Empirical Research Partners

Rochester Cahan 212 803-7973
Sungsoo Yang 212 803-7925

March 2, 2017

Stock Selection: Research and Results March 2017

Big Data, Little Alpha? The Robots Ate My Alpha: The Rise of Quant Funds

Bobby Axelrod Has It, Do You?

- Big Data is so hot right now it even made a cameo appearance in the latest episode of *Billions*. When Bobby Axelrod, never one to let a grudge fester, goes out looking for revenge against a rival hedge fund manager his mega-short trade hinges on satellite imagery of truck movements at an obscure factor in China. It's entertaining stuff but unfortunately in the real world the endorsement of Hollywood scriptwriters isn't necessarily enough to send us scurrying to the cloud to sift through terabytes of geospatial data.
- To try to dimension the legitimate opportunity in Big Data we conducted a comprehensive review of the academic research that's been done on the topic over the past four years. The 40 papers we read were all fascinating and covered a vast range of data source, from the use of facial recognition software during CEO interviews to text-mining posts on the SeekingAlpha website in order to extract the commentators' sentiment towards stocks. However, one thing stood out above all else: only 15% of the papers found predictive power over a year and most had alpha measured in days or at best a month or two.
- Appendix 1 on page 10 lists all the papers we reviewed along with a brief synopsis of their results and a note on the holding period the signal is applicable over.

Anything In It For Investors?

- Among the handful of papers that found alpha at *investment* holding periods, the most promising used Natural Language Processing (NLP) techniques to parse the textual content of 10-K filings and conference call transcripts. For example, one interesting paper studied bank stocks, and found that an increase in the number of negative words used in their 10-Ks predicts a significant increase in the probability of a distressed delisting in the next three years. Rising negativity also seems to precede missed dividend payments, higher year-ahead loan losses, and lower return on assets.
- We think there's enough potential in this avenue of research to add it to our research agenda for this year. We've long found that watching what companies do with their capital is much more useful that listening to what they say they'll do, but a preponderance of academic evidence suggests there might be merit in also capturing the nuance and context embedded in the text that surrounds the cold, hard numbers.

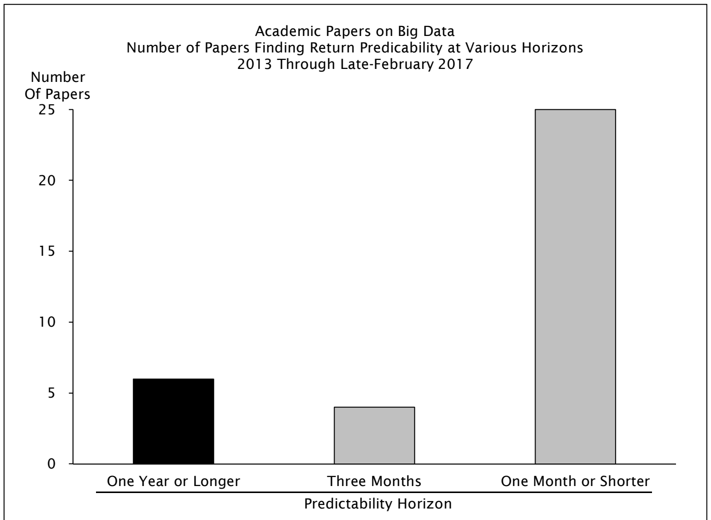
The Robots Ate My Alpha!

- There's a growing unease among fundamental investors that the success of a handful of quant hedge funds has altered the rules of the game. It's certainly true that quant investors have grown in stature: in the fourth quarter of last year they controlled 24% of all hedge fund U.S. equity assets, their highest-ever share. Furthermore, the five largest funds manage 70% of the quant pie, meaning the actions of a few behemoths matter.
- To get a handle on what the quants are actually up to we ran the portfolios of the four largest funds through our Portfolio Analytics framework. All four managers have carried a consistent momentum bias, which isn't a huge surprise. What's more interesting is the fact that two of them have had significant, decade-long bets towards fundamentally-stable stocks, a position that probably helped them for much of the post-Crisis era.
- It's also evident that the big quants have an ironclad discipline in avoiding stocks that have a high risk of failure, as measured by our Failure model. Exhibit 17 on page 9 presents stocks that are heavily owned by the Big Five quant hedge funds. None of the top 20 screen as Failure candidates and 15 of them score in the best two quintiles of our Core model. Verisign, Spirit AeroSystems, Hyatt Hotels, and U.S. Steel feature, among others.

Nicole Price (212) 803-7935 Yi Liu (212) 803-7942 Yu Bai (212) 803-7919 Iwona Scanzillo (212) 803-7915

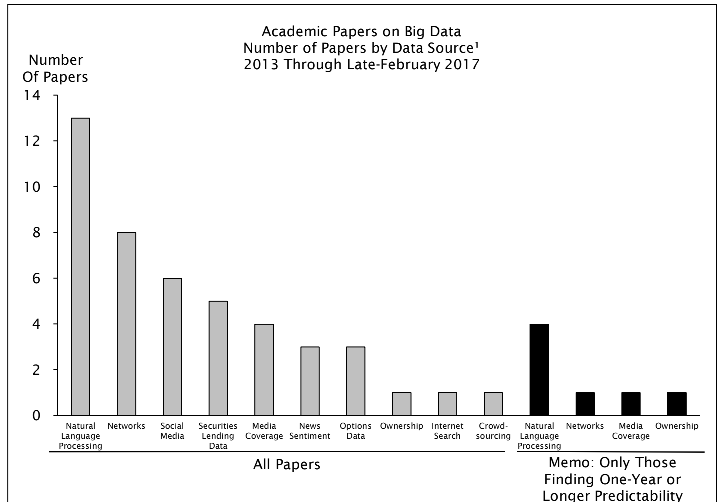
Conclusions in Brief

- Research on Big Data suggests it's mostly useful for traders…



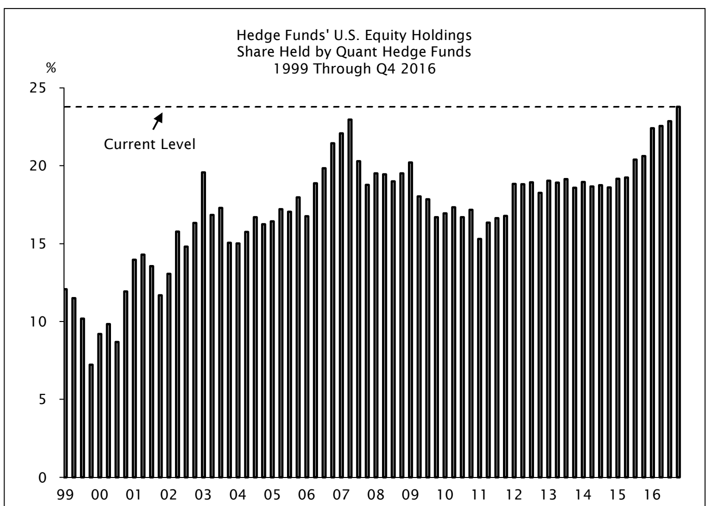
Source: Empirical Research Partners Analysis.

