growth was at risk and Eagle Alpha data insight analysts anticipated that management outlook for the following quarter could disappoint.

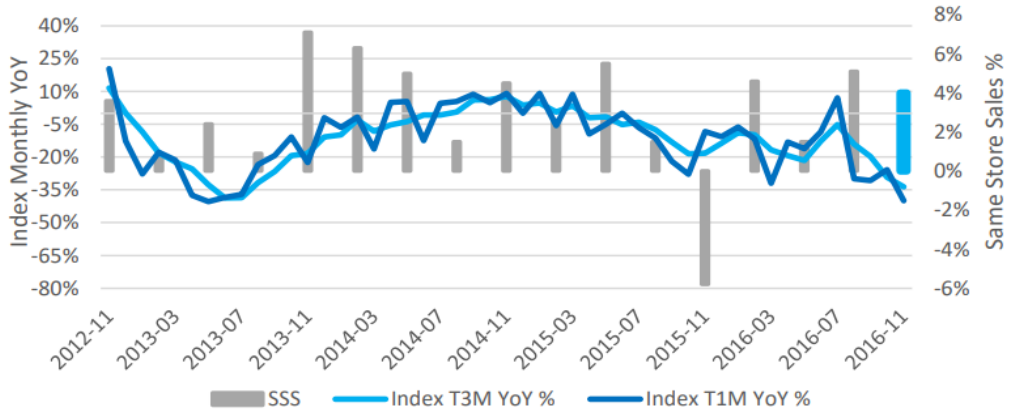
On December 21st 2016, Finish Line reported a weak quarter and gave poor guidance for the following quarter. "We are disappointed that our third quarter sales and earnings fell short of our expectations," said Sam Sato, Chief Executive Officer of Finish Line.

Eagle Alpha's Data Analytics report on December 5th 2016, indicated that the search index, an indicator of consumer demand based on online search data, was pointing to weakness in SSS for Finish Line. The report stated: "the sharp downtrend observed in the index in recent months (Figure 55) indicates that expectations for SSS growth at FINL are at risk, or that management outlook for the February quarter may disappoint".

Finish line reported SSS of 0.7% for the November 2016 quarter compared to expectations for growth in excess of 4%. The company also reported a significant drop in SSS in the February quarter of -4.5%.

The search indicator correctly predicted weakness in SSS ahead of earnings. The company's stock reacted to the change in fundamentals and is down significantly since then.

Figure 55: Search Signal Index for FINL



Source: Eagle Alpha Analysis

31. Equity (Burberry BRBY) > Discretionary (Long-Term) > Online Search Data

Key Takeaway

Citi concluded that the short-term 1-month YoY observation crossing over the 3-month moving average YoY indicates major inflection points of same store sales growth for Burberry.

Dataset

Google Trends is a public web tool based on Google Search that shows how often a particular search term is entered relative to the total search-volume over time across various regions of the world. Using Google Trends, Eagle Alpha has built company specific indices based on search terms that are related to a given retailer's product offering. This involves an exhaustive process for identifying search terms related to a company's revenues using both internal and third party tools.

Geography	Coverage	Mapped to Tickers	History	Collection Frequency	Delivery Frequency	Lag Time	Delivery Method
Global	B2B, Retail, Luxury, Restaurants	Yes	Since 2006	Daily, Weekly and Monthly	Monthly	1-3 days	JPG, PNG, SVG, PDF, CSV

Backtesting/Significance

Observing crossing points of three-month and one-month moving averages of one of these search indices has proved predictive of inflection points in revenue growth for companies across a broad range of sectors including retail, luxury goods, restaurants and 2B software.

Over the last 19 quarters the Burberry indicator demonstrated a two-quarter hit rate of 58% in in-sample testing, and three-quarter hit rate of 68%. This means that 58% of the time the three-month moving average has moved in the same direction as same store sales over a two-quarter period.

We prefer to measure the accuracy of our search signals indicators by reference to the hit rate, as we believe it better captures the goal of the tool to identify inflection points in growth for a company. However, below we present analysis from Citi's March 2017 report "Searching for Alpha: Big Data" which focuses on correlations.

Figure 56 appears to support our prior that short-term 1-month YoY observation crossing over the 3-month moving average YoY indicates major inflection points of same store sales growth as depicted by the green circles on the chart. Another interesting finding is that consensus[29] is pretty bad at predicting same store sales growth. In fact, it only achieves close to 20% of correlation with actual reported figures. With either the 1-month YoY or 3-month moving average YoY measures based on EA equity index for Burberry, the correlation jumps to over 70% which is a significant improvement. The additional advantage of this search data is its timeliness – the data at the end of the quarter is available immediately, whilst official figures typically are announced at least 3 weeks after quarter-ends. The timeliness and much improved correlation of the new dataset with actual reported figures make such an offering appealing.

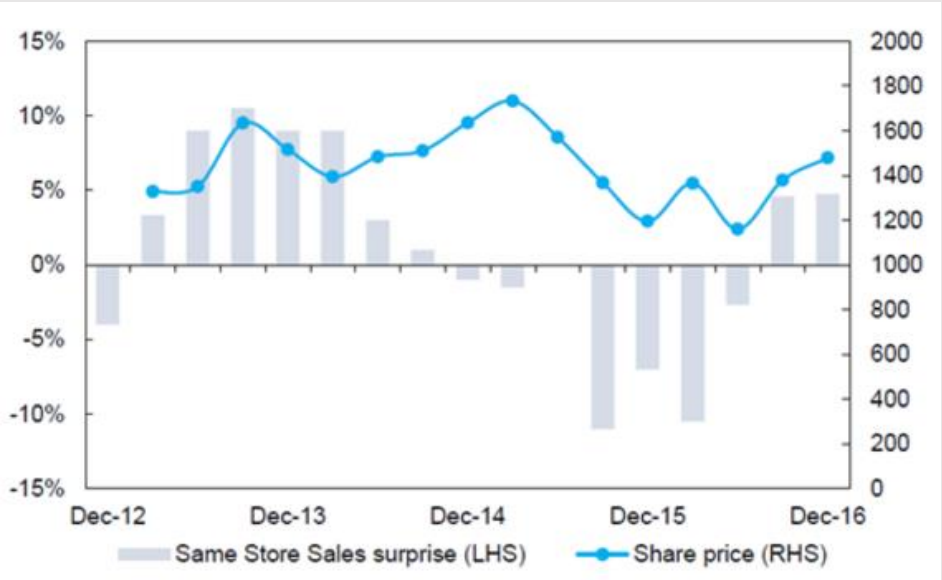
In Figure 57 we highlight that same store sales does have pricing impact especially the surprise element of it. That is, markets react to the positive/negative sales surprise. This suggests that, if we are able to predict the inflection points better with Google Trends data, there could be pricing implications from being able to act sooner and more accurately than the bulk of investors.

[29] *The consensus data is sourced from Bloomberg but it has low analyst coverage issues as not all analysts publish their estimates for same store sales.*

Figure 56: Burberry Same Store Sales vs Eagle Alpha Stock Index

Source: Eagle Alpha, Citi Research

Figure 57: Burberry's Same Store Sales Surprise vs Share Price



Source: Bloomberg, Citi Research

32. Equity (Sportswear Industry) > Discretionary (Long-Term) > Online Search Data

Key Takeaway

Online search data provided early indicator of weakness in sportswear sector. The industry analysis supported our case for fundamental weakness for FINL and FL and points to longer term fundamental issues for the sportswear space.

Dataset

Google Trends is a public web tool based on Google Search that shows how often a particular search term is entered relative to the total search-volume over time across various regions of the world. Using Google Trends, Eagle Alpha has built company specific indices based on search terms that are related to a given retailer's product offering. This involves an exhaustive process for identifying search terms related to a company's revenues using both internal and third party tools.

Geography	Coverage	Mapped to Tickers	History	Collection Frequency	Delivery Frequency	Lag Time	Delivery Method
Global	B2B, Retail, Luxury, Restaurants	Yes	Since 2006	Daily, Weekly and Monthly	Monthly	1-3 days	JPG, PNG, SVG, PDF, CSV

Backtesting/Significance

Observing crossing points of three-month and one-month moving averages of one of these search indices has proved predictive of inflection points in revenue growth metrics for companies across a broad range of sectors including retail, luxury goods and restaurants. Over a nineteen quarter period a Search Signal for Finish Line (FINL) and Foot Locker (FL) has demonstrated a three-quarter hit rate of 74% in in-sample testing. This means that 74% of the time the three-month moving average has moved in the same direction as same store sales (SSS) over a three-quarter period. Over a twelve quarter period the indicator for Under Armour (UAA) has demonstrated a three-quarter hit rate of 67% in in-sample testing.

Note: we have published 36 quarterly reports for consumer companies incorporating Google search data. 25 of these indicators proved accurate (69% hit rate).

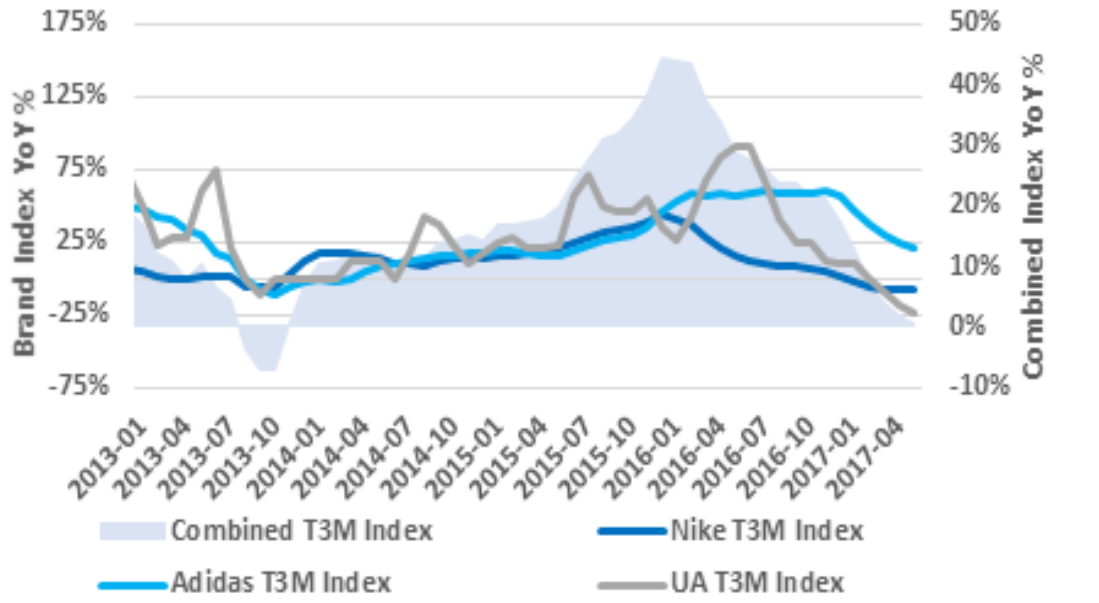
Case Study

In December 2016, Eagle Alpha published a report on US sports retailer FINL and FL using Google Trends data. The Search Signal for Finish Line was indicating that SSS growth was at risk and could disappoint management expectations. Subsequently Finish Line reported a poor quarter and offered weak guidance for the March quarter. In February of 2017 Eagle Alpha published a follow up report on the sports retailers where we highlighted that the Search Signal for FL had deteriorated and pointed to a risk of the company missing expectations. The company reported strong SSS and offered positive guidance for the April quarter. However, in April the company preannounced negatively for the quarter, reducing SSS guidance from 5% to 2%. When the company reported the quarter SSS was just 0.5%. Search Signals proved to be an early indicator of fundamental weakness for both FL and FINL.

