



MICHAEL RECCE
Chief Data Scientist

The Data Science Revolution

How the new richness and accessibility of data, and advances in data science, are enhancing both quantitative and traditional fundamental investment research—and sparking a revolution in active management.

So-called “big data”—the residue of information that we all leave behind as we buy things, sell things, browse Main Street and the internet, use our smartphones and generally live our modern lives—is proliferating. At the same time, advances in cloud computing, machine learning and artificial intelligence allow us to extract coherent, strategic insights from these digital residues. Combined, as data science, they have the potential to be a richly enhanced source of information about our world—information that is deeper and more detailed than we have ever had before, and yet also broader and more comprehensive.

NEUBERGER	BERMAN
-----------	--------





The Data Science Revolution is upon us. Harnessing its potential is the next big challenge for the active asset management industry.

Five Things We Believe

- ① Big data can be useful; data science can make it powerful
- ② Data science makes our picture of the world both more detailed and more comprehensive
- ③ Data science brings extra breadth to fundamental research and extra depth to quantitative methodologies
- ④ Data science enhances the existing information flow between fundamental and quantitative research teams
- ⑤ A centralized research platform, feeding into both fundamental and quantitative strategies, can get the most out of a data science capability

Spreading from its origins in technology and retail, data science is now making waves in the information- and research-driven world of finance and investment.

Here, we look at what big data means in this new context. With real-world examples, we explore why data science informs rather than replaces traditional investment research. We describe the importance of information flows between quantitative and traditional fundamental research and how they can be enhanced still further with data science. In fact, we argue that the true power of data science lies there—which has implications for the way we believe a modern active management business should structure its research efforts.

Amid the hype, many will forget that big data is of little use without the combined insights of data scientists and investment professionals. They will fail to organize their research efforts to get the most out of it. At the same time, it is important to be sensitive to data privacy concerns and continue to monitor the evolving regulatory and legislative landscape. Ultimately, we believe those who engage with big data seriously and ethically will find it transformative.



Data Science: More Depth, No Loss of Breadth

In 2018, we addressed what we regard as the most prominent long-term themes shaping the economy and the investment industry. One theme we identified was that “Big data will define our economic ecosystem and transform investing.”

We pointed to the European Union’s Markets in Financial Instruments Directive (MiFID II). Among other things, this bans brokerages from giving research away to asset managers as a benefit for trading securities through them.

“That is likely to accelerate the ongoing cutback in securities-research resources within sell-side institutions... We expect more research to be conducted within asset management firms... [and believe that active managers will seek] more efficient ways to gather increasingly granular and timely information about how companies are performing, often in a raw form that can be sliced-and-diced to generate genuinely proprietary insights. They will be able to do this because, for the first time, that information exists and is accessible. It is called big data, and we think it will be the battleground for alpha over the coming years.”¹

Big data is spreading everywhere in the modern, networked world, and advances in cloud computing, machine learning and artificial intelligence are enabling us to extract coherent, strategic insights from it.

Think about how we used to learn about a flu outbreak, for example. Sick people would go to the local doctor, who would notice an uptick in flu sufferers and speak with colleagues who were seeing the same thing. Today, clusters of Google searches can reveal the same dynamics many months or even years before these things would become evident in the old way.

Worldwide, users visit a new web page a million times per second. Just like the Google searches for “aching limbs and sore throat,” these visits carry information about what interests us at one moment in time. That is how the static, print media-like online advertisements of 2010 quickly progressed to the dynamic, personalized ads that we are familiar with today. Similarly, in the analog world we had an inkling that people were more likely to go shopping when the weather was nice. In the big data world of credit card transaction records, we can find out what has been bought, where it was bought and how much it cost.

Importantly, data science can give us a deeper and more detailed picture of the world without giving up any breadth or comprehensiveness.

¹“2008, 2018, 2028: The Dissolving Divides That Will Shape The Post-Crisis Investment Era” (September 2018) at <https://www.nb.com/pages/public/global/four-trends-for-the-next-ten-years.aspx>



Data Science and Asset Management: Five Examples of Relevant Big Data

A wide range of alternative data is potentially useful in the investment context. It includes “hard” data such as supplier payments, prescriptions and healthcare insurance claims, and laboratory results. But there is also rich information in “softer” data such as news and social media keywords or satellite imagery. Satellite photos can tell us how much ore is being removed from a remote strip mine, how much heat is being emitted from local smelting plants and how many workers are parking their cars in the parking lot, for example.

We find that the following five data sources—which can include numbers, written text, spoken words, digital information and pictures—are the most prominent:

① Credit/debit card and bank account transaction data

- Transactions include date, amount and description
- Bank accounts include deposits, such as paychecks
- Detailed geographic information (to the individual store level)
- Detailed demographic information (age, gender, income)
- Online vs. in-store sales (in most cases)

“Are men really going to buy clothes from Lululemon?” [READ MORE](#)

② Online transaction data

- SKU-level data on Amazon purchases from 2014
- 1.3m worldwide users covering approx. 2.7m unique brands and 28m unique products

“Are Amazon’s new batteries eating into Energizer’s market?” [READ MORE](#)

③ Internet content and search data

- 32 web content datasets from 2015 including:
 - Social media engagement on company home pages
 - Store location reference data
 - Car inventory via leading car sales websites
 - App ratings (Google Play Store and Apple Store)
 - Company discounts
- 100m worldwide users’ Google Search terms from 2014
- ~50m unique keywords on a daily basis
- Click timestamp, type (organic or paid), geographic information

“Are employees buying into the big changes at Procter & Gamble?” [READ MORE](#)

④ Online job post data

- 4m worldwide (70% of U.S. total job postings)
- Daily collection from 32K+ companies’ websites (8000+ public companies)
- Data since 2007, detailed job descriptions from 2014
- Job level, category, region

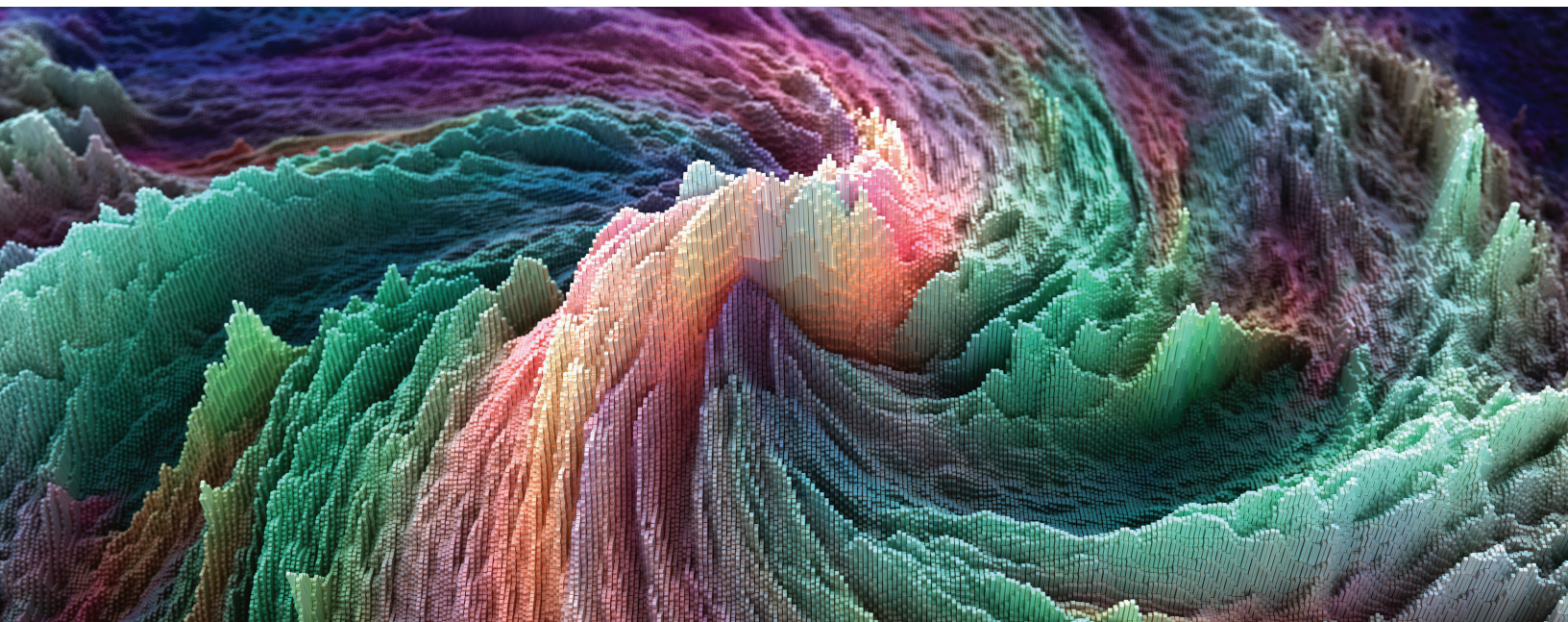
“How can I tell if companies are investing in the post-capex era?” [READ MORE](#)

⑤ Conference call transcripts and filings

- Earnings and acquisition announcement calls
- Monthly update calls (e.g. sales)
- Company conference presentations
- 1500+ U.S. public companies from 2010 to present
- All corporate SEC filings

“Why do I not want to hear the words ‘very optimistic’ on an earnings call?” [READ MORE](#)

While rarely decisive, these data can provide another corroborative layer to the insights gleaned from traditional data sources and fieldwork.