- …But text-mining shows some efficacy at investment horizons:



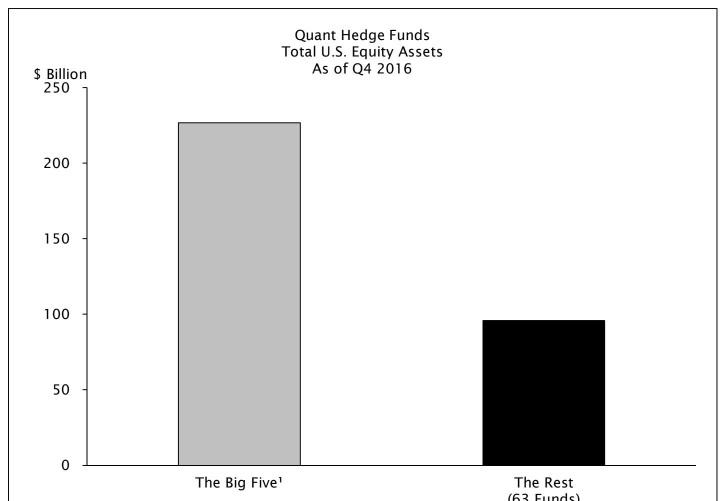
Source: Empirical Research Partners Analysis.
[1] Sums to more than total because papers can be classified in multiple categories.

- The quants' share of hedge fund equity assets has reached a new high…



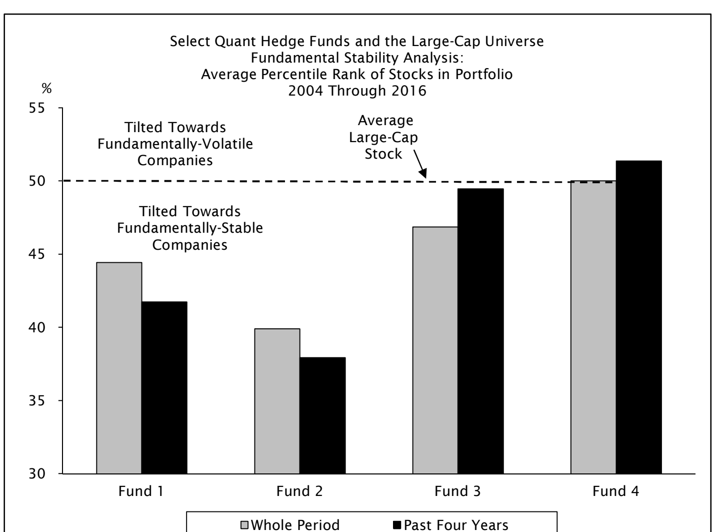
Source: FactSet Research Systems, Empirical Research Partners Analysis.

- …Driven by the Big Five quant funds:



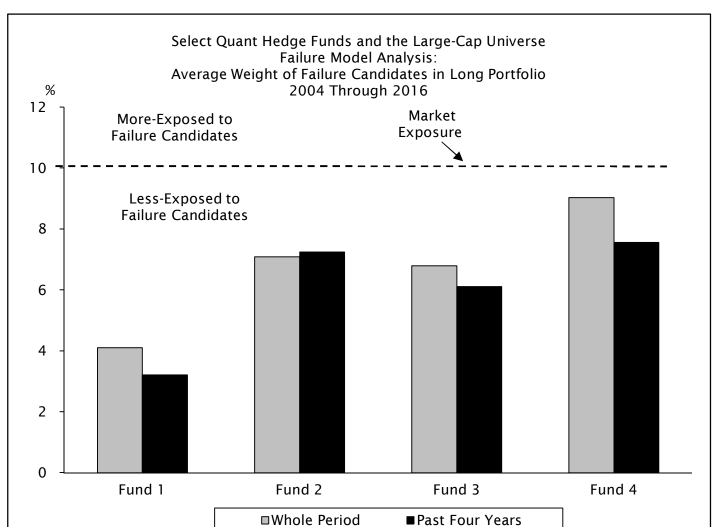
Source: 13F Filings, Empirical Research Partners Analysis.
[1] Five largest quant managers by U.S. equity assets.

- Two of the big quants have a significant stability bias...



Source: Empirical Research Partners Analysis.

- …And all of them are wary of Failure candidates:



Source: Empirical Research Partners Analysis.

Big Data, Little Alpha?

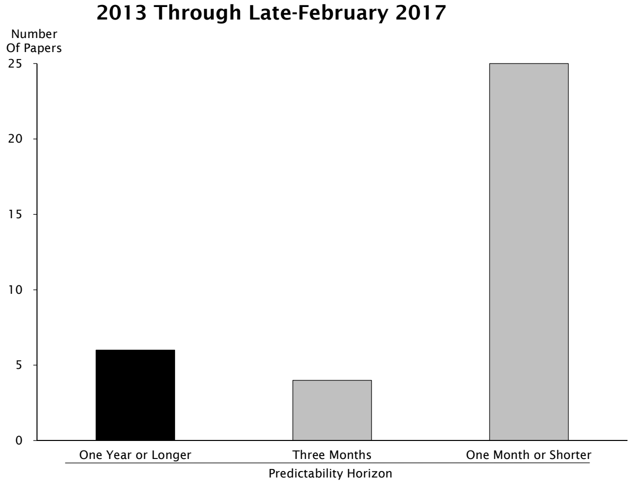
Bobby Axelrod Has It, Do You?

In the latest episode of *Billions*, a rollicking subplot sees Bobby Axelrod crafting a revenge trade against a rival hedge fund manager using satellite imagery of truck movements (or the lack thereof) at an obscure factory in China. But in a telling exchange with his analyst, Bobby is dismissive of the data itself: everyone has satellite data these days, give me more! Such is the conundrum of Big Data. If everyone has it the Edge is gone, but if you don't have it you're a Luddite clinging to your knitting needles while the robots pass you by, or so the data vendors and quants will tell you. In fact, the buzz around Big Data has reached the point where you need a cloud-based, NoSQL database just to keep track of it all.