The negative trends seen at FL and FINL triggered Eagle Alpha to take a deeper dive into the industry to see what could be catching company management off guard. Analysis of industry trends led to the conclusion that the problems for FINL and FL were more rooted in industry dynamics than company specific issues. This can be seen in Figure 58. Using Google Trends, we created indices of sports apparel and sports footwear for the big three global sports brands of Adidas, Nike and Under Armour.

We also created a combined index of all three brands. As can be seen in Figure 58, a three-month moving average of our index began to roll over in early 2016, stabilized somewhat in mid-2016 and then accelerated to the downside in late 2016. The decline in the combined index happened in waves. This can be observed by looking at the individual brands. Nike rolled over in early 2016, followed by Under Armour in mid-2016 and finally by Adidas in late 2016.

Figure 58: Big 3 Global Brands in Decline Through 2016



Source: Eagle Alpha Analysis

This industry analysis supported our case for fundamental weakness for FINL and FL and pointed to longer term fundamental issues for both companies. The Search Signal indicators for both companies remained weak in June, indicating further potential weakness in SSS. Subsequently both FL and FINL reported very weak quarters and the stocks traded off significantly. From the time of Eagle Alpha's first Search Signal report on FINL and FL in December of 2016 to September of 2017 the stocks declined 65% and 55% respectively.

Investors have only recently started to consider the implications for Adidas and Nike of the negative results at the retailers and Under Armour. These stocks have been weak in recent months and have seen broker downgrades. However, headwinds for the sector were apparent in search data over 12 months before.

33. Equity (Time Warner TWX) > Discretionary (Long-Term) > Mobile App Data

Key Takeaway

App data showed an early indicator of a positive inflection in revenue growth for HBO, one of Time Warner's largest divisions.

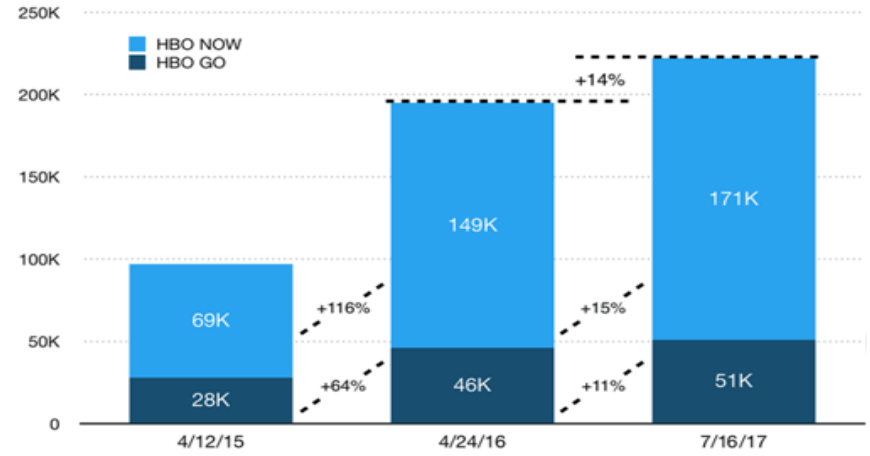
Dataset

App download and usage data is an example of an alternative data source that can be used to analyze product adoption and brand satisfaction. This type of data enables investors to track activity in a wide range of sectors from banking to food delivery to online entertainment.

Case Study

On July 19th 2017, an app data provider published a report highlighting record number of new installations of both HBO NOW and HBO GO. The network's streaming services reached the peak of approximately 222,000 installs due to the premiere of season seven of Game of Thrones (Figure 59).

Figure 59: HBO Mobile App Peak New Installs (United States)



Source: App Data Provider

The subsequent report publish on August 18th stated: "The sizable influx of new subscribers reflected in our download estimates at the time drove record single-day revenue one month later as their 30-day free trials converted to paid subscriptions." HBO NOW rose to the top of the App Store revenue charts and grossed around $2.6 million across both the App Store and Google Play on August 16th 2017. This represents a 41% increase over the previous record of $1.3 million recorded on June 24th 2016.

HBO is one of Time Warner's largest divisions but growth for the division has been slowing in recent quarters, therefore this analysis could be an early indicator of a positive inflection in revenue growth for HBO.

34. Equity (Activision Blizzard) > Discretionary (Long-Term) > Social Media Data

Key Takeaway

Using social media data, we correctly highlighted that the Overwatch game was well positioned to set a new sales record for Activision Blizzard (ATVI).

Backtesting/Significance

Eagle Alpha's Data Analytics team has published 15 quarterly reports incorporating social data. 11 of these indicators proved accurate equating to a hit rate of 73%.

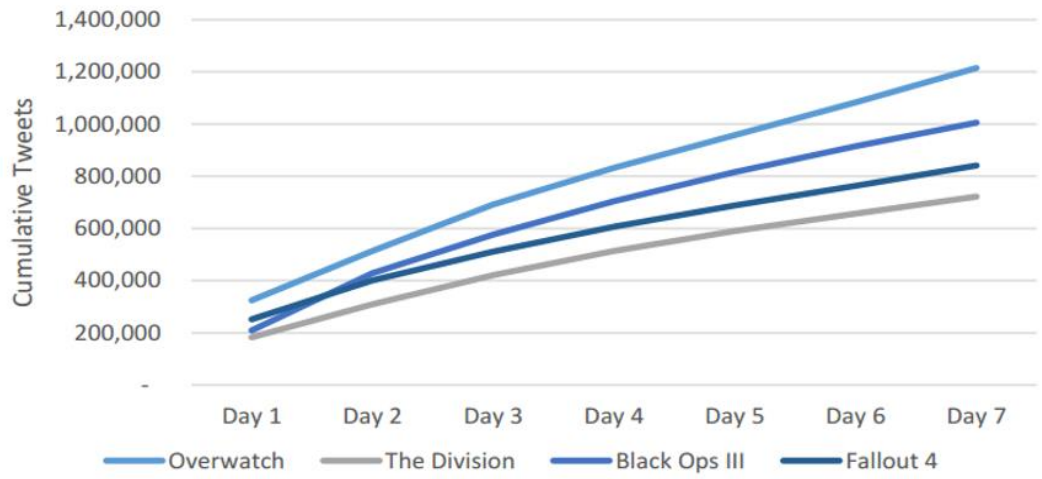
Dataset

Social media data is created through the public's interaction with social media platforms such as Twitter, Facebook, YouTube and Twitch. Access to this data can be obtained directly from the provider themselves or often through third party platforms. This case study also incorporated Eagle Alpha's Web Queries tool. Web Queries is a query based tool that enables clients to obtain analytics on over 90 million web sources including: blogs, image and video sites, forums, review sites, social media and news sites.

Case Study

On June 3rd 2016, Eagle Alpha published a Data Analytics report on video games publisher Activision Blizzard (ATVI US) titled "Activision Blizzard: Overwatch Positioned to Set New Sales Record." The report highlighted the potential success of the company's latest game, Overwatch. Overwatch enjoyed stronger Twitter visibility when compared to the competition and the strong positive sentiment towards the title pointed to positive consumer reaction to the title.

Figure 60: Overwatch Twitter Visibility Much Stronger than Comp Titles

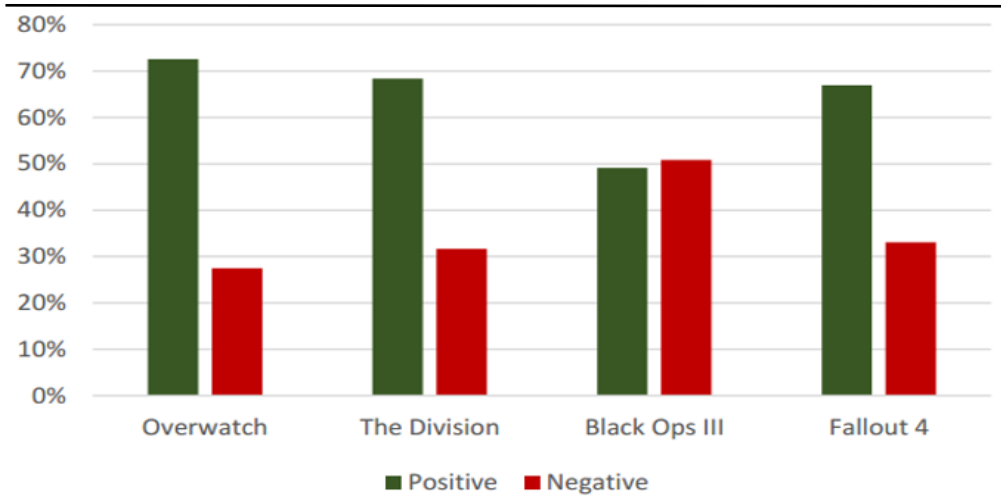


Source: Eagle Alpha Web Queries

Over 1.2m tweets mentioned Overwatch in the first week post the game's launch, while the competition gathered much lower number of mentions: Black Ops III (1m), Fallout (840k) and the Division (720k).

The Eagle Alpha Web Queries tool was also used to analyse consumer comments and sentiment around Overwatch. The first week comparative statistics are presented below, and clearly show a stronger performance for Overwatch compared to rival titles.

Figure 61: Overwatch Leads on Twitter Sentiment for First Week



Source: Eagle Alpha Web Queries, Twitter

On June 14th 2016, Activision reported that Overwatch had over 10 million players after only three weeks of the launch. The Division, on the other hand, was launched in the beginning of March 2016 and had 9.5 million players more than two months after the launch.

In August 2016, Overwatch hit the 15 million user mark, reaching the milestone faster than any other game in Blizzard's history. More recently, January 27th 2017, Activision announced that Overwatch reached another milestone by amassing more than 25 million players.

35. Equity > Discretionary (Long-Term) > Trade Data

Key Takeaway

The South Korea real-time export data accurately tracked revenue of construction machinery companies throughout the Q1'12 – Q2'17 period.

Dataset

This vendor delivers high frequency and comprehensive South Korea export data. Preliminary export data includes total volume and value of all products, at all country destinations that are exported from South Korea. The dataset is created by aggregating and analyzing customs declaration forms. History: Since 2003.