Getting the Most out of Big Data: The Real Alpha Is Not in the Data, but in the Data Science

Alternative data become most useful when they are processed with data science techniques. A swatch of credit card transactions is just a crowd of individuals buying things: it takes a lot of computing power, and often machine-learning techniques, to identify coherent subsets of that sample that are tractable to market and security analysis.

That is not to say that raw, individual data points are of no use at all. Securities analysts have been gleaning actionable insights from individual earnings call transcripts for decades. The ability to apply natural language processing to them at huge scale over time, however, transforms them into an entirely new type of resource. Raw or individual data points are like pebbles on the beach—there today, perhaps washed away tomorrow. But data science applied to big datasets can reveal the shifting contours of the beach itself. Big data can be useful in the short-term context of tactics and trading; data science moves it to the longer-term context of strategy and investing.

Moreover, because so much of the value of alternative data is realized through the data science techniques applied by individual asset managers, it is not unusual for different investors to mine the same data for entirely different insights—which, in our view, substantially reduces the potential for rapid “alpha decay.”

The first major task of any data scientist is to determine which data have the potential to be useful.

Usefulness can be limited for many reasons. A dataset may simply be of irretrievably poor quality. It may be very interesting, but immaterial to any investment question. It may represent excellent value for one of our competitors and their investment focus and style, but not for us. A hedge fund may wish to pay extra for the freshest possible data, while we, as a long-term investor, can glean the strategic insights we need from less costly data that are two or three quarters old. Neuberger Berman’s data scientists estimate that only 5% of the swatches it sees has the potential to inform our research effort.

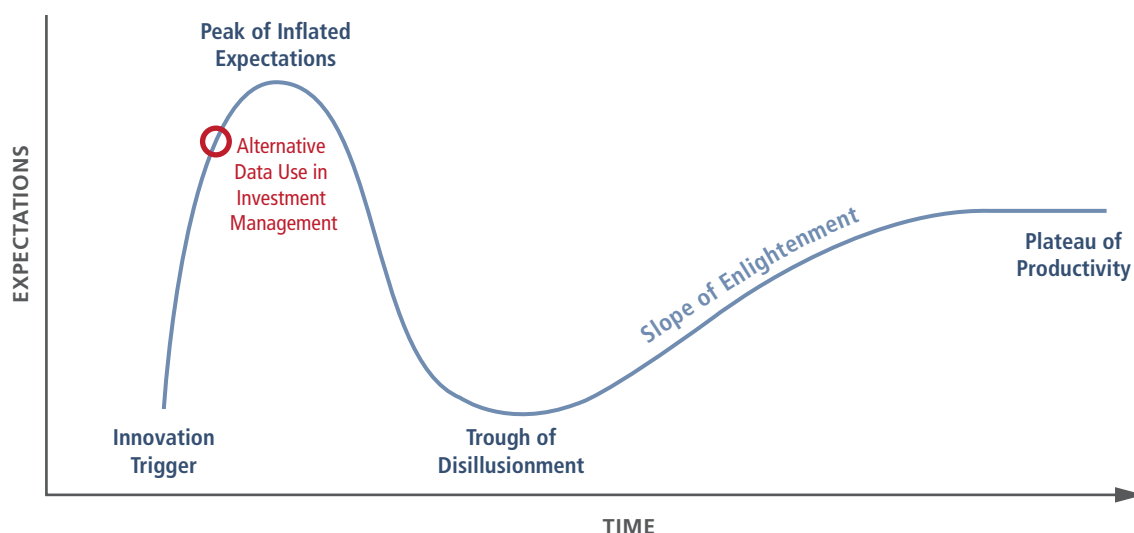
The next major task is to take potentially useful data and make them actually useful. That requires the development of clear research hypotheses—which is where existing investment teams have a critical role to play in getting actionable results from alternative data. It also requires the cleaning and interpretation of raw datasets. Having the expertise and the infrastructure to do that is clearly vital, but so is being an engaged, responsive and demanding consumer of data, rather than a passive one.

For example, it is clear that there is good information in the number of cars pulling into a store. One raw data sample we acquired appeared to show a large store getting peak customer traffic at 8:00 a.m. Oddly, that was two hours before the store opens. We called the data provider to ask whether they were mistakenly counting cars traveling on a nearby highway. At first, the data provider rejected this suggestion, insisting that vehicles travelling faster than 20mph were excluded. However, after further investigation, we established that, in the morning rush hour, this particular highway slows to a crawl—enough to place all of its traffic in the data set.

The engaged and active approach to alternative data is important for its development in the investment management context. Despite the buzz that big data is generating, achieving this engaged and active approach to using it is far from easy. It is challenging to attract the right mix of investment and data science skills—especially when we are competing for it against academia and the world's biggest technology firms.

That is why, to borrow from the “Hype Cycle” concept put forward by the technology consultancy Gartner, big data is enjoying its initial surge in our industry, but it is very likely to peak and begin a steep decline into disillusionment once the first movers realize that alternative data can be opaque or even misleading without the right investment in people and infrastructure to make sense of it. The recovery will happen only after those who are not wholly committed to data science have dropped out.

Alternative Data Use in Investment Management: The Gartner “Hype Cycle”



Source: Gartner, “Understanding Gartner’s Hype Cycles” (August 2018), at <https://www.gartner.com/en/documents/3887767>, Neuberger Berman.

Getting the Most out of Big Data: Beyond “Quantamental”

At Neuberger Berman we believe that making the most out of alternative data is not just a matter of investing in data science, but also of integrating data science more fully into the traditional investment management research process, both quantitative and fundamental.

What do we mean by that?

It helps to think about the characteristics of alternative data and data science alongside the characteristics of traditional, fundamental security analysis and the characteristics of factor- or risk premia-based quantitative investing.

Fundamental analysts’ knowledge is very deep, but comparatively narrow: they know a lot about a small number of companies. Quantitative investors’ knowledge is very broad, but comparatively shallow: they know a few key things about every company in their investment universe.

As we have seen already, data science has the power to give us a deeper and more detailed picture of the world without giving up any of its breadth and comprehensiveness. Data on the performance of every product from every competitor company in the marketplace can bring extra breadth to fundamental research. Data that corroborate or challenge signals from traditional metrics and ratios, or fill gaps in company reporting, can bring extra depth to quantitative investing.

This, in turn, enhances the existing capacity for quantitative and fundamental processes to inform one another. Data science can therefore take us beyond the crudest kind of “quantamental” approaches, which merely lump these two research efforts together, to a more complete integration of research and flow of information across an organization.