We've been keeping an eye out for any research that suggests Big Data can help at *investment* time horizons, i.e., those measured in years rather than days or minutes or, increasingly, ticks. Since 2013 we've reviewed scores of academic papers, touting everything from facial recognition during CEO interviews to the singular importance of the word 'but' in conference call Q&A with analysts. All are fascinating but far fewer show any kind of predictive power beyond a week (see Exhibit 1). Appendix 1 on pages 10 through 12 presents the complete list of papers along with a brief description of their findings and the predictive horizon of their signals. Most are measured in days.

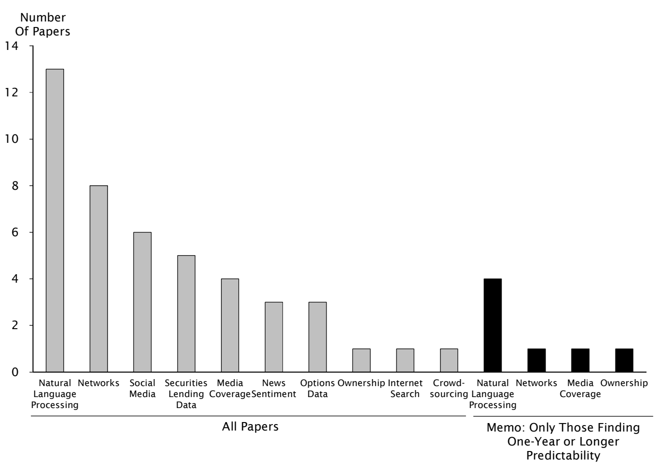
Exhibit 2 breaks down the papers based on the type of data used. Text-based content, which can be processed by increasingly sophisticated Natural Language Processing (NLP) algorithms, leads the pack. A lot of the recent academic work has focused on scraping the content of 10-Ks and conference calls, looking for nuggets of qualitative information that might be discounted in the stock price more slowly than hard, quantitative information. Networks are also a popular topic, the basic idea being that firms are much more interconnected than simple industry classifications can capture. Can the performance of a supplier or a company that does business in the same city tell us more about a firm than its somewhat contrived GICs peer group? Other popular topics include counting the number of 'likes' a firm gets on social media and delving into detailed securities lending data that sheds some light on the confidence of short sellers.

Exhibit 1: Academic Papers on Big Data
Number of Papers Finding Return Predictability at Various Horizons
2013 Through Late-February 2017



Source: Empirical Research Partners Analysis.

Exhibit 2: Academic Papers on Big Data
Number of Papers by Data Source[1]
2013 Through Late-February 2017



Source: Empirical Research Partners Analysis.

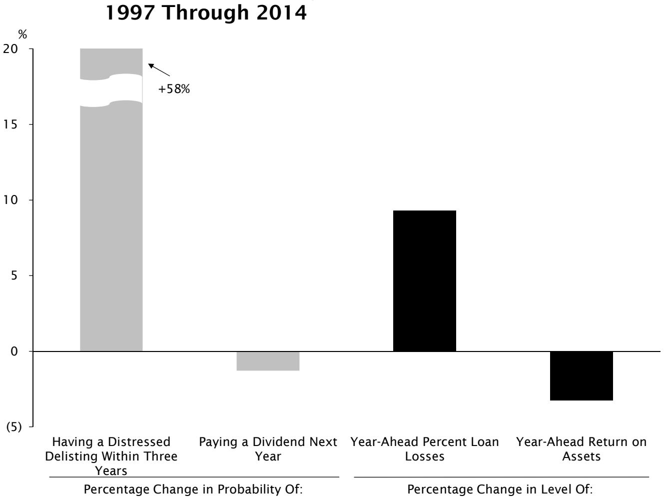
[1] Sums to more than total because papers can be classified in multiple categories.

Word Games

The black bars on the right of the second chart above show the breakdown for papers that found longer-term predictive power. It's worth taking a closer look at some of these because they offer a potential starting point for investors as opposed to traders. A good example is a recent working paper that scraped the text of bank's 10-K filings, looking for lots of negative words. The authors found that a one standard deviation increase in the number of negative words used led to a hefty +58% increase in the probability the bank would suffer a distressed delisting in the next three years (see Exhibit 3). Increasing 10-K negativity also implied a lower chance of paying dividends in the following year, higher year-ahead loan loss provisions, and a lower return on assets.

Last year we used a similar approach, and the same dictionary of positive and negative words, to parse the language used in company conference calls with analysts.[1] Using a broader set of companies than just banks we found that future failure stocks (defined as those in the worst 10% of performers in the following year) tended to have more *positive* conference call sentiment whereas future winners were generally more dour (see Exhibit 4). Unsurprisingly, company managements tend to be more positive in their prepared remarks at the start of the conference call than in the Q&A session at the end.

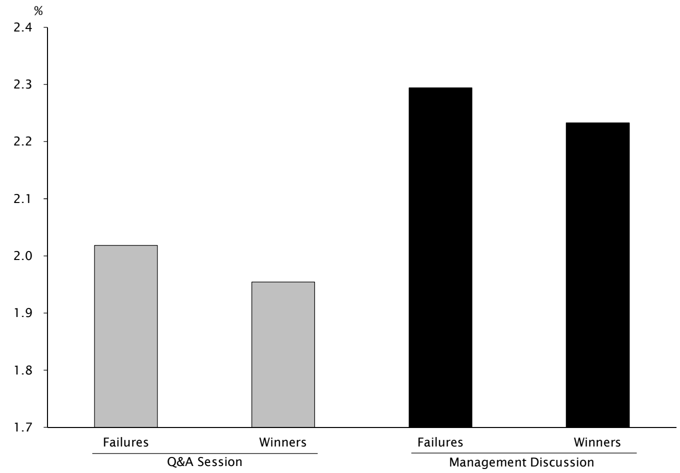
Exhibit 3: U.S. Publicly-Traded Banks
Change in Key Variables for a One Standard Deviation Increase in the Percent of Negative Words Used in the 10-K[1]
1997 Through 2014



Source: Gandhi, P., Lougran, T., and Bill McDonald, 2017. "Using Annual Report Sentiment as a Proxy for Financial Distress in U.S. Banks." Working Paper.

[1] After controlling for capital adequacy, market capitalization, 10-K readability, and prior year relative return.