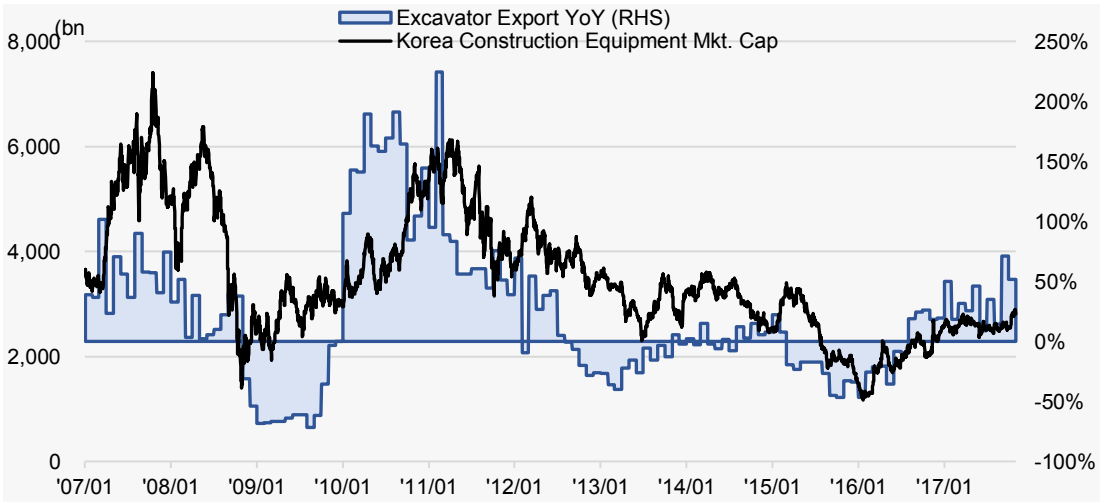
Backtesting/Significance

The South Korea real-time export data shows R-squared of 0.86 with revenues for construction machinery companies over a 5 year period and R-squared of 0.75 with YoY revenue growth. The calculated export value also correctly projected the directional movement of reported revenues for construction machinery companies in all quarters between Q1 2012 and Q2 2017.

Case Study

Monthly export value of excavator showed correlation of 82% with total market capital of S. Korean construction machinery companies from Jan. 2012 ~ Sep. 2017. Data was also indicative to Volvo's (STO:VOLV-A) market cap. with 62% correlation to S. Korean excavator export value. Volvo's Changwon plant accounts 40% of group's total revenue and exports key components to 8 other production facilities.

Figure 62: Excavator Qtr Export vs Revenue Aggregate of Related Companies



Source: Trade Data Provider

36. Equity (Lululemon LULU) > Discretionary (Long-Term) > Online Search, Social Media, Pricing Data

Key Takeaway

Eagle Alpha's analysis proved correct, i.e. Lululemon (LULU) reported sales growth of 13% YoY in Q3 2016 which was in line with our expectations.

Datasets

Google Trends is a public web tool based on Google Search that shows how often a particular search-term is entered relative to the total search-volume over time across various regions of the world.

Social media data is created through the public's interaction with social media platforms such as Twitter, Facebook and YouTube. Access to this data can be obtained directly from the provider themselves or often through third party platforms.

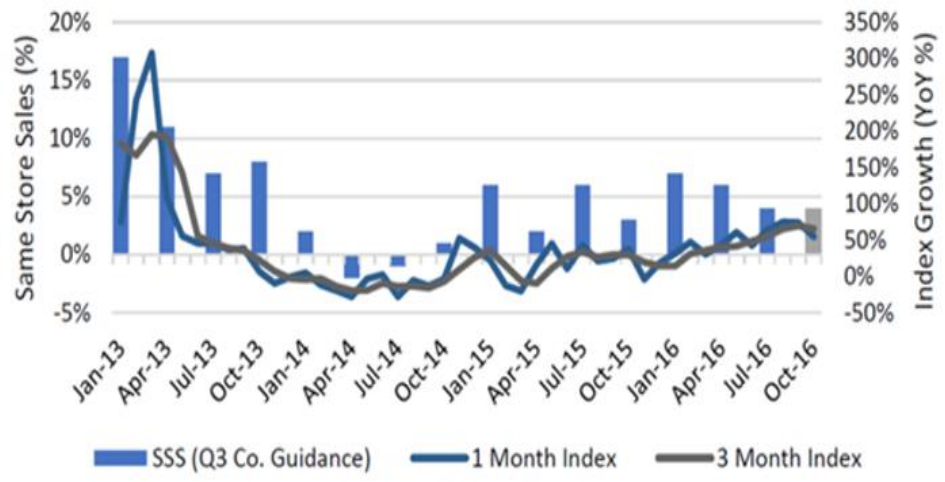
Pricing data is crawled from the websites of large retailers. This can provide an insight into long term trends, as well as the most recent trading performance, and is based on metrics such as average selling price and share of bestselling products in a category. Eagle Alpha owns this dataset.

This case study also incorporated Eagle Alpha's Web Queries tool. Web Queries is a query based tool that enables clients to search over 90 million web sources including: blogs, image and video sites, forums, review sites, social media and news sites.

Case Study

Lululemon's (LULU) same-store sales (SSS) have been decelerating since the beginning of 2016 with Q2 numbers particularly disappointing. On November 22nd 2016, Eagle Alpha published a note on Lululemon highlighting stabilized momentum in Q3. Figure 63 below shows that search data suggested stable to improving SSS in Q3 2016.

Figure 63: Search Signal Index for LULU

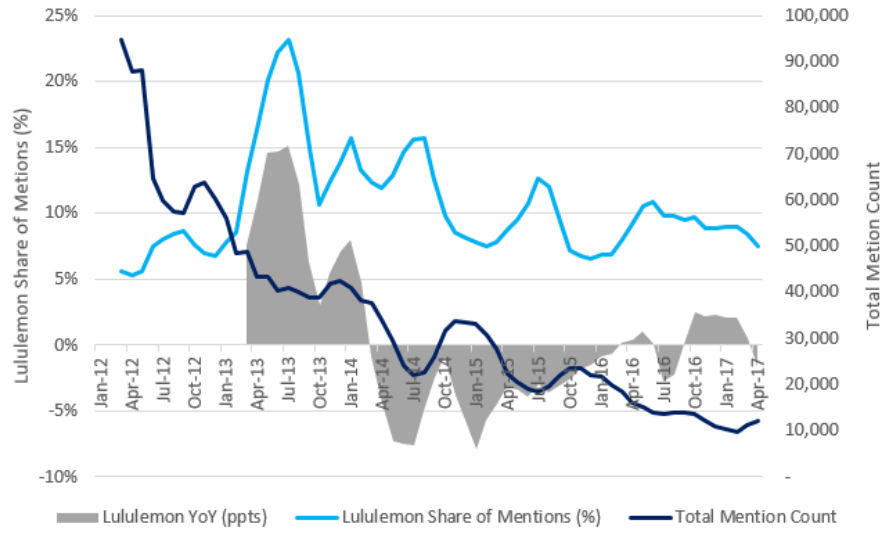


Source: Eagle Alpha Analysis

According to social media data, the athleisure trend was experiencing continued pressure. However, Lululemon improved its competitive position as the company received a meaningful share of mentions across blogs and forums for the first time since 2014 (Figure 64).

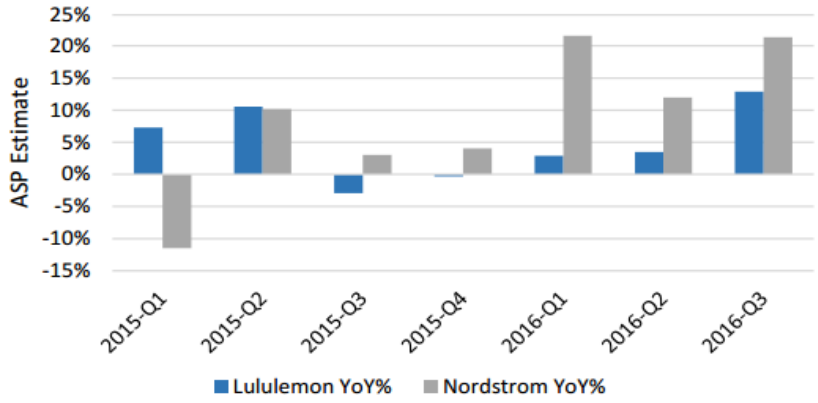
Eagle Alpha's YoY average selling price (ASP) estimate for Lululemon grew by 13% in Q3 2016, representing the fastest change since Q1 2015 (Figure 65).

Figure 64: Lululemon Share of Mentions Across Blogs and Forums



Source: Eagle Alpha Web Queries

Figure 65: Lululemon Price Growth Accelerated



Source: Eagle Alpha Analysis, Online Pricing Data

In our November 2016 report, we concluded: “consensus estimates are for Q3 2016 revenues growth of 13% YoY, slightly behind the 14% number reported in Q2. This revenue growth looks achievable.” On December 7th 2016, Lululemon reported sales growth of 13% YoY – in line with our expectations.

37. Equity > Discretionary (Long-Term) > Online Reviews

Key Takeaway

Companies that get better reviews from employees post better share price performance.

Dataset

This vendor operates a website where employees and former employees review companies and management of companies on an anonymous basis.

Case Study

The vendor researched the performance of 'Best Places to Work' companies. Companies getting the award have satisfied employees and also turned out to be good investments. The table below shows the relative performance of an equally-weighted portfolio of the 26 public companies included in the 2016 Best Places to Work list,

Figure 66: Stock Returns for 2016 Best Places to Work Companies vs. the S&P 500

Employer Name	January 2, 2015 Stock Price	November 18, 2015 Stock Price	Percentage Gain (Loss)
Paycom	$26.06	$42.00	61.2%
Hubspot	$33.48	$52.91	58.0%
Expedia	$85.76	$125.30	46.1%
Google	$529.55	$760.01	43.5%
Facebook	$78.45	$107.77	37.4%
NIKE	$95.03	$125.78	32.4%
Salesforce	$59.24	$77.35	30.6%
Adobe	$72.34	$91.27	26.2%
Guidewire	$49.97	$58.62	17.3%
Red Hat	$68.99	$80.53	16.7%
Costco Wholesale	$141.61	$161.25	13.9%
Southwest Airlines	$42.69	$46.31	8.5%
LinkedIn	$229.65	$248.76	8.3%
Apple	$109.33	$117.29	7.3%
Gartner	$83.60	$87.87	5.1%
Workday	$80.41	$83.68	4.1%
Stryker	$93.99	$97.01	3.2%
Delta Air Lines	$49.18	$47.75	-2.9%
Akamai	$63.25	$59.25	-6.3%
Eastman Chemical	$76.48	$71.62	-6.4%
Vivint Solar	$9.24	$7.99	-13.5%
Zillow	$32.40	$26.79	-17.3%
Chevron	$112.58	$92.21	-18.1%
F5 Networks	$130.33	$104.13	-20.1%
Twitter	$36.56	$25.90	-29.2%
SolarCity	$52.92	$27.75	-47.6%
2016 Best Places to Work Portfolio Return			*9.9%*
S&P 500 Return			*1.2%*

Source: ESG Data Vendor

38. Equity > Discretionary (Long-Term) > Online Reviews

Key Takeaway

The assessment of service and product quality of auto lenders has proven to be indicative of deeper operational issues which are then reflected in stock prices.

Datasets

There are three datasets of relevance:

1. Consumer Affairs: The dataset contains consumer complaints or reviews and ratings for over 3,300 U.S companies and product lines. History: since January 1999.

2. Better Business Bureau: The BBB serves as an intermediary between consumers and businesses in consumer disputes, as well as collecting and providing business ratings. History: since January 2012.

3. Consumer Finance Protection Bureau: This U.S. government regulatory agency provides access to its database of customer complaints relating to financial services companies. History: since January 2011.