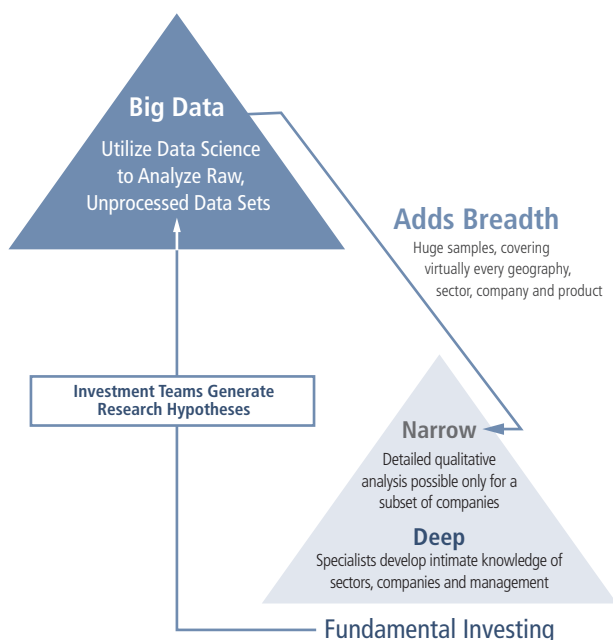
Just as important, investment management professionals from both camps are the most vital source of the research hypotheses that define the scope of data scientists’ work and bring shape and coherence to otherwise raw alternative data sets. Data science can provide lots of answers, but it would not even know the questions without the experience and expertise of investment professionals.

In short, we think data science is only truly additive to the investment management process when information flows freely within an integrated research platform. But we also believe that, when different research centers share a data science resource, those information flows can be greatly enhanced.

Information Flows: Big Data ↔ Fundamental Investing

As we described, data science can bring new breadth to fundamental analysts’ research. The starting point for this work is often a research hypothesis originating among the analysts themselves.

Most hypotheses are concerned with how well certain products are competing, or the immediate results achieved by new strategic initiatives. Some have their origins in research into environmental, social and governance (ESG) factors: relevant and material data on these factors are still often missing from corporate reporting, and third-party ESG ratings typically rely on corporate reporting and regulatory filings that can be immaterial, incomplete, non-standardized and backward-looking; alternative data that we identify and analyze ourselves can potentially be powerfully additive.



Overall, we find that many of our data science research hypotheses fall into one of the following six analytical categories:

Competitive Threat	Are Amazon's new batteries eating into Energizer's market? READ MORE
--------------------	--

Customer Mix	Are men really going to buy clothes from Lululemon? READ MORE
--------------	---

Customer Loyalty	Are rewarded customers good customers? READ MORE
------------------	--

New Business Initiative	How do customers like the new-look McDonalds? READ MORE
-------------------------	---

Pre-IPO	Why did we doubt that investors would pick up Lyft and Uber? READ MORE
---------	--

ESG	Is Comcast a good place to work? READ MORE
-----	--

In simple terms, the data science insights that feed into our fundamental research aim to enhance our understanding of economy-wide, sector-wide and company fundamentals so that we are better able to pick out who is going to win in the marketplace. We want to pick the next winners in a particular sector before anyone else is aware they exist; and we want to identify threats to incumbent businesses before they make their impact. In our new world of diminishing sell-side research and the intensifying focus on efficiency in investors' search for alpha, data science has a clear role.

These insights are additive to our research and stock-selection efforts, rather than decisive in themselves. Different strategists working to different time horizons will make different buy and sell decisions—but whatever they do, they will now be informed, to a greater or lesser extent, by the additional depth and breadth that these data science insights bring.

“Data science fills the gap between fundamental investors and quants. To make it a success, it cannot simply be a toe in the water for the traditional research team on the one hand or the quants on the other—this is about capturing that middle ground of greater breadth and greater depth for both of them.”

– Michael Recce

Chief Data Scientist, Neuberger Berman

“In quant investing you make a hypothesis about the metrics you want to forecast as a key component of the performance of a company, and try to find evidence that those metrics have some predictive value. Data is what you need to test the metrics.”

– Ray Carroll

Chief Investment Officer, Neuberger Berman Breton Hill

Information Flows: Big Data ↔ Quantitative Investing

Just as there is a direct flow of information between our data science team and our team of fundamental analysts, so there are direct flows between data science and our quantitative team.

In some cases the links are so immediate that they effectively create data science strategies, such as high frequency trading. Alternative data can also be scoured for profit opportunities around the quarterly reporting season, potentially by anticipating earnings surprises.

At Neuberger Berman, quantitative investing operates with a longer time horizon. We are most focused on factor- or risk premia-based quantitative investing, which tries to identify risks that, over long time periods, have systematically been rewarded in excess of the market return.

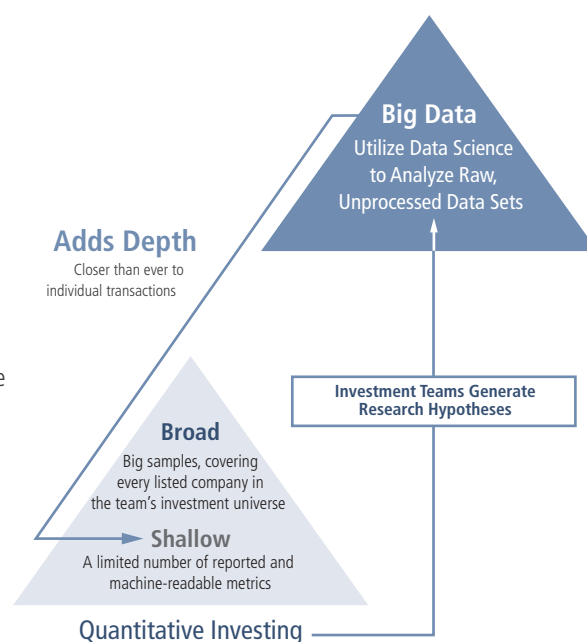
We think the most thoroughly tested risk premia are the value, carry, quality, momentum and insurance risk premia. Once we have identified these risks, we then try to identify stocks that map closely onto them. That involves gathering data about those stocks. These are primarily traditional fundamental metrics and ratios, but alternative data have at least three roles to play:

First, they can provide entirely new inputs directly into the mapping process. A good example is how our team has enriched its momentum signal by looking beyond stock price momentum, and even beyond earnings momentum, to include signs of momentum in credit card transactions and in keywords from management conference calls.

Second, alternative data can fill gaps that companies leave in their standard reporting. That can be relatively simple things, such as getting around Apple’s recent decision to report aggregate volumes rather than the number of phone handsets it has sold—information that may still be useful to us. It can also cover more complex issues. At a general level, ESG factors are often still missing or non-standardized in current corporate reporting, for example. More specifically, our quantitative team finds the research and development costs line item useful for analysis in the technology and healthcare sectors that tend to report it; alternative data, particularly data such as job postings and other items linked to operational expenditure, can act as a proxy for R&D in other sectors.

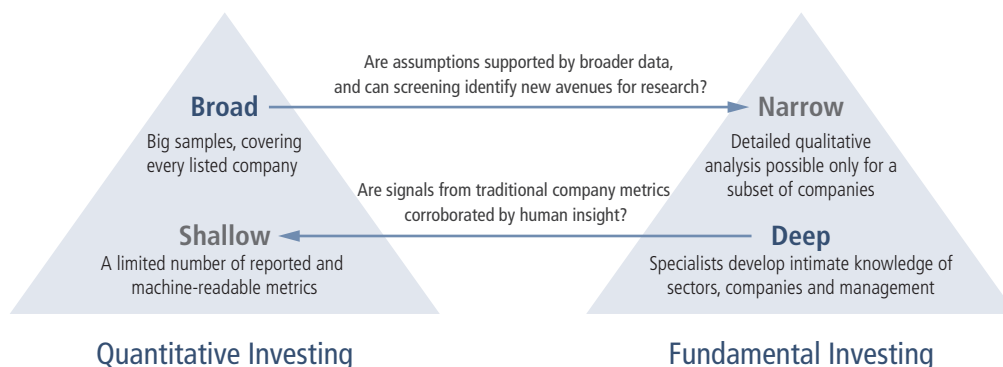
Third, and most commonly, alternative data simply enrich our quantitative team’s dataset for and knowledge of a company or sector. It can potentially identify risks that are not discounted before those risks exert their impact; and it can identify idiosyncratic drivers of return in excess of that predicted by the presence of systematic risk.

These insights can come to the quantitative team direct from the data science team, and often they come in response to specific research hypotheses developed by our quantitative team. However, we find that the additional information from alternative data is often most powerful when it comes mediated by the human insights of our team of fundamental analysts.