Exhibit 4: Large-Capitalization Stocks (ex-Energy & Materials)
Net Positive Conference Call Sentiment[1]
2010 Through 2015



Source: FactSet Research Systems, Empirical Research Partners Analysis.

[1] Sentiment based on the Loughran-McDonald Dictionary, available at https://www3.nd.edu/~mcdonald/Word_Lists.html.

We can also aggregate conference call sentiment to the market level, something that one of the papers in Appendix 1 also tried.[2] Exhibit 5 shows the aggregate conference call sentiment for global stocks, overlaid with a market index. In our analysis we found the sentiment expressed in conference calls closely tracks the performance of stocks, but it's mostly coincident rather than leading. On the other hand, the academics did find some evidence that high aggregate sentiment is associated with lower future market returns, so the topic is potentially worth a closer look.

Another interesting paper found that *changes* in the management discussion and analysis section (MD&A) and footnotes of 10-K filings are associated with negative year-ahead returns, irrespective of the sentiment embedded in the text (see Exhibit 6). For example, a one standard deviation change in the textual content of the MD&A section was associated with a stock price underperformance of (3) percentage points in the following year, see the left-most bar, and an ROA decline of more than (50) basis points, see the third bar. The authors weren't able to prove *why* that was the case, but one possibility is that managements mostly copy-and-paste last year's disclosures unless something has gone wrong.

Six Degrees of Separation

Network effects are another popular Big Data topic these days. A good example of a paper from that genre is one that used NLP to scrape the text of 10-K filings looking for any reference to states where the firm does business. For each of the 50 states the author used an econometric model to forecast future economic activity and then at the company level he weighted those forecasts by the text-based exposure of the company to each state. Thus the forecast for each stock becomes its exposure to the expected economic activity of the network of states that it operates in. It turned out that stocks in the highest quintile of this so-called Predicted Regional Economic Activity (PREA) factor outperformed those in the lowest quintile by around +6 percentage points on an equally-weighted basis or +5 points cap-weighted over the following year (see Exhibit 7).

---

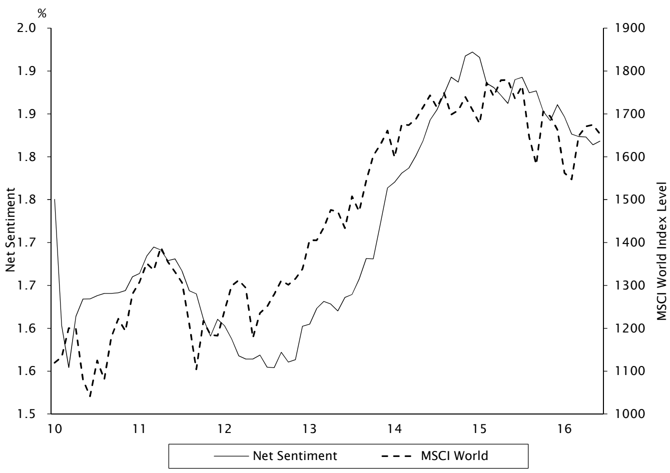
[1] *Stock Selection: Research and* Results January 2016. "The Sound of Failure: Parsing Conference Call Language for Red Flags."

[2] Jiang, F., Lee, J., Martin, X., and Guofu Zhou, 2015. "Manager Sentiment and Stock Returns." Working Paper.

There are a number of other network effect papers listed in Appendix 1, ranging from networks of firms selling similar products to those that are co-mentioned in the same news stories, but none of them found much alpha beyond a one-month holding period.

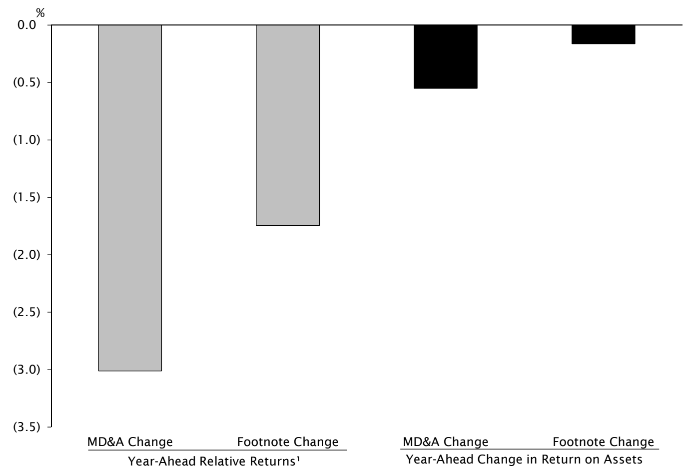
A final paper that did show some longer-term efficacy used abnormally high search volume on Yahoo Finance as a measure of investor attention (see Exhibit 8). The authors sorted stocks by their abnormal search volume (i.e., controlling for the search activity the company normally generates) right before earnings announcements and then tracked the stocks' returns over the following year. To control for the actual earnings surprise they further sorted stocks based on that variable. As the chart shows, regardless of the actual surprise the stocks in the highest quintile of abnormal search volume (grey bars) performed better than those in the lowest quintile (black bars) over the subsequent year.

Exhibit 5: Global Stocks
Average Net Sentiment in Conference Calls and MSCI World Index
2010 Through June 2016



Source: Empirical Research Partners Analysis.

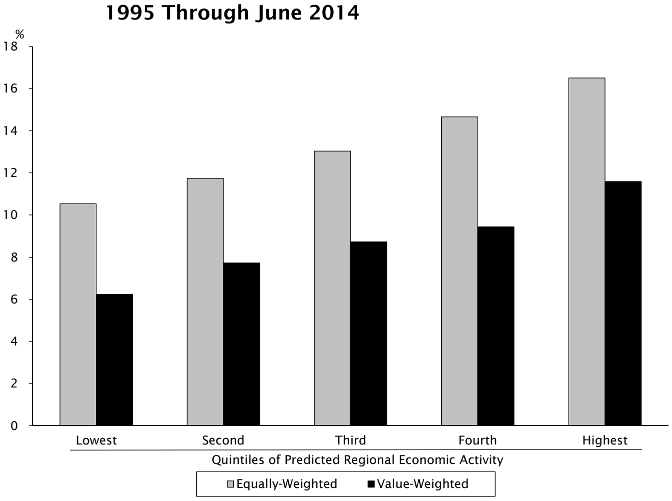
Exhibit 6: U.S. Stocks
Impact on Year-Ahead Relative Returns and Year-Ahead Change in ROA for a One Standard Deviation Change in the Textual Content of the 10-K MD&A and Footnotes
1994 Through 2014



Source: Amel-Zadeh, A., and Jonathan Faassee, 2016. "The Information Content of 10-K Narratives: Comparing MD&A and Footnotes Disclosures." Working Paper.