Case Study

On October 13th 2015, Eagle Alpha's Data Analytics team published a research note on U.S. auto lenders highlighting customer service issues. Santander Consumer (SC) had the worst consumer complaint and review profile, followed by Ally Financial (ALLY). The two companies had:

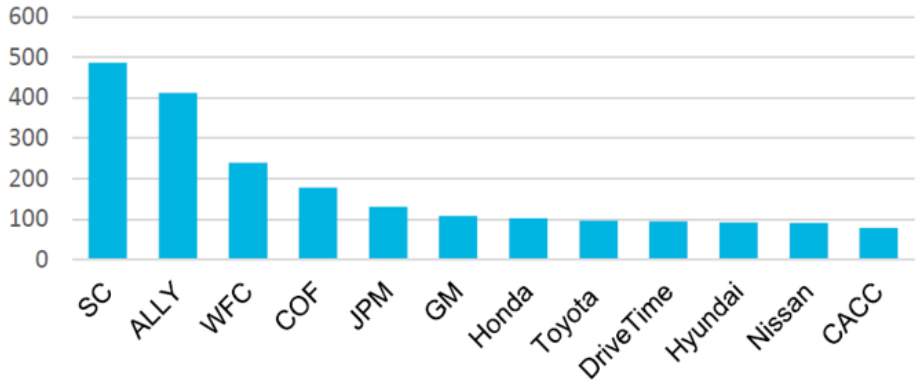
• The highest number of complaints to Consumer Finance Protection Bureau (CFPB).

• The highest number of complaints to Better Business Bureau (BBB).

• The highest number of poor reviews on Consumer Affairs website.

• Among the lowest ratings from consumer finance websites.

Figure 67: Complaints to CFPB Regarding Vehicle Loans & Leases, 2015 YTD



Source: CFPB, Eagle Alpha. These figures were high in absolute terms and in Santander Consumer's case, significantly higher relative to its loan issuance.

Santander Consumer's stock price was down 45% nine months after the publication of our report. Part of this decline is attributable to a general selloff in auto lenders as the market is pricing in a cyclical peak in US auto sales. However, Santander Consumer has declined more than the other auto lending stocks in the 9 months after Eagle Alpha issued its report (Figure 68).

Although Santander Consumer's business suffers from a number of troubles not directly related to its poor reviews, including capitalization issues, the reviews and complaints were a "canary in the coalmine" signal, supporting the idea that there is value in understanding a company's quality ratings as it can affect operations and earnings.

Figure 68: Performance of SC Share Relative to the Peer Group since October 13, 2015

Source: Bloomberg

The assessment of service and product quality of auto lenders, in this case, has proven to be indicative of deeper operational issues which are then reflected in stock prices.

39. Macro > Discretionary (Short-Term) > Sentiment Data

Key Takeaway

Analysis and scoring of central bank communications yields more accurate predictions of central bank policy versus market consensus expectations.

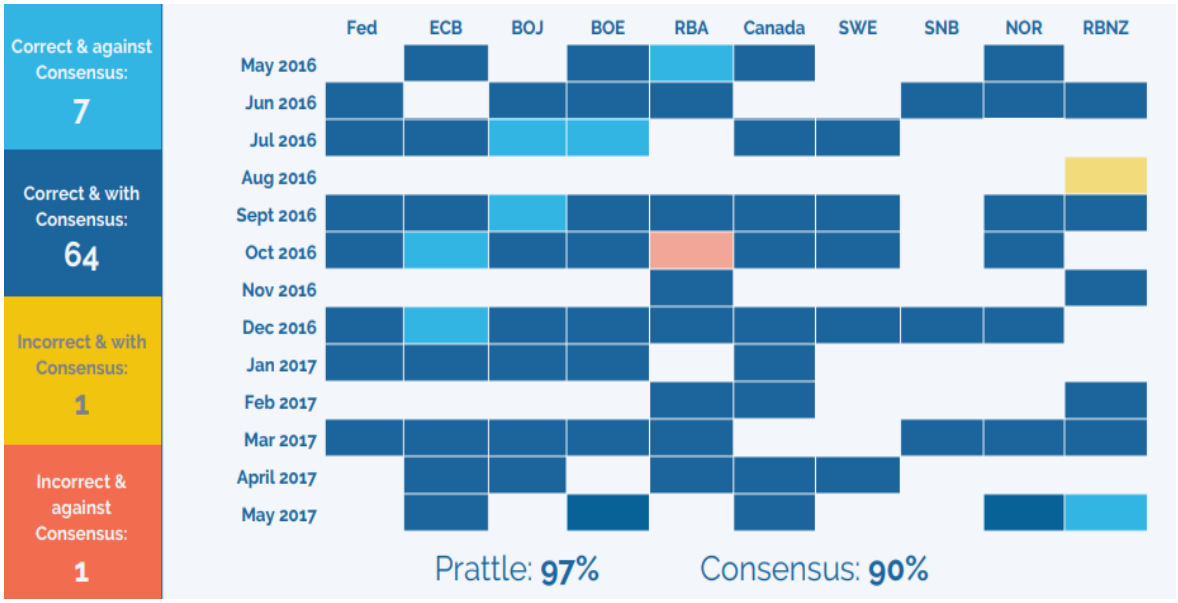
Dataset

This data provider produces analytics that predict the market impact of central bank public statements. Through machine learning algorithms, this analytics provider transforms complex communications into actionable scores.

Case Study

This data provider contends that central banks publish 177,000 words per week with market-moving potential; far too much information for a human analyst to comprehensively and objectively review. Algorithm based analysis can yield a more accurate interpretation. The chart below shows this data provider's historical track record of predicting central bank policy moves based solely on its analysis of central bank communications. Strong asset price volatility was seen in instances where central banks acted contrary to consensus expectations.

Figure 69: Central Bank Actions, Predictions And Consensus Expectations



Source: Sentiment Data Provider

40. Macro (China) > Quantitative, Discretionary > Satellite

Key Takeaway

SpaceKnow uses satellite imagery to predict China's PMI indices.

Dataset

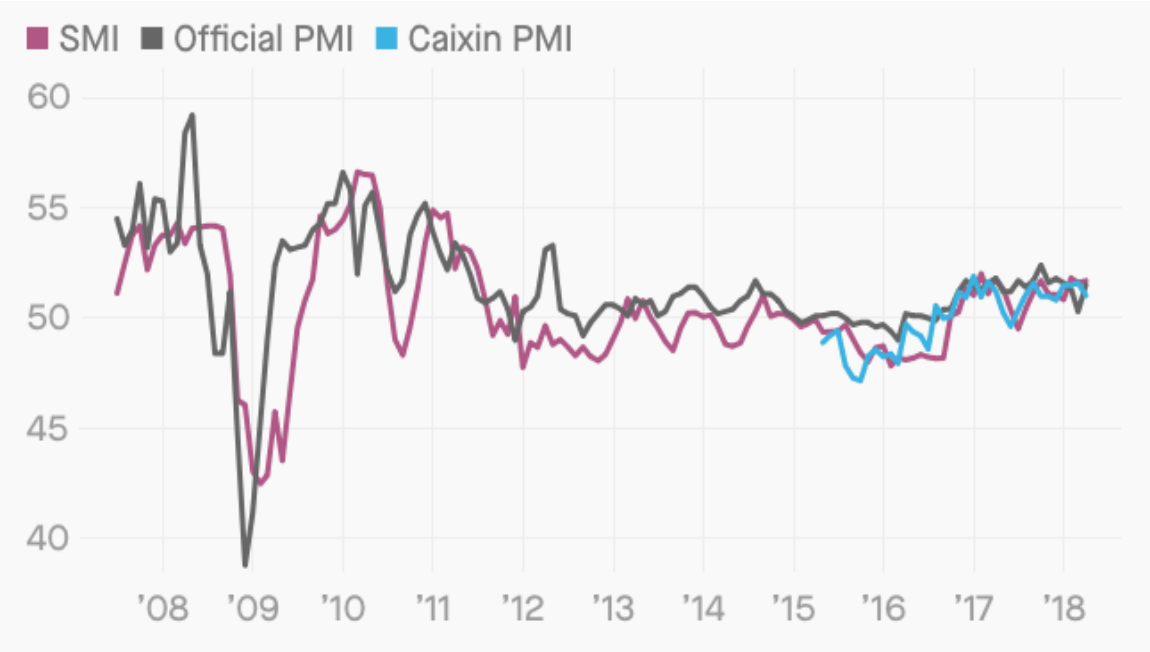
SpaceKnow was founded by a former NASA remote scientist and is a start-up backed by Silicon Valley venture capital. SpaceKnow produces its China Satellite Manufacturing Index which, according to CEO Pavel Machalek, is more reliable than other sources: "We don't conduct a survey, and our index is completely objective and automated. What we have is an independent third look at the facilities in China using a completely different methodology."

Case Study

China Manufacturing PMI and Caixin Manufacturing PMI are indices that analysts use to monitor industrial activity in China. Purchasing manager indices are based on surveys when managers are asked what they think about changing conditions. China's official PMI index focuses on state-owned enterprises and large companies, while Caixin concentrates on small and medium-sized companies.

When an index goes above 50, analysts conclude that the Chinese manufacturing sector is in expansion compared to the previous month. PMI indices are divided into sub-indices: New Orders (30%), Production (25%), Employment (20%), Suppliers' Delivery Times (15%), and Raw Materials Inventory (10%).1

Figure 70: China SMI Compared to PMI Indices



Source: SpaceKnow, FactSet

SpaceKnow's algorithms, on the other hand, are able to identify signs of industrial activity such as geographic changes, new construction objects or inventory accumulation. SpaceKnow used these methods to analyze 2.2 billion observation points gathered from satellite images of over 500,000 square kilometers spanning the period of 14 years.

An algorithm comparing images of more than 6,000 industrial facilities is used to calculate the SMI (Satellite Manufacturing Index). SpaceKnow is able to give an early read as monthly estimates are updated with a 10-day lag every Wednesday vs official and Caixin PMI indices which are published once a month with a 1-month lag

41. Macro > Discretionary (Short-Term) > Satellite Imagery Data

Key Takeaway

A relationship was identified between copper prices and estimates of copper inventories using satellite imagery data.

Dataset

Satellite data company provides investors with actionable insights on year-over-year traffic trends based on proprietary quantitative analysis of satellite data, retail parking lot imagery, store locations, and climate data.

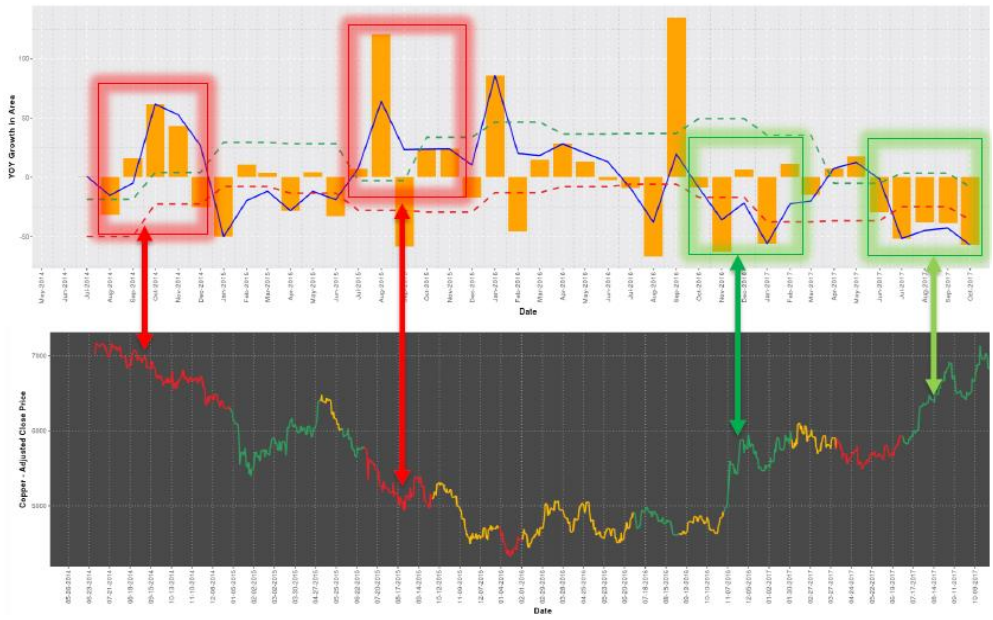
The traffic data can be used to compare the current y/y traffic growth data against the historical average of the y/y traffic growth measurements. A signal is generated when the current y/y traffic growth rate is above or below certain standard deviation levels, indicating a significant change in shopper behavior.