Information Flows:

Quantitative Investing ↔ Fundamental Investing



Information flows both ways between our fundamental and quantitative research teams. There has been a quantitative risk-management overlay to fundamental investing for half a century—the world of benchmarking, tracking error, information ratios and portfolio factor sensitivities. When these two functions are integrated more fully, however, quantitative screens can suggest entirely new avenues for bottom-up research, or risks that may be hidden to the analyst’s eyes. The quantitative team may see something at the (broad) macro level that is not being picked up by a fundamental analyst’s (deep) company- and sector-level view, for example.

Fundamental analysts’ insights can also inform quantitative investing. They can often confirm whether or not signals from fundamental metrics and ratios are corroborated by deeper human insight into individual companies.

For example, the quantitative investing team may think that a stock maps well onto the metrics that define a value or quality company, but their fundamental analyst colleagues may have reason to differ. That is why we have a process that automatically reports meaningful differences between the views of Neuberger’s fundamental analysts and the quantitative metrics we have for the stocks we are researching. The most obvious example is when a stock that trades with a low price-to-book or price-to-earnings ratio does so because the market has already discerned that its future earnings prospects are poor—a “value trap.” Similarly, low free cash flow often identifies a low-quality company, but it may simply reflect a high level of capital expenditure. This often does act as a drag on performance, but some management teams are great allocators of capital. Fundamental analysts may have an insight into these outliers.

Quantitative investing seeks to generate returns from known risks. Fundamental analysts, who enjoy deeper insights into individual companies’ prospects, can help in seeking to eliminate some of the most extreme downside outcomes from that risk distribution (or “cut off the left tail”).

Alternative data reflecting real-time sales, competitive threats or the performance of new investments can help make these human insights more comprehensive and immediate. As such, alternative data can further enrich the insights that fundamental analysts bring to quantitative investing.

“The alternative data footprint of several companies in our investible universe has grown exponentially over the last few years. Understanding and making sense of such data provides important insights in relation to the employee organizations and customer behavior underlying the operations of these companies.”

– Hari Ramanan

Portfolio Manager, Equity Research

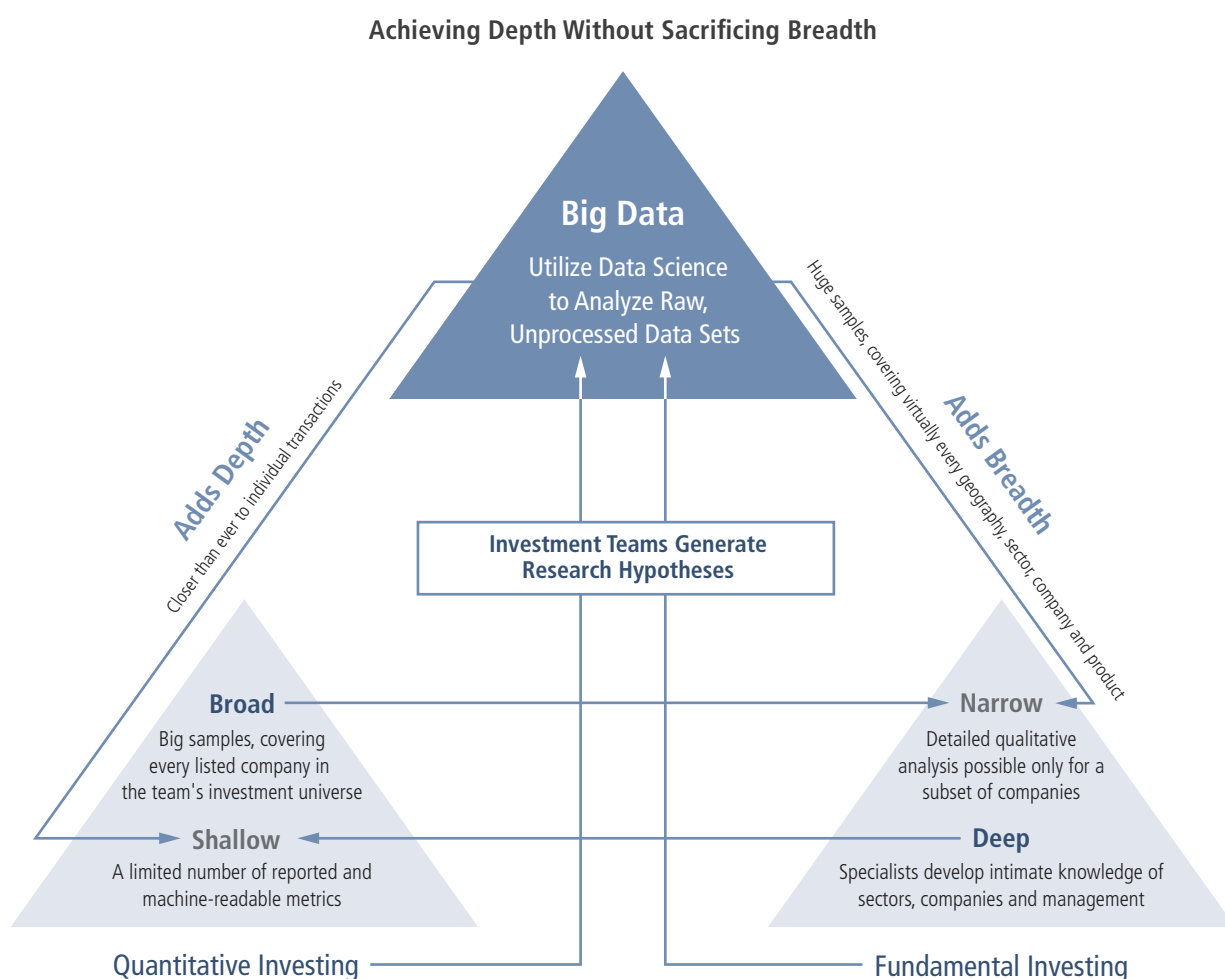
Integration Helps Information Flow

To get the most out of data science, practitioners should recognize that it is not a replacement for traditional investment research, but a complement to it; and not a technology support function for investment professionals, but an extension of what they already do. For the same reasons, simply hiring a team of data scientists and setting them to work is not necessarily going to enhance an investment manager's search for alpha. Finding a common language with which to integrate that team into the existing information flows that investment teams generate is critical.

We think this has both strategy and governance implications.

In strategy terms, we think the application of data science fits an already advanced trend in our industry. Many investors are abandoning the low-tracking error middle ground in favor of highly active, higher-conviction portfolios on the one hand and more cost-efficient and risk-efficient quantitatively managed portfolios on the other. Alternative data has a role to play in both of these approaches, and in enhancing the dialogue between them: boutiques that are active in only one or the other may not glean enough benefit from data science to justify the necessary investments.

The governance point is related to this. A centralized research capability—as opposed to siloed teams of analysts supporting individual strategies, groups of strategies, or portfolio management teams—is arguably best for facilitating the flow of information that data science stimulates. We believe centralized research can get the most from the growing wealth of alternative data in the world.



Data Science and Investment Management: Beyond the Hype

We are probably not the only asset manager telling you about the revolutionary potential of big data. We may be in a minority, however, when we say that we believe a lot of this commentary will be hype.

Data that are as varied, immediate, detailed and comprehensive as those becoming available to us today will certainly extend the reach and deepen the insights of investment research. They are already opening up entirely new avenues for research and analysis. Big data has a clear role in the world of diminishing sell-side research and the intensifying focus on efficiency in investors' search for alpha.

We believe that is unlikely to be the story for most practitioners, however. Poorly selected swatches of data, acquired passively and uncritically, compromised by gaps and anomalies and hosted on inappropriate platforms, will only take investors so far—and probably not far enough to justify the costs. Disillusionment will follow and the hype will burn itself out.

The ones left to benefit will be those that not only make the investments in data science expertise and technology infrastructure that are necessary to glean strategic insights from the datasets, but also recognize its structural and governance implications.

Data science helps us see the world both more comprehensively and in more detail. That makes it additive to both fundamental and quantitative research efforts in investment management. We believe that recognizing this dual role for data science is important; and that a firm hosting both quantitative and fundamental capabilities, and nurturing dialogue between them, is more likely to realize its full potential.

Most importantly, data science is a new tool in traditional investment research efforts, not a replacement for them. Data science has to ask the right questions. That requires investment professionals to generate research hypotheses that are relevant to investment objectives and within the scope of data science and its datasets. We believe that achieving a common language with which to nurture that information flow is much easier when a firm maintains a centralized research program that feeds the full range of its fundamental and quantitative strategies.