[1] After controlling for the Fama-French-Carhart factors.

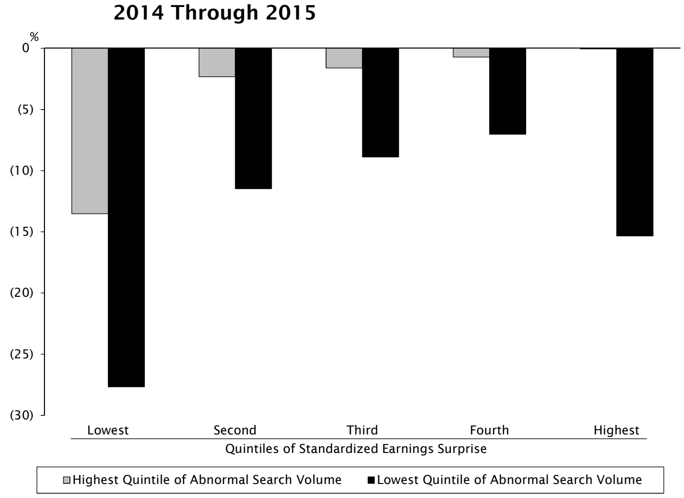
Exhibit 7: U.S. Stocks
Nominal Returns by Quintile of Predicted Regional Economic Activity (PREA)[1]
Monthly Data Compounded to Annual Periods
1995 Through June 2014



Source: Esad Smajlbegovic, 2016. "Regional Economic Activity and Stock Returns." Working Paper.

[1] Nominal returns are less the risk-free rate. PREA is a measures a firm's exposure to leading economic indicators for each state it is economically active in, where the weights are based on a textual analysis of the firm's 10-K filings.

Exhibit 8: U.S. Stocks
Relative Returns in the Year Following an Earnings Announcement by Quintile of Earnings Surprise and Abnormal Yahoo Finance Search Volume[1]
2014 Through 2015



Source: Empirical Research Partners Analysis.

[1] Relative returns begin two days after the announcement day, i.e., they capture post-announcement drift but not the announcement-day reaction.

Other papers have studied Google search volume and EDGAR searches (the SEC's filing website) but this paper shows that Yahoo Finance search volume subsumes both in terms of predictive power for future returns. However, a couple of major caveats apply: the study only covered a short two-year period and Yahoo Finance search data isn't readily available online like Google search data is.

Conclusion: Text-Mining Shows the Most Promise

The bulk of the research on Big Data gravitates towards trading rather than investing time-horizons. Our read of the literature so far is that the most promising avenue for long-term investors is to use text-mining to capture the nuance and context of what company managers are saying. We've long found that it's more important to watch what managements actually do with their capital rather than listen to what they say they'll do, but a preponderance of academic work suggests there's probably some incremental value in parsing the spin that managers put on the cold, hard numbers. We've added the topic to our research agenda for this year and we'll extend the technology we built to process conference call transcripts to other text-based data sources.

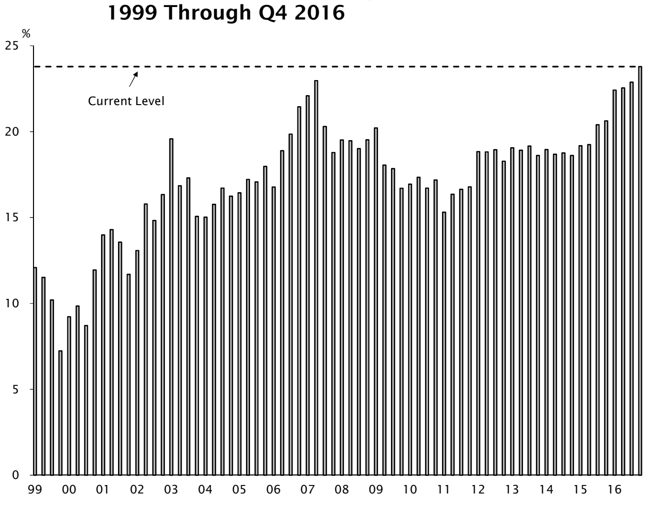
The Robots Ate My Alpha: The Rise of the Quant Funds

Putting the Quants Under the Portfolio Analytics Microscope

Your author's 14-month-old son is currently obsessed with "Robots, Robots Everywhere," a story about robots that can travel up in space, dive under the sea, and even milk a cow! In talking with our fundamental clients many share a similar obsession; there's a perception that the rise of the big quant managers has altered the game in some hard-to-pin-down way. It is certainly true that quants have gained a bigger share of the pie recently, as of the end of the fourth quarter last year quants accounted for 24% of the U.S. equity holdings of all hedge funds (see Exhibit 9). That's a new high-water mark, having now exceeded the previous peak in the summer of 2007, right before the so-called Quant Crisis in August of that year. We've found in past research the quants represent something like half the turnover of all hedge funds, even though they only control a quarter of hedge fund equity assets.

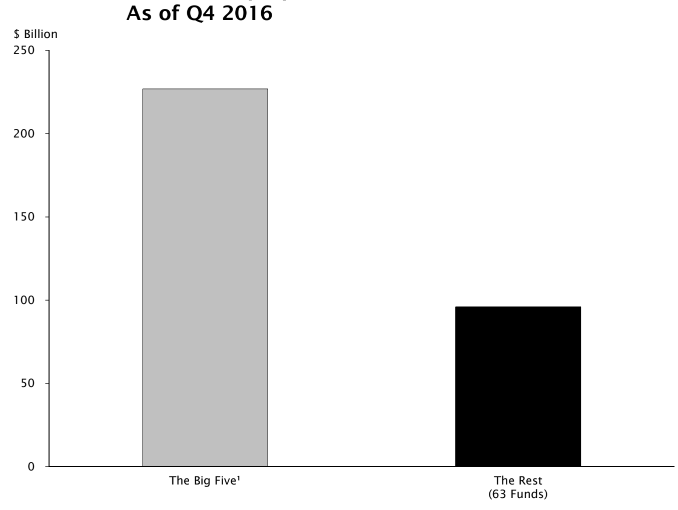
It's also noteworthy that quant equity assets are very concentrated in the five largest funds (see Exhibit 10). Collectively the Big Five control 70% of all quant equity assets, whereas the top five fundamental hedge funds only manage 15% of non-quant equity assets. All of the Big Five quants rank in the top 10 hedge funds by U.S. equity assets and currently the two largest quants occupy the first two spots on the list.