Case Study

During periods with significant increases in inventories (highlighted by red boxes in the figure below), the expectation is that copper prices are going to decrease.

During periods with significant decreases in inventories (highlighted by green boxes), on the other hand, the opposite is true with the expectation for copper prices to rise.

Figure 71: Satellite Imagery Metal Signals



Source: Satellite Data

42. Macro > Discretionary (Long-Term) > Online Search Data

Key Takeaway

Eagle Alpha's US Unemployment index has a 5-year correlation of 0.9 with the US Unemployment Rate, with an out-of-sample prediction improvement of 14% compared with a baseline ARIMA model.

Dataset

Google Trends is a public web tool based on Google Search that shows how often a particular search-term is entered relative to the total search-volume over time across various regions of the world. There are several advantages to using search data as a determinant of economic activity. The data is very timely, it has over 10 years of history, it is unique and offers unparalleled flexibility in terms of the variety of issues that can be analysed. The data is also generated as a by-product of people's normal day-to-day activity, as opposed to traditional survey methods which rely on individuals or firms responding to survey questions after the event. This can avoid problems associated with non-response or inaccurate responses.

Geography	Coverage	Mapped to Index	History	Collection Frequency	Delivery Frequency	Lag Time	Delivery Method
U.S	US Labor Market	Yes	Since 2006	Monthly	Monthly	1 to 14 days	CSV

Backtesting/Significance

Eagle Alpha data scientists and data insight analysts have invested three years into finding the best way to use online search data to predict economic indicators. We have devised our own proprietary methodology that leverages all relevant academic research, as well as accepted best practices in the field.

Each index is built using a rigorous process: 1) generate relevant search terms; 2) source the search volume for each term dating back to 2004; 3) clean the data and adjust for outliers and seasonality; 4) search terms are ranked by their predictive scores; and 5) final index includes a selected basket of terms, and measures co-movement of search activity with a particular economic indicator. To obtain a detailed overview of our methodology please contact us.

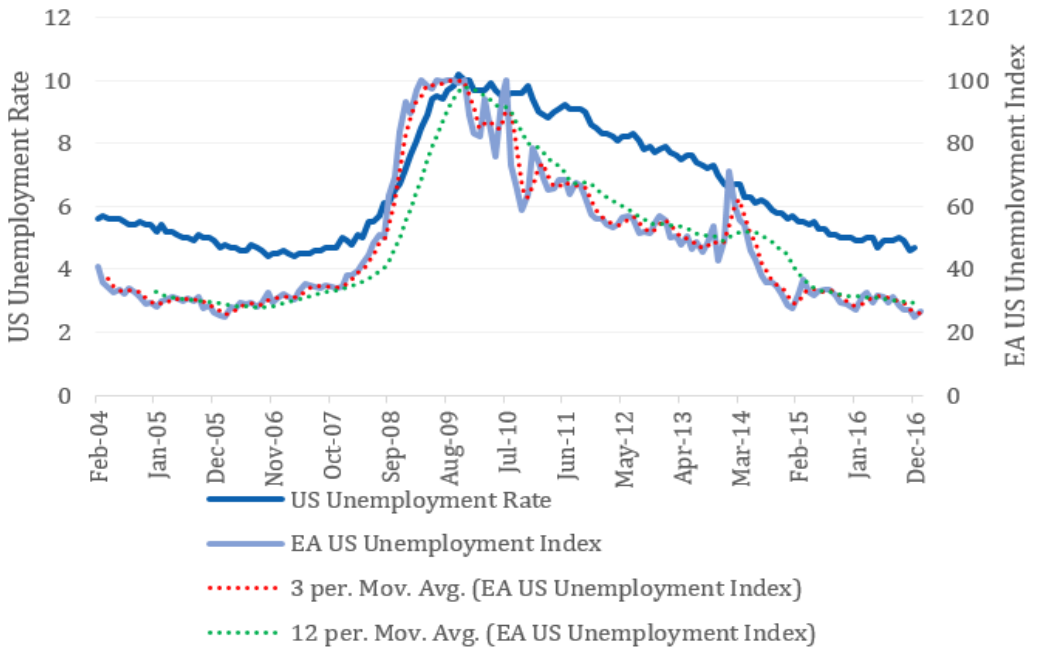
These indices are not designed to provide point estimates for macro investors, but instead add value by improving the predictive power of clients' estimation models.

Case Study

Eagle Alpha's US Unemployment Index (Figure 72) is a measure of online search activity relating to the claiming of unemployment benefits. The index has a 5-year correlation of 0.9 with the US Unemployment Rate and testing shows an out-of-sample prediction improvement of 14% over a baseline ARIMA model of unemployment over the same period.

In January 2017, the index increased by 1.8 points. This is the largest MoM change since August 2016. The 3 and 12-month moving averages mirror each other closely, which is consistent with the US economy operating close to full employment.

Figure 72: Eagle Alpha US Unemployment Index vs. Unemployment Rate (US)



Sources: Eagle Alpha, Google, BLS, Bloomberg

43. Macro > Discretionary (long-term) > Employment Data

Key Takeaway

Employment data enables more granular analysis of the labor market by sector.

Datasets

Completely unique in the industry, the job listing dataset only indexes jobs directly from employer websites. Updated daily with over 4 million jobs from more than 30,000 employers, the platform eliminates duplicate and expired job listings, as well as job pollution. From the core platform, the company has developed an array of products and services and achieved significant market traction in two primary business units: Candidate Sourcing and Job Market Data and Analytics.

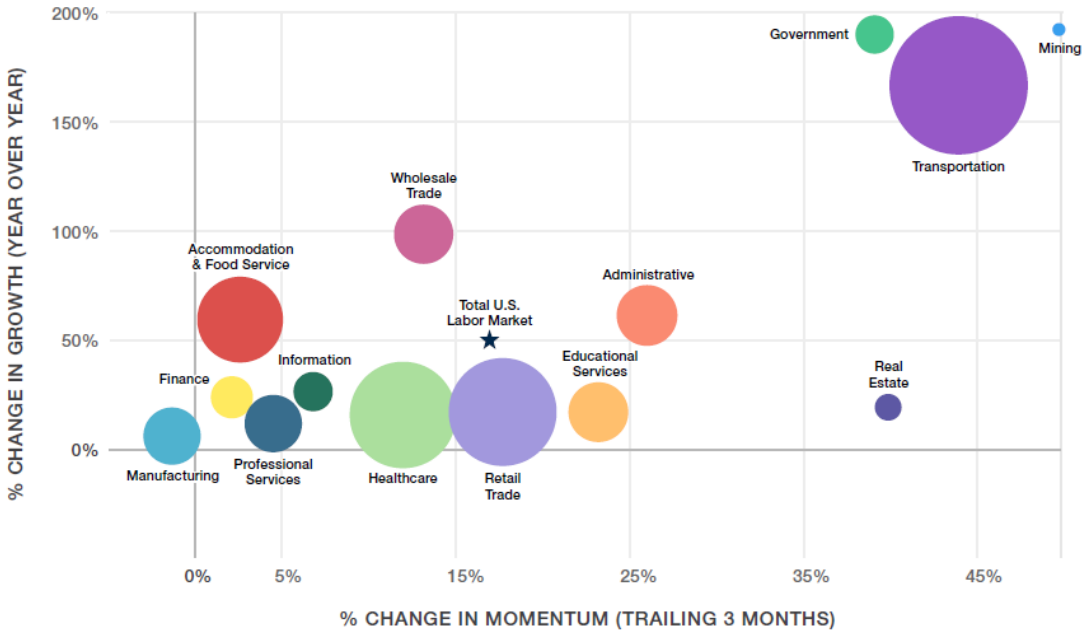
Backtesting/Significance

Backtesting of the dataset demonstrated alpha in the dataset, with "Jobs Active" producing the highest and most consistent returns. Yearly hedge returns were between 6-8%.

Case Study

Official figures by the end of August 2017 were showing low levels of unemployment in the US. However, employment data showed that not all industries were on the hiring spree. The figure below shows a sector-level look at the labor market, focusing on 12-month and three-month growth rates. Retail and manufacturing were bringing up the rear, while transportation, national security and mining were gaining momentum.

Figure 73: Change in Labor Demand by Industry



Source: Employment Data

44. Macro > Discretionary (Long-Term) > Trade Data

Key Takeaway

Statistical backtesting has proven the indicators are frequently a better predictor than street-mean estimates, while also having the advantage of substantial lead time.

Dataset

Trade nowcasting indicators provide valuable insights into international trade and industrial production. Our data partner employs big data and predictive analytics techniques, and more than 25,000 times series, to forecast trade balance and industrial production statistics. The data is also available at the level of an individual shipping port, and bespoke feeds can be provided.

Geography	Coverage	Mapped to Tickers	History	Collection Frequency	Delivery Frequency	Lag Time	Delivery Method
Global	12 Countries e.g. USA, CAN, JPA, GB, GER, TAI	No	Since 2010	Daily	Daily	1 day	API, CSV

Backtesting/Significance

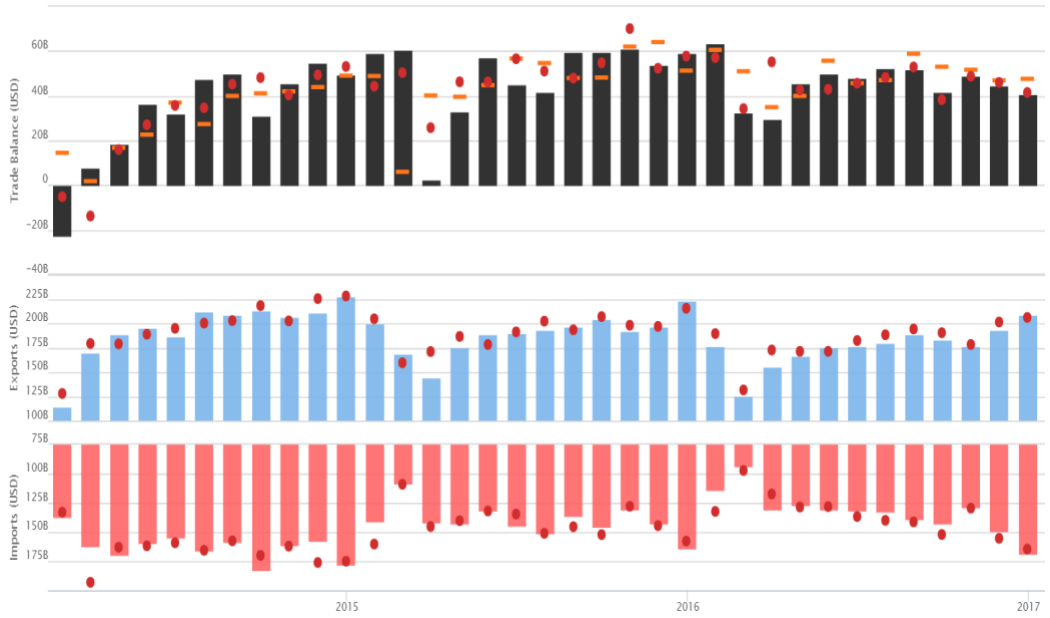
Statistical backtesting has proven the indicators are frequently a better predictor than street-mean estimates, while also having the advantage of substantial lead time. Over the last 3 years, the indicators were better than street estimates 67% of the time.