Big data is out there, waiting to change our expectations of what investment research can achieve. To harness it, active managers need to embrace data science, but also structure their research efforts to exploit its full potential.

Data Science at Work

Use Cases at Neuberger Berman

References to any securities in this documents are for illustrative purposes only and do not constitute a recommendation to investors. Case studies shown were selected based on the statistical significance of the data trend identified from the analysis.

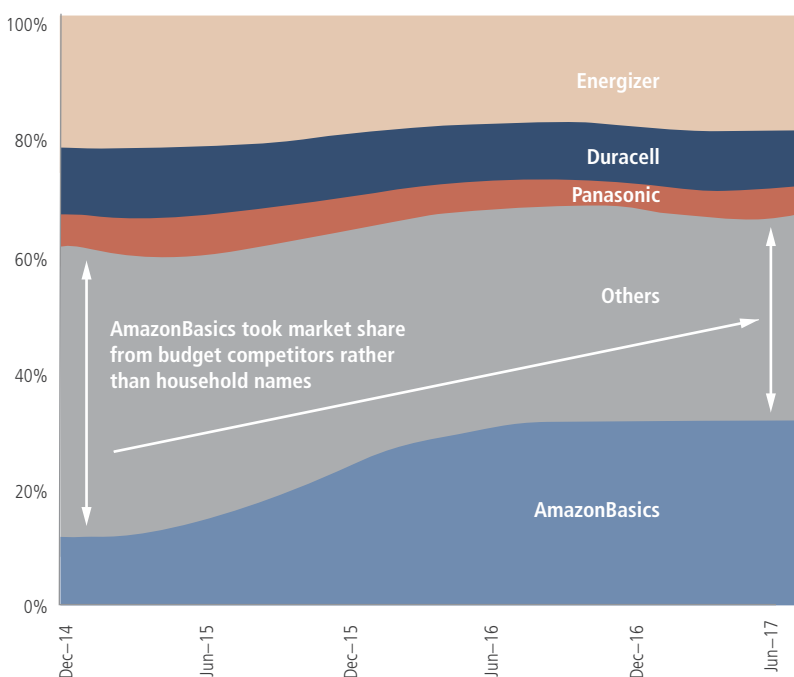


Are Amazon's new batteries eating into Energizer's market?

We have become used to game-changing disruption whenever Amazon enters a new market. So when it began to report growing market share for its own brand, AmazonBasics batteries, it was natural to anticipate havoc for household names such as Energizer, Duracell and Panasonic.

Analysis of transaction data on Amazon's own platform—35,000 daily unique users over the course of five years—showed that AmazonBasics was taking share from smaller brands, however. While Energizer's market share did fluctuate a little over the period we looked at, there was very little churn with AmazonBasics.

Battery Market Share of Transactions



Source: Amazon, Jumpshot.

This example shows how alternative data can potentially reveal the reality of competitive threats long before it becomes evident in standard corporate reporting.

KEY DATA SOURCE

Online transaction data

FEATURES

- SKU-level data on Amazon purchases since 2014
- 1.3m worldwide users covering approx. 2.7m unique brands and 28m unique products

WHAT WERE WE LOOKING FOR?

Evidence that incumbent brands could withstand Amazon's entry into the batteries market.

WHAT DID WE LEARN?

Amazon's brand eats into smaller brands' market share rather than the market leaders'.

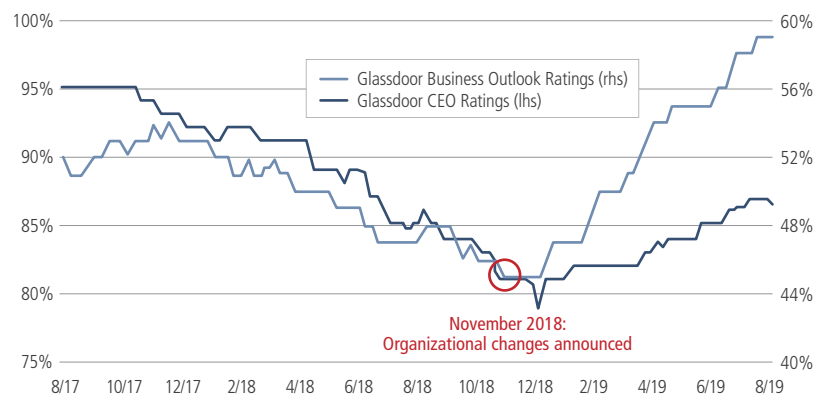
Are employees buying into the big changes at Procter & Gamble?

In 2016, the new CEO of Procter & Gamble started to implement organizational restructuring in an effort to make the business simpler and more responsive. By the middle of 2017, however, the slow pace and lack of clarity in the restructuring led to an activist investor taking a large stake and initiating a proxy contest to gain a seat on the company's board. That bid ultimately failed, but, interestingly, the activist was offered a seat on the board in any case. In November 2018, P&G announced its largest organizational change in 20 years, implementing many of the measures for which the activist had advocated.

By the first quarter of 2019, P&G was reporting its best quarterly organic sales in over a decade. Investors had already decided to get behind the stock—but how confident could they be that this was the sustainable result of a genuine turnaround within the business?

One way to get additional insight is to ask the people working there. Gathering the ratings that employees leave on the recruitment website Glassdoor is one way to do that. While the sample size was limited, the results were resoundingly clear.

Glassdoor Business Outlook and CEO Ratings – P&G

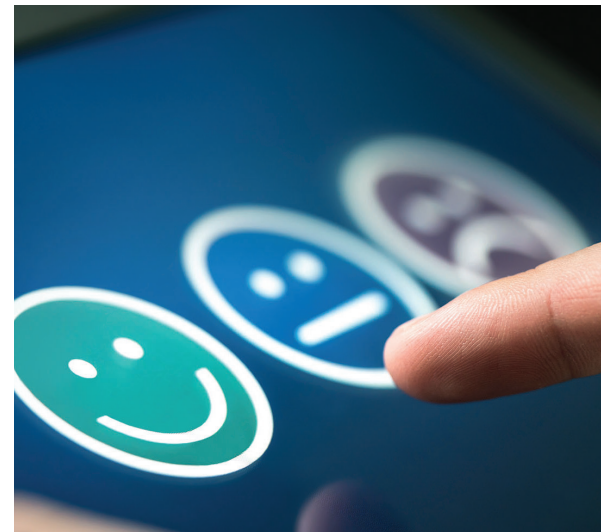


Source: Glassdoor.

Asked if the company's business outlook will improve in the next six months, the proportion of returns with a "positive" response began to climb steadily, coinciding precisely with the announcement of organizational changes. Moreover, despite the proxy battle, when asked if they approved of the CEO, after falling from 95% to 81% between mid-2017 and the end of 2018, the proportion of employees who approve has climbed back to 87%. The average CEO approval rating for consumer staples companies is 77%.

Alongside traditional reported data, these Glassdoor survey results seem to indicate that P&G's employees are buying into the major changes underway at the company, and enjoying the new freedom and responsibilities they have in their product categories.

This example is a reminder that employees generate useful data as well as consumers. This data can offer an insider's perspective on the ethos and performance of a company, especially at times of change, long before these things show up in traditional corporate reporting.



KEY DATA SOURCE

Internet content and search data

FEATURES

- 32 web content datasets since 2015 including:
 - Social media engagement on company home pages
 - Store location reference data
 - Car inventory via leading car sales websites
 - App ratings (Google Play Store and Apple Store)
 - Company discounts
- 100m worldwide users' Google Search terms from 2014
- ~50m unique keywords on a daily basis
- Click timestamp, type (organic or paid), geographic information

WHAT WERE WE LOOKING FOR?

The inside view on Procter & Gamble's organizational changes from employee ratings on Glassdoor.

WHAT DID WE LEARN?

A turnaround in employee sentiment supported our positive view on the restructuring.



KEY DATA SOURCE

Online job post data

FEATURES

- 4m worldwide (70% of U.S. total job postings)
- Daily collection from 32,000+ companies' websites (8,000+ public companies)
- Data since 2007, detailed job descriptions since 2014
- Job level, category, region

WHAT WERE WE LOOKING FOR?