**Exhibit 9: Hedge Funds' U.S. Equity Holdings
Share Held by Quant Hedge Funds
1999 Through Q4 2016**



Source: FactSet Research Systems, Empirical Research Partners Analysis.

**Exhibit 10: Quant Hedge Funds
Total U.S. Equity Assets
As of Q4 2016**



Source: 13F Filings, Empirical Research Partners Analysis.

[1] Five largest quant managers by U.S. equity assets.

Last year we wrote in detail on the quants, but with the launch of our new Portfolio Analytics toolkit we thought it was worth running their portfolios through that framework to see if we could learn anything about the bets they're taking.[3,4] Before diving in, a few of the usual caveats are worth reiterating: 13F filings are lagged, they only capture

[3] *Portfolio Strategy* July 2016. "Quant Hedge Funds: Menacing Machines?"

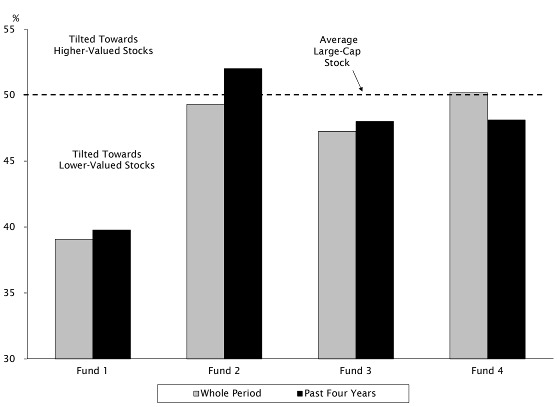
[4] *Portfolio Analytics* January 2017. "Empirical Analyses of Decision Making."

the long side, and they don't tell us anything about what happens intra-quarter. The latter two are particularly problematic for quant managers because they tend to be market neutral and generally have high turnover rates. A factor exposure that's observed on the long side could very well be completely hedged on the short side; after all quants think in terms of factors and they optimize their portfolios to take deliberate bets on some factors and neutralize others. Nonetheless, we thought the exercise was worth doing, just in case anything interesting jumped out of the black boxes. We focused on the four funds from the Big Five that have been around since 2004, giving us more than a decade of data to work with.

Exhibits 11 show the average exposure of the four quant funds to our valuation framework. The bars in the chart show the weighted-average percentile rank of the stocks in the managers' long portfolios, scored on our valuation factor. We've averaged across two time periods to get a better sense for whether the funds have changed their bets over time. For example, Fund 1 has very consistently held stocks that score around the 40th percentile market-wide on valuation, i.e., lower-valued stocks, whereas the other funds have kept their value bets much closer to neutral.

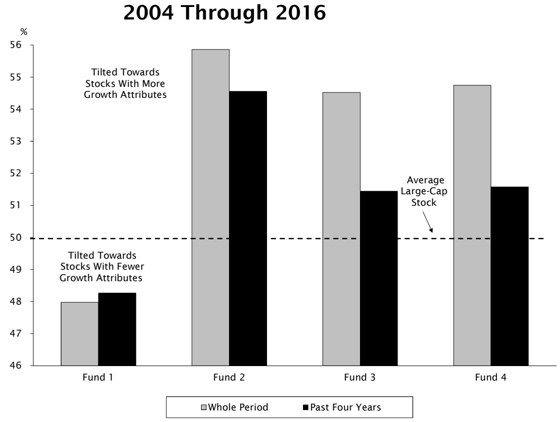
Taking the flipside, Exhibits 12 show the funds' exposure to our growth score, a framework we use to measure the all-around "growthiness" of stocks based on attributes like their past and expected top-line growth rates, their ROE and reinvestment rates, and their valuation multiples. As we'd expect, the value-oriented Fund 1 tends to be underweight stocks with lofty growth credentials whereas the other funds have mostly had a pro-growth stance, although Funds 3 and 4 have reined in that exposure in the past four years.

Exhibit 11: Select Quant Hedge Funds and the Large-Cap Universe Valuation Analysis: Average Percentile Rank of Stocks in Portfolio 2004 Through 2016



Source: Empirical Research Partners Analysis.

Exhibit 12: Select Quant Hedge Funds and the Large-Cap Universe Growth Score Analysis: Average Percentile Rank of Stocks in Portfolio 2004 Through 2016



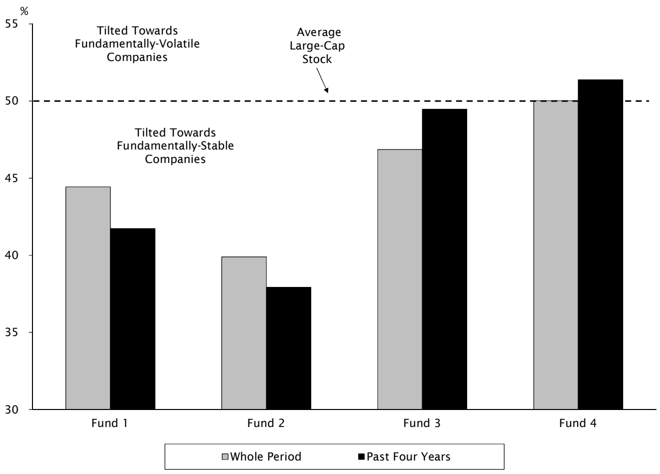
Source: Empirical Research Partners Analysis.

One of the defining characteristics of equity markets post-Crisis has been the quest for stability, often articulated by holding bond-like stocks for their income and purported safe-haven status. We measure fundamental stability using a scorecard that screens for firms with consistent growth rates and ROEs, predictable earnings, low financial leverage, and low betas. We took a look at how the quants have been positioned relative to that measure of fundamental stability (see Exhibit 13). There's a clear bifurcation among the funds: Funds 1 and 2 have been big proponents of stability while Funds 3 and 4 have been more circumspect.

What's interesting is that Fund 2, while embracing *fundamental* stability has also recently sought out companies with above-average arbitrage risk, a metric we use to measure controversy (see Exhibit 14). That's not necessarily mutually exclusive, because arbitrage risk is a shorter-term measure that looks at whether recent trading activity can be explained by the stock's beta and the market's move. If it can't, then something is often afoot, and sometimes not in a good way. If you have an Edge in stock-picking then it can make sense to embrace high arbitrage risk, and it's telling that Fund 3 and 4 also lean that way.