Case Study

Considering China's trade balance, Figure 74 below shows monthly values from February 2014 onwards. Red dots show the trade nowcasting forecasts, while orange dashes indicate street estimates.

In January 2017, it was reported that China's trade remained sluggish in December 2016 with exports decreasing 6.1% YoY. Imports, on the other hand, increased by 3.1% YoY vs. 6.7% YoY in November 2016.

Figure 74: China's Trade Balance



Source: Trade Nowcasting Indicators

45. Macro > Discretionary (long-term) > Trade Data

Key Takeaway

The South Korea real-time export data accurately tracked overall China exports.

Dataset

This vendor delivers high frequency and comprehensive South Korea export data. Preliminary export data includes total volume and value of all products, at all country destinations that are exported from South Korea. The dataset is created by aggregating and analyzing customs declaration forms.

Dataset Overview:

• History: Since 2003.

• Geography: South Korea export destinations.

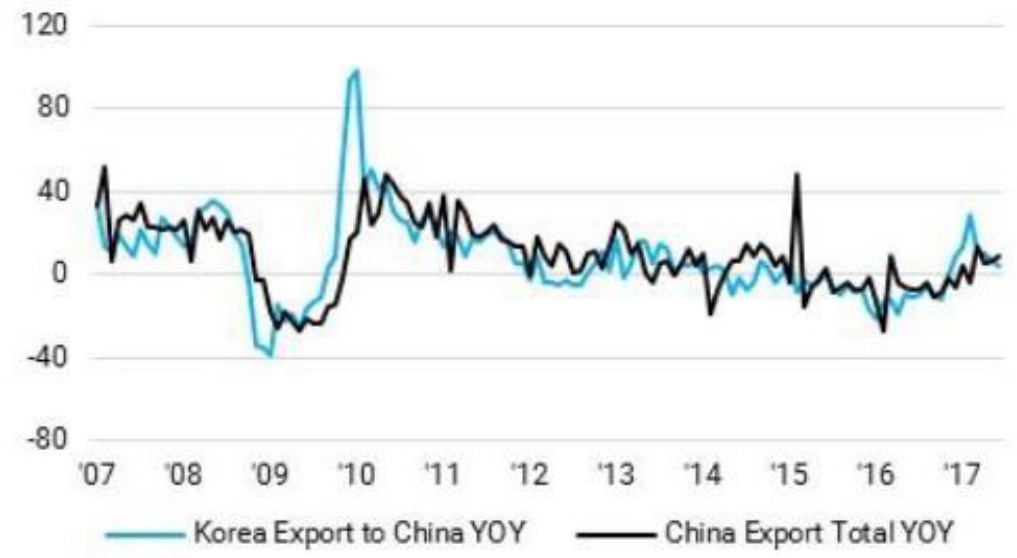
• Delivery: CSV.

• Frequency: Weekly.

Case Study

The South Korea real-time export data accurately tracked overall China exports in the 2007-2017 period.

Figure 75: South Korea Export vs China Total Exports



Source: Trade Data Provider

46. Macro > Discretionary (Long-Term) > Employment

Key Takeaway

Using payroll processor microdata improves forecast accuracy of an aggregate labor market activity measure.

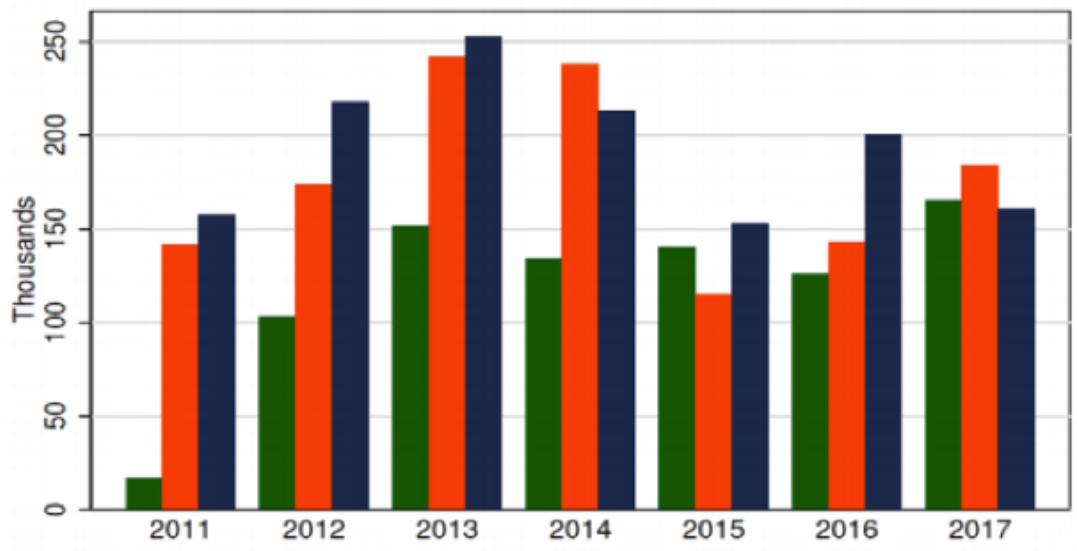
Dataset

The vendor is one of the largest providers of business processing and cloud-based solutions – including payroll, talent management, human resource management, benefits administration and time and attendance – to employers and automotive dealerships around the world. History: since January 2014.

Case Study

Members of the Federal Reserve Board evaluated value of the payroll processor microdata and how it could be used to measure aggregate labor activity. They highlighted that the timeliness and frequency of the payroll data improved forecast accuracy for the official figures (Current Employment Statistics published by BLS, the Bureau of Labor Statistics). Some of the results are presented below.

Figure 76: Real-Time Data (Blue) vs First BLS Print (Green) vs Latest BLS Figures (Orange)



Source: BLS, Employment Data Provider.

47. Macro > Discretionary (Long-Term) > Sentiment

Key Takeaway

Quantitative metrics of emotional content in market narratives may complement other indicators and analysis in helping to gauge systemic risk.

Datasets

Three data sources were analysed: 1) Internal Bank of England daily commentary on market news and events, i.e. 26 documents per month on average from January 2000 until July 2010; 2) Reuters' news wire articles in the United Kingdom, consisting of over 17 million English news wire articles; and 3) Broker research reports, which are from 14 brokers from June 2010 until June 2013, approximately 100 documents per month.

Case Study

The Bank of England researchers analysed unstructured text-based market data to explore how narratives and sentiments drive developments in the financial system. They concluded: "We find that changes in the emotional content in market narratives are highly correlated across data sources. They show clearly the formation (and subsequent collapse) of very high levels of sentiment – high excitement relative to anxiety – prior to the global financial crisis."

The researchers used a word count methodology (see Figure 77 below), which classifies words as representing anxiety or excitement, to calculate an emotional statistic of sentiment (Relative Sentiment Shift or RSS). RSS equals the difference between excitement words and anxiety divided by the total number of characters in a text.

Figure 77: Word Count Methodology Used

Table 1: Emotion dictionary samples

Anxiety		Excitement	
Jitter	Terrors	Excited	Excels
Threatening	Worries	Incredible	Impressively
Distrusted	Panics	Ideal	Encouraging
Jeopardized	Eroding	Attract	Impress

Source: The Bank Of England

Further robustness tests showed that the sentiment metrics have some predictive power for standard aggregated measures of consumer confidence, market volatility, and some relevant but more theoretic measures of uncertainty from the literature exploiting text-based information. Furthermore, their measures also contain useful information in explaining economic data in a simple VAR model. Therefore, the authors of this paper believe that these quantitative metrics of emotional content in market narratives may complement other indicators and analysis in helping to gauge systemic risk.

48. Macro > Discretionary (Long-Term) > Trade

Key Takeaway

Trade data was used to track the surprising macro-economic turnaround for Brazil in 2016.

Dataset

Trade nowcasting indicators provide valuable insights into international trade and industrial production. Our data partner employs big data and predictive analytics techniques, and more than 25,000 times series, to forecast trade balance and industrial production statistics. The data is also available at the level of an individual shipping port, and bespoke feeds can be provided.

Geography	Coverage	Mapped to Tickers	History	Collection Frequency	Delivery Frequency	Lag Time	Delivery Method
Global	12 Countries e.g. USA, CAN, JPA, GB, GER, TAI	No	Since 2010	Daily	Daily	1 day	API, CSV

Backtesting/Significance

Statistical backtesting has proven the indicators are frequently a better predictor than street-mean estimates, while also having the advantage of substantial lead time. Over the last 3 years, the indicators were better than street estimates 67% of the time.

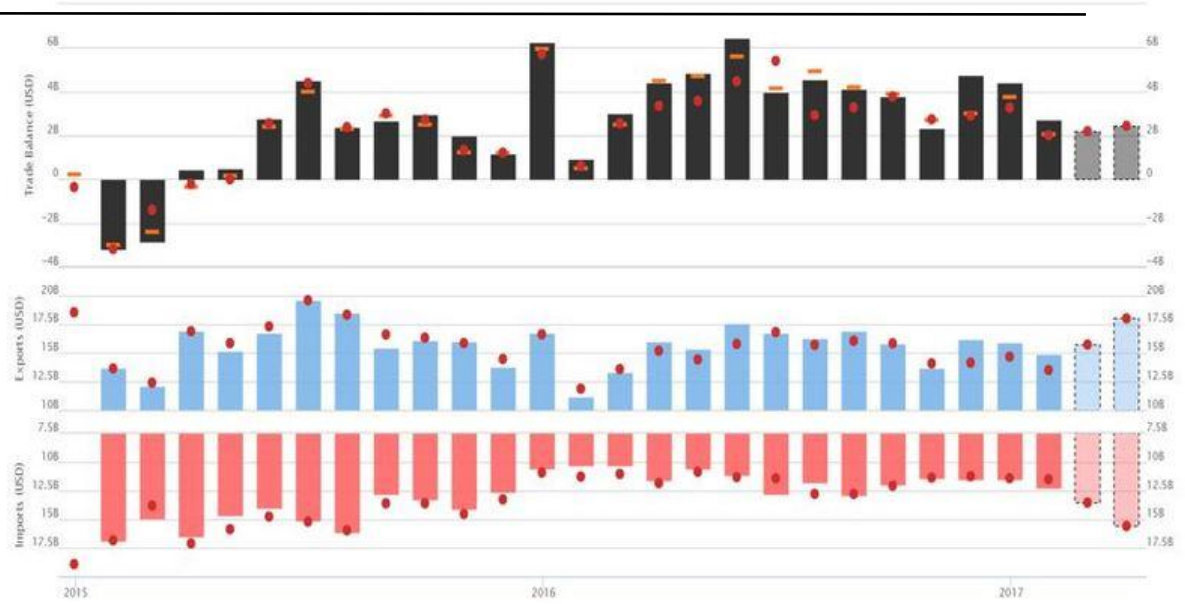
Case Study

One of the most surprising macro-economic turnarounds in 2016 occurred in Brazil. In April 2016, Brazil's economy was in a free-fall, exacerbated by political turmoil. However, the trade nowcasting indicators showed that there were hints of moderation in the steep drop in shipments into the country.