A clearer view of expansionary operational expenditure by services companies, and of who is winning the race to provide cloud computing solutions.

WHAT DID WE LEARN?

Hiring patterns indicated to us that Microsoft Azure can sustain its recent success, especially in the financial services sector.

How can I tell if companies are investing in the post-capex era?

When developed economies were powered by traditional manufacturing, we could get a good sense of business confidence from the capital expenditure that companies reported. Good investments would start to show up in revenues six or 12 months later, and in aggregate, that reporting also gave us an insight into where we were in the business cycle.

Those signals are much weaker in our new economy of services and technology. That is a world of operational expenditure—wages, salaries, rents and the like—rather than capital expenditure. Hiring engineers and designers is the modern world's capex, but that information is not reported in quarterly corporate filings.

It is available in the form of job postings, however. We can now collect these from more than 8,000 U.S. public companies, a sample that represents almost three-quarters of all the job advertisements in the U.S.

That information can enable us to see whether or not a company is hiring lots of engineers and designers in its early life and then filling positions in sales and marketing as it matures, giving us an objective insight into management's confidence in its business model.

Job postings can also tell us about the performance of certain suppliers to the companies that are hiring. For example, when Microsoft's Azure started to surprise with its success in the cloud computing market, our technology analysts looked for real-time confirmation of the sustainability of this trend in job postings for I.T. positions, which often specify whether the hiring company is looking for someone experienced at working with Oracle, Google, AWS or Azure. That can be an indicator of future cloud-computing investment at the hiring company. Based on what we found, we believed that Azure was likely to sustain its success—but also that it was the preferred solution in the financial sector, versus AWS in the technology sector.

These examples are a good reminder that alternative data do not always come in the form of numbers, and that gathering it and understanding it can suggest new metrics for analyzing individual companies and the wider economy—metrics that often fill gaps in public reporting for the first time.

Why do I not want to hear the words 'very optimistic' on an earnings call?

At Neuberger Berman we have developed a natural language processing model that can identify more than 5,000 phrases of two or more words, which, when they appear in the transcripts of quarterly earnings calls, it considers to carry positive or negative implications.

It will come as no surprise that "raise guidance" is seen as positive and "reduce guidance" is seen as negative. Similarly, "ahead of schedule" is better than "slower than." We would rather see "repurchase shares" than "equity offering."

However, is "very optimistic" a good or a bad sign? Correlation between this "bigram" and subsequent stock performance, picked up from thousands upon thousands of transcripts, indicates that it is a bad sign, on average. Management tends to say it when results have been poor and they want to persuade investors that a turnaround is on the way. Bigrams such as "closer look," "briefly review" or "quick update" may seem completely neutral, but analysis shows that the first tends to be positive and the other two negative—management likes to get into details that reflect well on the business and skirt over the problems.

Extract from the AT&T Inc., Q2 2019 Earnings Call, July 24, 2019

half of the year. Let's now look at our **segment operating** results, starting with our Communications segment on Slide 8. The story of our Communications segment this quarter is stable **revenues, growing EBITDA and expanding margins**. Mobility continues to **build momentum** and deliver solid results across the board with revenue, EBITDA and margin growth while adding phone subscribers. Our Entertainment Group is delivering **EBITDA growth** and Business Wireline revenue trends improved in the quarter, thanks to strength in strategic and managed services and about \$125 million from IP licensing. But even without those licensing proceeds, Business Wireline revenue trends were the best that we've seen in years. And when you factor in strong business wireless performance, our Business Solutions revenue grew 2.3%. On the cost side, the team is doing great work in controlling content, promotions and **other operating** costs. Solid **cost management** was evident throughout the business, especially in our Entertainment and Business Wireline units. Let me give you some more details, starting with Mobility on Slide 9. Our Mobility business continues to **perform very** well. Service revenues grew by 2.4%. **EBITDA growth** was even higher at 3.1% and **EBITDA margins margins** expanded by 80 basis points with service margins of 56.1%. We had a **strong quarter** with 355,000 phone net adds, including 72,000 postpaid and 283,000 prepaid. And we added 388,000 smartphones in the quarter, further strengthening our customer base. Postpaid phone churn was up slightly to 0.86%, but was down sequentially. And at the same time, our prepaid business, especially Cricket, continues to perform at strong, consistent levels. Prepaid revenues were up nearly 10%. We had our 18th consecutive quarter of phone growth, and churn hit an all-time low at both Cricket and AT&T prepaid. With the network leadership and FirstNet expansion that Randall talked about earlier, we're confident that our wireless business **will get** even stronger as we evolve to 5G. In short, our network investments, particularly our spectrum deployment, are paying off, and **we're not** done yet. Now let's go on to our Entertainment Group results on Slide 10. Our focus on long-term customer value continues to impact our Entertainment Group. **EBITDA grew** both year-over-year and sequentially. This was the second **straight quarter of EBITDA growth**. Year-to-date, EBITDA is up about 4%. Expense reduction outpaced revenue declines, setting the stage for **EBITDA growth**. Broadband **revenue growth** helped us as well as **continued growth** in video ARPU's. The number of premium TV customers on the 2-year

Source: S&P Capital IQ. For illustrative purposes only.

The earnings call extract above shows our model in action, picking out positive bigrams in green and negative bigrams in red.

Again, there are intuitive findings: "expanding margins" and "EBITDA growth" tend to be positive. And there are apparently neutral phrases that contain hidden information: "will get" and "other operating" are interpreted negatively because they are associated with the non-specific language management uses to gloss over challenges, whereas "segment operating" is associated with laying out details of how different parts of the business have been performing; a bigram such as "straight quarter" is positive because few CEOs like to boast about their "third straight quarter of declining margins."

One can immediately see the preponderance of green over red bigrams in this extract. Our model classified this transcript in the top 5% of more than 800 that it processed in July.

These examples are a good reminder that alternative data are not always numbers, and that the techniques applied to interpret the data are as important as the data themselves.



KEY DATA SOURCE

Conference call transcripts and filings

FEATURES

- Earnings and acquisition announcement calls
- Monthly update calls (e.g. sales)
- Company conference presentations
- 1,500+ U.S. public companies from 2010 to present
- All corporate SEC filings

WHAT WERE WE LOOKING FOR?

Indications of management sentiment from earnings call transcripts.

WHAT DID WE LEARN?

The AT&T conference call for Q2 2019 revealed genuine optimism from management, which was justified by subsequent company performance.



Are men really going to buy clothes from Lululemon?

When Neuberger Berman's fundamental consumer sector analyst began to see positive signs coming out of the apparel company Lululemon Athletica, he turned to the data science team to find out whether more recent sales confirmed the apparent trend for growth. We scoured six billion credit card transactions and also worked with a data vendor in Asia to scrape SKU-level data on volumes and prices from the Chinese ecommerce site Tmall.

The findings were promising, and these growing ecommerce sales were reflected in its quarterly results two months later.

But our analysis was able to go still further and look into factors underpinning the sustainability of this sales growth.

Our researcher wanted to examine the company's claim that it was quietly growing its menswear sales, having built its reputation in womenswear. This sort of question is often posed by data scientists, who apply artificial intelligence and machine learning to uncover patterns in large datasets. Data scientists have learned to identify bank accounts that are likely being used for organized crime, for example: they don't write checks, they do a lot of cash-only transacting, they get mysterious deposits from charities, and so on. We can also identify with some confidence whether a bank account or credit card is owned by a baby-boomer or a Millennial, or by a man or a woman, based on their shopping habits. And it turned out that credit card transactions revealed Lululemon not only increasing sales but also selling to a growing number of men.

This example shows how alternative data and machine learning can combine to discover not only whether a company's sales, are growing, but also how its customers are changing over time.

KEY DATA SOURCE

Credit/debit card and bank account transaction data

FEATURES

- Transactions include date, amount and description
- Bank accounts include deposits, such as paychecks
- Detailed geographic information (to the individual store level)
- Detailed demographic information (age, gender, income)
- Online vs. in-store sales (in most cases)

WHAT WERE WE LOOKING FOR?

Early and objective evidence of sales growth at Lululemon.

WHAT DID WE LEARN?

Credit card transaction data predicted sales growth reported by the company two months later, but analysis also revealed a marked increase in sales to men.