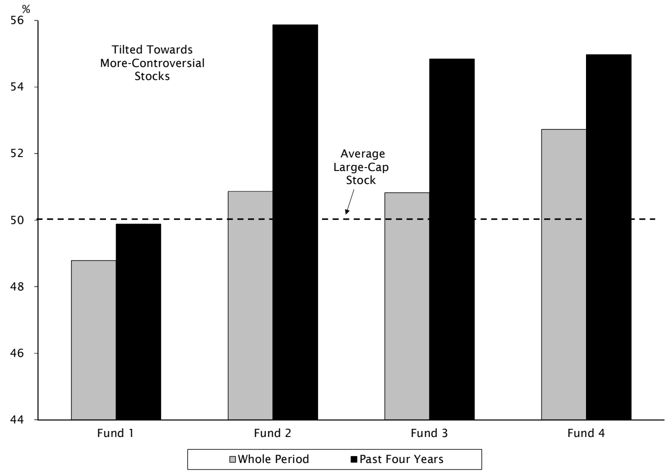
Quants have long had the reputation of being momentum junkies and that does show up in the data, although probably not to the extent that popular lore would suggest (see Exhibit 15). Here we've plotted the exposure of each fund to nine-month price momentum and it turns out Fund 1 has the largest and most consistent tilt towards high-momentum stocks. So Fund 1 would seem to deploy something like the classic value-momentum combination (recall Exhibit 11). The other funds are on average trend-followers too, but not to the same degree.

Exhibit 13: Select Quant Hedge Funds and the Large-Cap Universe
Fundamental Stability Analysis:
Average Percentile Rank of Stocks in Portfolio
2004 Through 2016



Source: Empirical Research Partners Analysis.

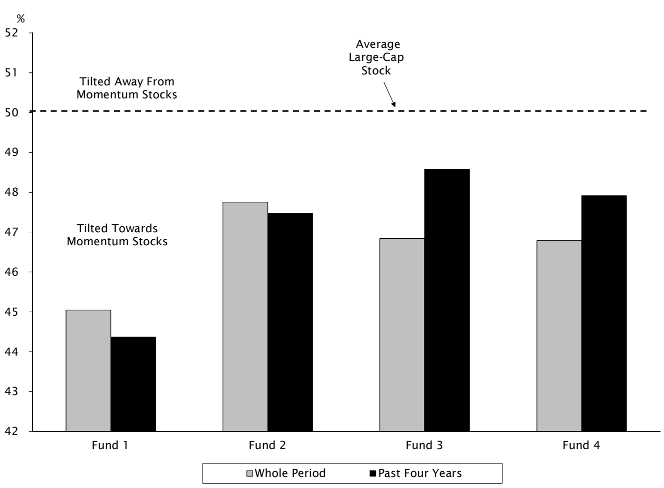
Exhibit 14: Select Quant Hedge Funds and the Large-Cap Universe
Arbitrage Risk Analysis:
Average Percentile Rank of Stocks in Portfolio
2004 Through 2016



Source: Empirical Research Partners Analysis.

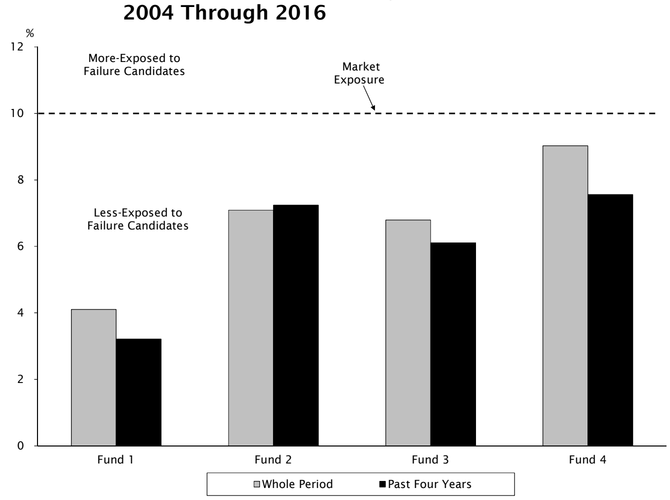
Finally we looked at the funds' exposure to Failure candidates, as identified by our Failure model (see Exhibit 16). This is another area where all the funds agree: avoiding failure is one of the best uses of quantitative tools. All four funds have a much lower-than-market exposure to Failure candidates and that stance has been fairly consistent over time. For Funds 2 through 4, which have tended to play in the high growth/high arbitrage risk space, this is reassuring because the odds in that sandbox are stacked against you and avoiding failure can go a long way towards swinging them back towards you.

Exhibit 15: Select Quant Hedge Funds and the Large-Cap Universe
Nine-Month Price Trends Analysis:
Average Percentile Rank of Stocks in Portfolio
2004 Through 2016



Source: Empirical Research Partners Analysis.

Exhibit 16: Select Quant Hedge Funds and the Large-Cap Universe
Failure Model Analysis: Average Weight of Failure Candidates in Long Portfolio
2004 Through 2016



Source: Empirical Research Partners Analysis.

Conclusion: Robots Have Feelings Too

Overall the robots don't look so different from you or me. At the micro-level their approaches can be quite different, even within the quant sphere. For example, Fund 1 represents what we would call more of a fundamental quant approach, taking fairly transparent, linear bets towards things like valuation and momentum. The other managers have gravitated towards Big Data, machine-learning strategies. That often entails running a stable of thousands of individual, and often non-linear, alpha signals that might apply to only a handful of stocks or all stocks. For example, one alpha signal might scrape blog postings for chatter about Apple's latest iPhone, in the hopes of getting a read on next quarter's earnings. Another alpha signal might use, as Bobby Axelrod did, satellite images to try to gauge manufacturing in China and then correlate that with a bunch of global growth-exposed stocks. Each signal on its own might have a hit rate only a smidgen above 50/50, but the idea is to flip a slightly biased coin thousands of times a day. If you can do that the odds start to skew in your favor. Instead of a portfolio of stocks, a Big Data quant builds a portfolio of little alpha drones and hopes they are uncorrelated enough that the law of large numbers kicks in.

Nonetheless, despite the differing approaches, once one rolls things up to the portfolio level the quants don't look that different from everybody else. For example, the profile of Fund 2 looks quite a bit like a typical long-only growth manager. Again, all the usual caveats apply of course since we have no idea what goes on intra-quarter or on the short side, but we'd be inclined to say the robots look a lot like humans, just with much better discipline, witness for example their near ironclad avoidance of Failure candidates. And as we've always said in our work, it's discipline that mostly separates investing success from failure. We can't all be Bobby Axelrod, but we all can use systematic tools to help skew the odds in our favor. They won't write a TV show about that but it works.