The nowcasting indicators appeared to be forming a bottom and pointed to signs of economic relief in May 2016. In July, the data was used to increase Brazil's industrial production forecast and in August the conclusion was made that the country's economy was exhibiting a turnaround.

Finally, in October 2016, data on Brazilian consumer goods imports signalled the return of the Brazilian consumer. In Figure 78 below, red dots show the trade nowcasting forecasts, while orange dashes indicate street estimates.

Figure 78: Brazil's Trade Balance



Source: Trade Nowcasting Data Provider

The Brazilian stock market surged as a result, increasing 66% in 2016 as measured by MSCI Brazil index.

49. Credit > Discretionary (Long-Term) > Pricing Data

Key Takeaway

Online property listings data was used to research buy-to-rent investors which led to the subsequent clamp down on mortgage lending.

Dataset

The case study below is based on an online property listings dataset that covers 80% of UK sales. This dataset is more timely than any publicly available dataset. Our partner's data can be actionable in various ways, such as analysing property deal flow and volume, tracking liquidity of the property market, tracking the liquidity and availability of secure consumer credit, analyzing consumer confidence around envisaged indebtedness, and advanced indicator of strength or weakness in housing related consumer discretionary stocks as well as telecom and utilities.

Geography	Coverage	Mapped to Tickers	History	Collection Frequency	Delivery Frequency	Lag Time	Delivery Method
U.K.	80% of U.K home sales	No	Since 2005	Daily	Daily	1 Day	FTTP, CSV

Case Study

Bank of England analysts demonstrated their research in several papers: 'How much do investors pay for houses?' (September 2015)[30], 'Five facts about buy-to-let' (July 2015)[31], 'Chance favours the prepared mind' (July 2015)[32]. These articles shed a light on the Bank's methodology and how various micro datasets were linked in order to gain valuable insights into UK housing market trends.

Buy-to-rent investors – known as buy-to-let (BTL) in the UK – were becoming large players in the UK housing market. Housing stock held for private renting went from 9% of the total stock in 2000 to 19% in 2013 (Bracke, 2015). The Bank of England decided to investigate whether BTL investors drove the housing prices up and could lead to macro instability.

The online property listings dataset was combined with Product Sales Data (PSD) and Land Registry's Price Paid data. The PSD data provided by the Financial Conduct Authority was used to gather details of loans for house purchases.

The combination of these datasets allowed the Bank of England to conduct a more precise analysis and get closer to the understanding of the individual agents' decision-making. Some valuable insights are presented below.

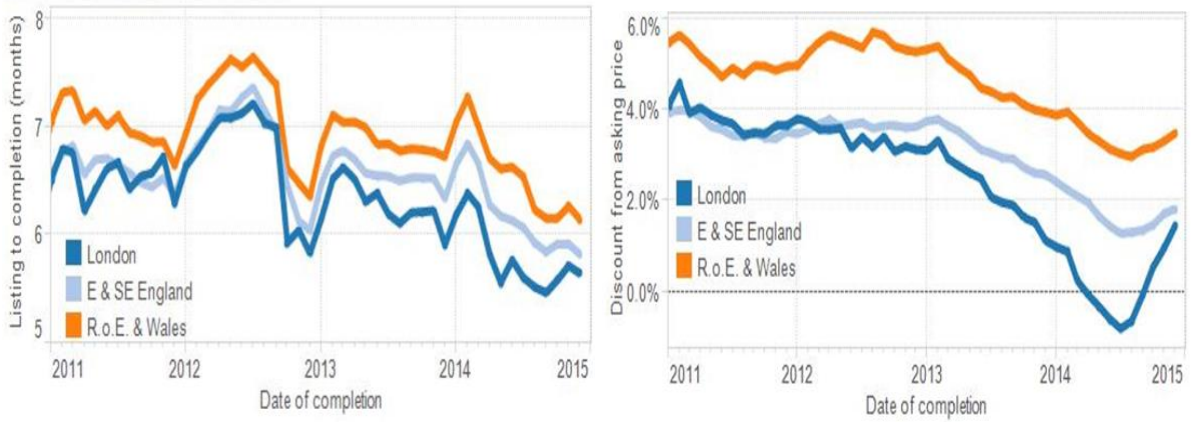
A decrease in average time from listing to completion of a sale highlighted increased purchasing activity.

[30] Source: 'How much do investors pay for houses?', September 2015

[31] Source: 'Five facts about buy-to-let', July 2015

[32] Source: 'Chance favours the prepared mind: What linked micro data can tell us about the housing market', July 2015

Figure 79: Months Listing to Completion & Discount From Asking Price



Sources: Online Property Listings Dataset, Land Registry Price Paid, ONS Postcode Directory

Buyers were able to negotiate and obtain 4-5% discounts in 2011-2012. However, the analysis showed that buyers lost some of the bargaining power as the speed of transactions increased. In London, the average discount even turned negative at one point.

Bracke (2015)[33] concluded: "The data show that BTL investors can accelerate the time it takes to sell a property, and BTL discounts are the implicit compensation for this contribution. However, investors' ability to `grease the wheels' of the housing market becomes limited when the market is already performing well. This is precisely when financial stability concerns become most important."

The Bank of England then used the above analysis to further investigate the mortgage-financed part of the housing market as half of all BTL transactions were supported. It then decided to take action and clamp down on BTL lending: "The PRA, the Bank's regulation arm, is concerned that changes to mortgage interest tax relief for landlords will strain buy-to-let borrowers and that only a few lenders include this risk when assessing mortgage applications."

[33] Source: 'How much do investors pay for houses?', September 2015

50. Credit > Discretionary (Long-Term) > Credit Risk Data

Key Takeaway

Testing showed that the probability of bankruptcy within 12 months ranges from 10% to 50% when a company gets a stress score of "1".

Dataset

Credit score data provider delivers predictive stress scores for credit, supply chain and financial professionals. Credit ratings are prepared for over 57,000 public companies worldwide and can be used to analyze equities with high bankruptcy risks.

Aggregate crowd-sourced usage data from the vendor's subscribers, credit managers and supply chain professionals from Fortune 1000 companies, is incorporated during the construction of stress scores. The provider found distinct behavioural patterns when its subscribers are concerned with certain companies and investigate them more closely.

Backtesting/Significance

Monthly credit scores have been proven to be highly accurate when predicting corporate failures. Testing showed that the probability of bankruptcy within 12 months ranges from 9.99% to 50% when a company gets a score of "1".

Case Study

On August 31st 2016, Hanjin Shipping, the world's seventh-largest container carrier at the time, filed for bankruptcy protection as it could not renegotiate its debts. On September 12th 2016, it was estimated[34] that $14 billion worth of cargo was stuck aboard Hanjin ships and the company lost a third of its market value in two weeks.

On 21st September 2016, the vendor published a case study with the post-filing analysis of Hanjin Shipping's bankruptcy. Figure 80 below shows that the credit score for Hanjing dropped to "2" in December 2015 and "1" in February 2016 well in advance of the bankruptcy filing.

[34] Source: 'A By-the-Numbers Look at Hanjin Shipping's Collapse', September 2016

Figure 80: Monthly Average Credit Scores

Business Name	2015 SEP	2015 OCT	2015 NOV	2015 DEC	2016 JAN	2016 FEB	2016 MAR	2016 APR	2016 MAY	2016 JUN	2016 JUL	2016 AUG	2016 SEP
2GO Group Inc	6	6	7	7	7	6	6	7	7	7	7	7	7
Abetrans Ltd.	3	3	3	3	3	3	3	3	3	3	3	3	3
AP Moeller Maersk A/S	9	9	9	10	9	9	9	9	9	9	9	9	10
Arpeni Pratama Ocean Line Tbk PT	2	2	2	2	2	2	2	2	2	2	2	2	2
Associated British Ports Holdings Plc				9	9	9	9	9	9	9	8	8	8
Atlantska Plovidba dd	4	4	4	4	4	4	4	4	4	4	4	5	5
Avance Gas Holding Ltd	7	8	8	8	8	7	7	7	5	4	3	3	3
Barska plovidba A.D. Bar	4	4	4	4	4	4	4	4	4	4	4	4	4
Bollore SA	6	6	6	6	6	6	5	5	5	5	4	5	5
Buana Listya Tama Tbk PT	6	6	6	6	6	6	4	4			5	5	6
China COSCO Holdings Company Limited	4	4	4	4	4	4	4	3	2	2	2	3	4
China Merchants Energy Shipping Co., Ltd	5	5	6	6	6	6	6	6	6	5	6	6	6
Chu Kong Shipping Enterprises(Grp)Co Ltd	7	7	7	7	7	7	7	7	7	6	7	8	8
Companhia Doca de Imbituba	3	3	3	3	3	3	3	3	3	3	3	3	3
Express Kenya Limited	3	3	3	3	3	3	3	3	3	3	3	3	3
FUSHIKI KAIRIKU UNSO CO., LTD.	5	5	5	5	5	5	5	5	5	5	5	5	5
Hanjin Shipping Co Ltd	3	3	3	2	2	1	2	1	1	2	1	1	1
Kirby Corporation	9	9	9	9	9	9	10	10	10	10	10	9	9
OT Logistics SA	6	6	6	6	6	6	6	6	6	6	6	6	6
South Logistics JSC	7	7	7	7	7	7	7	7	5	5	5	5	5