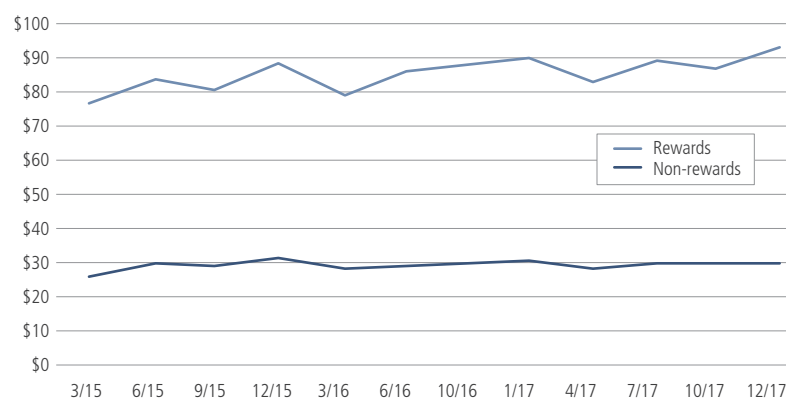
Are rewarded customers good customers?

One of our large-cap portfolio management teams wanted to know what sort of benefit Starbucks itself was getting from its customer loyalty program.

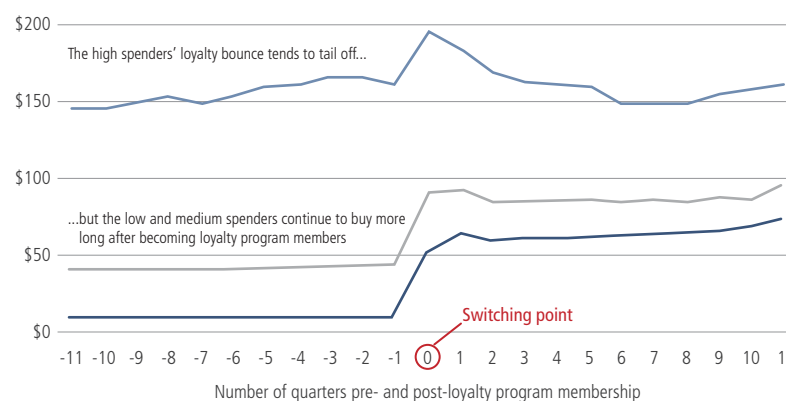
Our data science team turned to credit and debit card transactions. After identifying loyalty-program members, it found that they spent around \$88 on average each quarter versus the non-member spend of around \$30. We were also able to show that average revenue per user jumped substantially at the point of conversion to loyalty-program membership and, for all but the very highest-spending customers, it remained high and even rose further over subsequent years.

Furthermore, the revenue growth associated with our sample of reward members closely tracked revenue growth subsequently reported by Starbucks. In our view, we had apparently found a robust new alternative data metric with which to forecast future quarterly performance.

Quarterly Average Revenue Per User



Quarterly Average Revenue Per User



Source: Second Measure.

This example shows how alternative data can be used to tease out the spending patterns of distinct subsets of a customer base, and track the performance of business initiatives in close to real time.



KEY DATA SOURCE

Credit/debit card and bank account transaction data

FEATURES

- Transactions include date, amount and description
- Bank accounts include deposits, such as paychecks
- Detailed geographic information (to the individual store level)
- Detailed demographic information (age, gender, income)
- Online vs. in-store sales (in most cases)

WHAT WERE WE LOOKING FOR?

The impact that being on a retail loyalty program has on spending.

WHAT DID WE LEARN?

Customers who join the Starbucks loyalty program tend to spend more—immediately and persistently.



KEY DATA SOURCE

Credit/debit card and bank account transaction data

FEATURES

- Transactions include date, amount and description
- Bank accounts include deposits, such as paychecks
- Detailed geographic information (to the individual store level)
- Detailed demographic information (age, gender, income)
- Online vs. in-store sales (in most cases)

WHAT WERE WE LOOKING FOR?

The customer's take on McDonald's pilot "Experience of the Future" outlets.

WHAT DID WE LEARN?

Traffic increased at the new outlets and repeat visits went up, giving us added confidence regarding the full rollout.

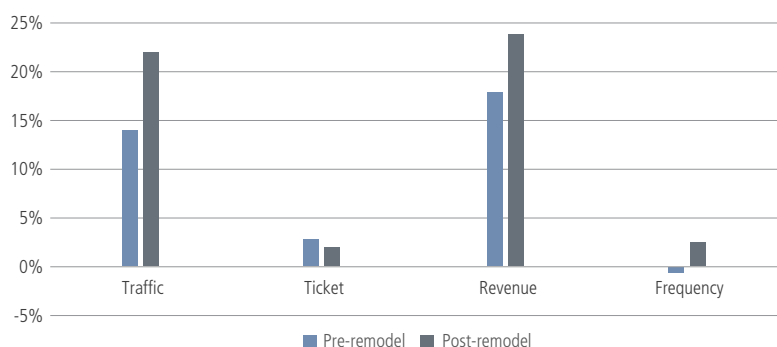
How do customers like the new-look McDonald's?

When McDonald's announced that it would spend \$6 billion on remodeling its restaurants, as investors we naturally wanted to know whether this substantial expenditure would add to the top line or tarnish the brand in the eyes of customers.

Our data science team was able to isolate the credit and debit card transactions made in the 839 "Experience of the Future" restaurants that were remodeled in the second half of 2017.

We found slightly slower growth in average spend in the new-look stores, with their new mix of self-service functionality and table service. They more than made up for it, however, with substantially increased growth in traffic. Whereas the number of repeat visits had been declining in the old-style restaurants, we saw it growing for the new model. Given the company's plans to roll the new model out for most of its U.S. restaurants by 2020, these findings gave our analyst more confidence in its potential to report earnings upside over the next two years, against a tough background for the sector.

Average Monthly Growth of McDonald's "Experience of the Future" Restaurants



Source: Second Measure.

This example shows the unique insights that store-level data can offer when a business introduces potentially high-impact new initiatives, insights unavailable through traditional data and research.

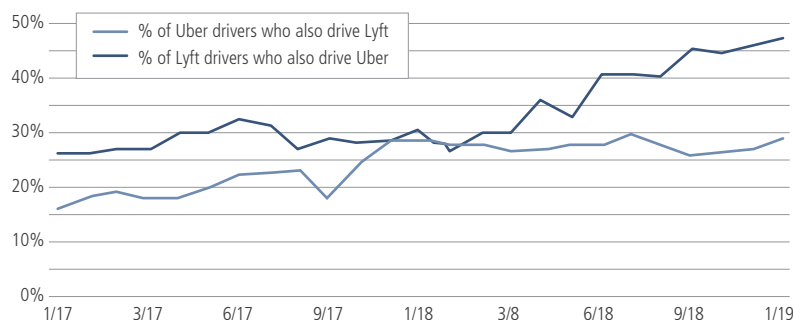
Why did we doubt that investors would pick up Lyft and Uber?

The market leaders in ride-hailing apps both took their stock public last year, and both similarly suffered inauspicious first-day trading. Our analysts were not surprised by this because our data science team, using two types of alternative data, had confirmed their existing doubts about potential rising costs and potential declining sales growth.

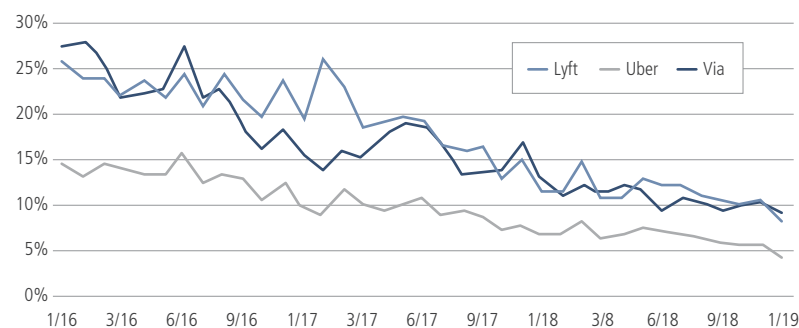
On the costs side, from direct deposit information, we were able to identify Lyft and Uber drivers from our sample of bank account data. Through 2017 and 2018 that data showed a rising percentage of drivers working for both companies. That implied declining driver loyalty and increasing driver churn even as the absolute number of drivers was rising.

On the sales side, from credit and debit card transactions, we found clear evidence of rapidly declining new customer growth for all three of the leading ride-hailing firms, which suggests that future revenue growth will be ever more reliant on increasing use by existing customers.

Driver Performance: Lyft vs. Uber



New Customer Rate: Rideshare Rivals



Source: Second Measure.

These findings were not decisive for all portfolio managers: some decided not to invest; others saw the findings as a counterpoint to a strongly held investment thesis. All would agree, however, that these cases show how alternative data can provide a unique window into the true performance of pre-IPO companies. This case also shows how there is more to bank account data than consumer spending insights (payments out of accounts). Payments into accounts (e.g., paychecks) are key to explore labor mix and labor costs for individual companies, sectors and the economy as a whole.



KEY DATA SOURCE

Credit/debit card and bank account transaction data

FEATURES

- Transactions include date, amount and description
- Bank accounts include deposits, such as paychecks
- Detailed geographic information (to the individual store level)
- Detailed demographic information (age, gender, income)
- Online vs. in-store sales (in most cases)

WHAT WERE WE LOOKING FOR?

Patterns in driver loyalty and new customer growth.

WHAT DID WE LEARN?

An increasing number of drivers were working for more than one company, and fewer new customers were subscribing.



KEY DATA SOURCE

Internet content and search data
Online job post data

FEATURES

- 4m worldwide (70% of U.S. total job postings)
- Daily collection from 32,000+ companies' websites (8,000+ public companies)
- Data since 2007, detailed job descriptions since 2014
- Job level, category, region

WHAT WERE WE LOOKING FOR?

An insider's view on the working environment at Comcast, against a background of contrasting third-party evidence.

WHAT DID WE LEARN?

Employees rate the company higher than the average for its sector and the firm appears to be hiring for genuine growth rather than because of employee churn.

Is Comcast a good place to work?

Alternative data have a key role to play in analyzing environmental, social and governance (ESG) factors because many of these factors are not covered in traditional company reporting, and a lack of standardization makes comparison difficult even when data is reported. The problem is most acute—and the data science potential most promising—when it comes to “softer” social factors such as human capital management.

For the third straight year, in 2019 U.S. telecommunications giant Comcast was named to LinkedIn's Top Companies list, ranking 15th. In the same year, *Fortune* ranked it third in its list of “Best Big Companies to Work For” as well as listing it among its “100 Best Companies to Work For.” In 2018 the firm was named a Leading Disability Employer by the National Organization on Disability.

Just as Comcast was receiving these accolades, however, MSCI was pointing to its planned acquisition of Sky and its “multiple labor controversies” and giving it a Labor Management ranking of 2.8—much lower than the sector average of 6.2.

Whom to believe? It is tempting to trust the professional rating agencies over the eye-catching awards. But we know that there are often severe discrepancies between different agencies' ratings of the same companies. Alternative data can give us a new, proprietary perspective on these questions. One source of data are the ratings that employees leave on the recruitment website Glassdoor. When we scraped those for the telecommunications sector, we found that Comcast ranked well above average.

Perhaps a more objective, less-biased source of data are active job postings. We can now collect a sample that represents almost three-quarters of all the job advertisements in the U.S. We find it revealing to compare the proportion of a company's workforce that is represented by currently live job postings with subsequent growth in Selling, General & Administrative expenses. We believe that gives us an insight into how many of those job postings relate to genuine expansion of employment at the firm and how many are due to churning of the same role—a high rate of churn would imply that employees don't believe the firm is a good place to work.

On that metric, Comcast's ratio compared well with its sector peers, putting it seventh out of 22 companies.

This example shows how alternative data can offer new, proprietary perspectives on often contentious ESG factors, where divergent views are commonplace. Together with data science techniques, they can contribute to a more holistic view of companies' exposures to ESG risks and their progress in addressing those risks.



About Neuberger Berman

Founded in 1939, Neuberger Berman is a private, 100% independent, employee-owned investment manager. From offices in 35 cities worldwide, the firm manages a range of strategies—including equity, fixed income, quantitative and multi-asset class, private equity, real estate and hedge funds—on behalf of institutions, advisors and individual investors globally. With more than 600 investment professionals and approximately 2,200 employees in total, Neuberger Berman has built a diverse team of individuals united in their commitment to delivering compelling investment results for our clients over the long term. That commitment includes active consideration of environmental, social and governance factors. Our culture has afforded us enviable retention rates among our senior investment staff and has helped earned us a citation from *Pensions & Investments* as a Best Place to Work in Money Management for six consecutive years. The firm managed \$356 billion in client assets as of December 31, 2019. For more information, please visit our website at www.nb.com.

All information is as of December 31, 2019 unless otherwise indicated. Firm data, including employee and assets under management figures, reflect collective data for the various affiliated investment advisers that are subsidiaries of Neuberger Berman Group LLC (the "firm"). Firm history and timelines includes the history and business expansions of all firm subsidiaries, including predecessor entities and acquisition entities. Investment professionals referenced include portfolio managers, research analysts/associates, traders, and product specialists and team-dedicated economists/strategists.

This material is provided for informational purposes only and nothing herein constitutes investment, legal, accounting or tax advice, or a recommendation to buy, sell or hold a security. This material is general in nature and is not directed to any category of investors and should not be regarded as individualized, a recommendation, investment advice or a suggestion to engage in or refrain from any investment-related course of action. Investment decisions and the appropriateness of this material should be made based on an investor's individual objectives and circumstances and in consultation with his or her advisors. Information is obtained from sources deemed reliable, but there is no representation or warranty as to its accuracy, completeness or reliability. All information is current as of the date of this material and is subject to change without notice. Any views or opinions expressed may not reflect those of the firm as a whole. Third-party economic or market estimates discussed herein may or may not be realized and no opinion or representation is being given regarding such estimates. Neuberger Berman products and services may not be available in all jurisdictions or to all client types. This material may include estimates, outlooks, projections and other "forward-looking statements." Due to a variety of factors, actual events or market behavior may differ significantly from any views expressed. Investing entails risks, including possible loss of principal. Past performance is no guarantee of future results.

Data science research case studies are for illustrative purposes only. Case studies shown were selected based on the statistical significance of the data trend identified from the analysis. This material is not intended as a formal research report and should not be relied upon as a basis for making an investment decision. The firm, its employees and advisory accounts may hold positions of any companies discussed. Specific securities identified and described do not represent all of the securities purchased, sold or recommended for advisory clients. It should not be assumed that any investments in securities, companies, sectors or markets identified and described were or will be profitable. Any discussion of data science analytical categories or environmental, social and governance (ESG) factor and ratings are for informational purposes only and should not be relied upon as a basis for making an investment decision. Data science analytical categories and ESG factors are one of many factors that may be considered when making investment decisions.

This material is being issued on a limited basis through various global subsidiaries and affiliates of Neuberger Berman Group LLC. Please visit www.nb.com/disclosure-global-communications-for-the-specific-entities-and-jurisdictional-limitations-and-restrictions.

The "Neuberger Berman" name and logo are registered service marks of Neuberger Berman Group LLC.

© 2020 Neuberger Berman Group LLC. All rights reserved



FIRM HEADQUARTERS

New York
800.223.6448

REGIONAL HEADQUARTERS

Hong Kong
+852 3664 8800

London
+44 20 3214 9000

Tokyo
+81 3 5218 1930

OFFICES

AMERICAS

Atlanta
Bermuda
Bogota
Boston
Buenos Aires
Chicago
Dallas
Los Angeles
New York
San Francisco
Sao Paulo
Tampa
Toronto
West Palm Beach
Wilmington

EUROPE, MIDDLE EAST & AFRICA

Dubai
Dublin
Frankfurt
London
Luxembourg
Madrid
Milan
Paris
Rome
The Hague
Zurich

PORTFOLIO MANAGEMENT CENTERS

Atlanta	Milan
Beijing	New York
Bermuda	Paris
Boston	San Francisco
Buenos Aires	Shanghai
Chicago	Singapore
Dallas	The Hague
Hong Kong	Tokyo
London	Toronto
Los Angeles	

NEUBERGER	BERMAN
-----------	--------

Neuberger Berman
1290 Avenue of the Americas
New York, NY 10104-0001

www.nb.com