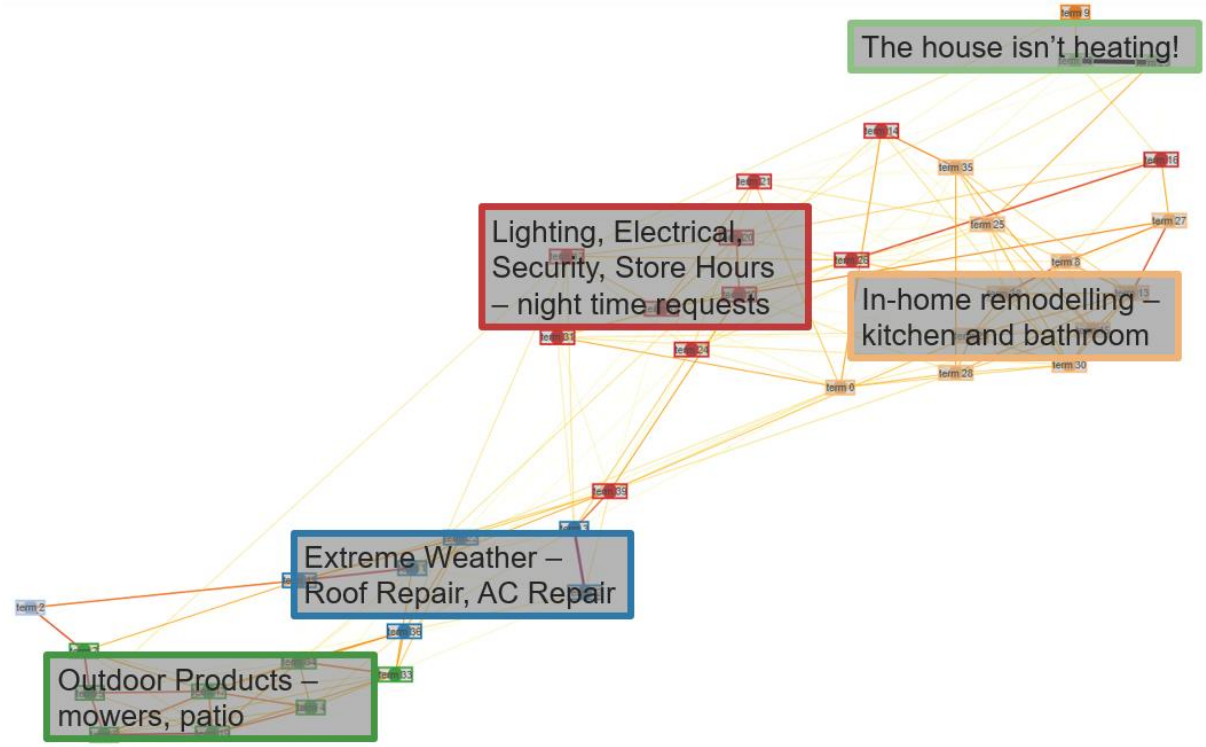
Source: Credit Risk Data Provider

Section 6: Quantitative Approaches to Online Search Data

Eagle Alpha specializes in the use of online search data for fundamental discretionary use cases (See Section 5). Another area that Eagle Alpha has been targeting involves utilizing quantitative methods to derive granular information about interest in specific company product categories and interest in a company's stock.

Using Clustering and Stochastic Neighbor Embedding to Understand Customer Search Behavior

Figure 81: Search Term Relationships for Home Depot

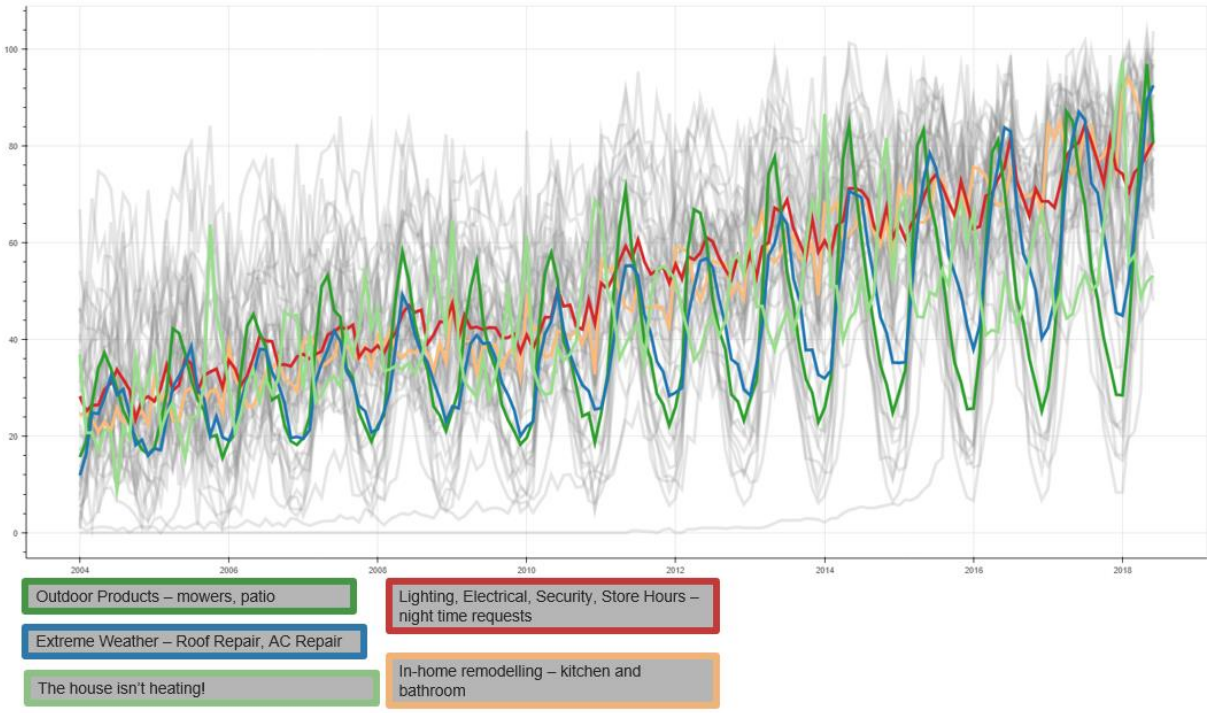


Profiling the target customers of a business is an essential part of understanding the fundamentals of a company. Determining how well the business is performing in different target customer groups is a challenge without having specific information in product spend breakdown.

Online search trends can be advantageous in determining the natural groups of customers a business maintains, as well as how interest from those groups changes over time. Relating the online search trends over time using a time series covariance method, unsupervised clustering techniques can be applied to group search terms that have similar trend behaviour. In the above example, the affinity propagation clustering technique is used as it does not rely on a pre-determined number of clusters when optimizing the groupings.

The estimated time series covariance between the search terms can also be used to visualize the high-dimensionality inherent in time series data. Using t-distributed Stochastic Neighbor Embedding, the relationship between the time series trends can be demonstrated in a digestible form. From the above analysis of common product search terms for the Home Depot (NYSE:HD), search terms for outdoor products shows natural separation and clustering apart from contractor remodelling projects, denoting two distinct customer search groups.

Figure 82: Differentiation of Home Depot's Customer Groups in Time Series Trends



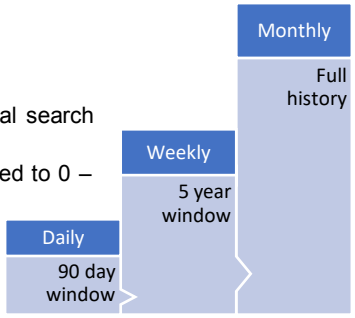
The time series trend of the groups illustrates the differentiation of customer behaviour. Outdoor product interest rises in the springtime, searches for roof and AC repairs peak in the late summer months, while search terms related to heating system repairs occur in the winter months. While these trends by themselves are relatively intuitive, growth or decline in interest in a specific category relative to other categories (or even other companies) can be highly indicative of company performance in that customer category.

Scales of Search: Multi-frequency and Relative Scale Search Term Processing

The primary source of online search data used at Eagle Alpha, Google Trends, supplies trend data at varying frequencies depending on the time span requested. For monthly data on a search term, Google allows for its entire history to be queried at one time, however for higher frequency trend data, the restrictions (5 years for weekly, 90 days for daily) can make deriving useful information challenging.

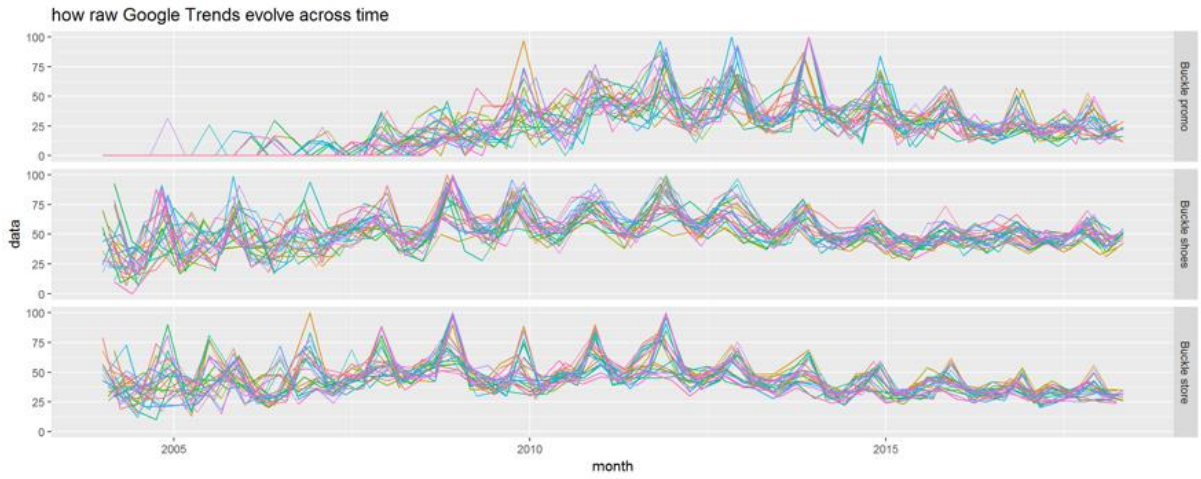
This is complicated by a few other factors defined by Google:

- The trend data, at any frequency, is only a dynamic sample of the total search volume.
- The scale of the trend, for a full or subset time span, is always normalized to 0 – 100.
- Relative scale trends between terms is limited to five terms at a time.



Eagle Alpha has developed a methodology for handling the sampling problem by using multi-sample averaging of identical time spans for the same term. This process involves downloading multiple samples of a search term on subsequent days (the sub samples rotate by day), and then solving for the scale and offset that best aligns the sample over the history with the previous samples.

Figure 83: Search Term Trends Vary Considerably Among Downloads



Handling the restriction on time span with increased frequency requires a complex approach utilizing multi-frequency scaling. Since the monthly data contains the longest time span with only a single value of 100, this data is utilized as the master source that data at higher frequency is scaled to and then concatenated together from. Multi-frequency scaling, offsetting, and averaging is required to deal with high frequency data to reduce the volatility of the high frequency trend data.

Analysing the relative scale factor among various search terms of a single company or among multiple companies is another useful technique. In this query technique, Google returns only one data point in time among all the search terms queried that is scaled to 100. However, due to the limit of five terms per query as well as the characteristic of sub sampling of the total search volume, the relative scales can be a challenge to utilize effectively. Eagle Alpha has developed methods for averaging the relative scale factors using a chaining approach that derives estimated global scale factor among any number of terms.

Reach out to the Data Science team at Eagle Alpha to learn more about our work with online search data and how we can help integrate it into your investment strategy.

Section 7: Overview of Eagle Alpha

Eagle Alpha provides alternative data solutions to the buyside.

Complimentary solutions are events, this use cases report and altdata.tv.

We execute a range of bespoke projects for clients. There are typically four types of projects: strategic investment integration projects, research questions, data assignments and alternative data driven models.

The core modules of Eagle Alpha are data sourcing, data analytics and data forum:

- Data Sourcing: this solution keeps clients on top of all the alternative datasets worldwide.
- Data Analytics: this solution enables discretionary managers to work with a variety of alternative datasets through a single dashboard platform.
- Data Forum: this is the leading alternative data industry group focused on legal & compliance, increasing efficiencies and driving best practices.

Eagle Alpha is a recognized leader in the alternative data space. For example, white papers by Citi, JPMorgan, BoA and Deloitte have all profiled the company.

This page was intentionally left blank.

How do you want to consume alternative data:

1. Raw and/or semi-processed data?
2. Use tools, indicators and analyst insights?
3. In a curated form?

Join the leading industry group

Since 2012 Eagle Alpha has provided a full service solution that enables asset managers to obtain alpha from alternative data

Eagle Alpha

SPONSORED BY

SMI | STANDARD MEDIA INDEX