Exhibit 17, below, lists stocks most heavily owned by the Big Five quant funds, as of their latest filings.

Exhibit 17: Large-Capitalization Stocks with High Quant Hedge Fund Ownership[1]
Sorted by Quant Hedge Fund Ownership
As of Early-March 2017

				Quintiles (1=Best; 5=Worst)								
				Select Metrics								
Symbol	Company	Price	Ownership By Big Five Quant Hedge Funds	Valuation (1=Lowest)	Growth Score (1=Highest)	Fundamental Stability (1=Most Stable)	Arbitrage Risk (1=Lowest)	Nine-Month Price Trends (1=Highest)	Core Model Rank (1=Best)	Failure Candidate?	YTD Returns	Market Capitalization ($ Billion)
VRSN	VERISIGN INC	$82.47	8.9 %	1	1	3	2	5	1	No	8.4 %	$8.5
SPR	SPIRIT AEROSYSTEMS HOLDINGS	61.61	7.9	2	2	5	4	1	1	No	5.6	7.5
H	HYATT HOTELS CORP	51.34	7.8	2	4	5	3	3	2	No	(7.1)	6.7
X	UNITED STATES STEEL CORP	38.72	7.0	4	5	5	5	1	1	No	17.5	6.7
VMW	VMWARE INC -CL A	89.89	6.6	2	2	1	3	1	1	No	14.2	36.7
UTHR	UNITED THERAPEUTICS CORP	147.72	5.7	1	1	4	5	1	1	No	3.0	6.6
OC	OWENS CORNING	58.49	5.7	1	4	3	3	3	2	No	13.4	6.6
DPZ	DOMINO'S PIZZA INC	189.81	5.5	5	1	2	3	1	1	No	19.2	9.1
FFIV	F5 NETWORKS INC	143.27	5.5	3	1	1	4	1	3	No	(1.0)	9.3
CBOE	CBOE HOLDINGS INC	78.05	5.5	5	1	1	3	2	5	No	5.6	6.3
HII	HUNTINGTON INGALLS IND INC	218.50	5.4	3	2	2	3	1	2	No	18.6	10.1
ETR	ENTERGY CORP	76.66	5.3	1	5	3	2	4	2	No	5.6	13.8
RIG	TRANSOCEAN LTD	13.82	5.3	1	5	5	5	1	1	No	(6.2)	5.4
STLD	STEEL DYNAMICS INC	36.60	5.2	2	3	4	5	1	1	No	2.9	8.9
LVNTA	LIBERTY VENTURES	43.86	5.1	5	1	5	4	1	3	No	19.0	12.5
WCG	WELLCARE HEALTH PLANS INC	141.20	5.0	1	3	3	2	1	1	No	3.0	6.3
FE	FIRSTENERGY CORP	32.43	4.8	1	5	4	3	5	3	No	6.0	14.3
ALSN	ALLISON TRANSMISSION HLDGS	35.98	4.8	3	3	4	3	1	2	No	6.8	6.0
JAZZ	JAZZ PHARMACEUTICALS PLC	132.62	4.7	2	1	4	5	5	3	No	21.6	7.9
EXEL	EXELIXIS INC	21.53	4.7	5	2	5	5	1	2	No	44.4	6.3
BURL	BURLINGTON STORES INC	89.01	4.6	4	1	1	3	1	1	No	5.0	6.3
FTNT	FORTINET INC	37.35	4.4	3	1	3	5	3	3	No	24.0	6.5
CCK	CROWN HOLDINGS INC	53.59	4.3	1	5	4	2	5	3	No	1.9	7.5
TSO	TESORO CORP	85.19	4.3	1	5	5	4	3	1	No	(2.0)	10.0
MNK	MALLINCKRODT PLC	52.42	4.2	1	5	5	5	5	2	No	5.2	5.6
TDC	TERADATA CORP	31.10	4.1	1	5	4	4	3	1	No	14.5	4.0
PVH	PVH CORP	91.60	4.1	1	4	2	5	5	1	No	1.5	7.3
PBI	PITNEY BOWES INC	13.64	3.9	1	5	1	5	5	1	No	(8.9)	2.5
SRCL	STERICYCLE INC	82.88	3.8	1	3	1	5	5	4	No	7.6	7.1
KLAC	KLA-TENCOR CORP	90.12	3.8	3	3	3	3	2	1	No	15.2	14.1
LEA	LEAR CORP	141.99	3.8	1	2	3	3	2	1	No	7.3	9.9
DHI	D R HORTON INC	32.00	3.8	4	2	2	4	4	5	Yes	17.5	11.9
AKAM	AKAMAI TECHNOLOGIES INC	62.60	3.7	2	2	1	5	3	3	No	(6.1)	10.9
BERY	BERRY PLASTICS GROUP INC	50.33	3.7	1	1	3	2	2	1	No	3.3	6.5
SEE	SEALED AIR CORP	46.48	3.6	3	5	4	4	5	4	No	2.5	9.0
LLTC	LINEAR TECHNOLOGY CORP	64.58	3.6	5	1	1	1	1	3	No	4.1	15.5
MRVL	MARVELL TECHNOLOGY GROUP LTD	15.60	3.6	5	5	4	3	1	3	No	12.5	7.9
PNW	PINNACLE WEST CAPITAL CORP	82.19	3.6	4	4	1	1	3	3	No	6.2	9.2
CE	CELANESE CORP	89.17	3.6	3	3	3	2	2	2	No	13.7	12.6
BBY	BEST BUY CO INC	44.13	3.6	1	4	4	4	1	1	No	3.4	13.8
LRCX	LAM RESEARCH CORP	118.54	3.5	2	2	3	4	1	1	No	12.1	19.3
MXIM	MAXIM INTEGRATED PRODUCTS	44.30	3.5	4	2	2	4	2	1	No	15.7	12.5
LSXMA	LIBERTY MEDIA SIRIUSXM GROUP	39.33	3.5	3	3	5	4	2	2	No	13.9	16.7
NFX	NEWFIELD EXPLORATION CO	36.46	3.4	5	5	5	5	5	5	No	(10.0)	7.3
MAS	MASCO CORP	33.78	3.4	4	5	2	3	4	4	No	7.2	10.8
DOX	AMDOCS	60.65	3.4	2	3	1	1	4	3	No	4.1	8.9
THO	THOR INDUSTRIES INC	110.82	3.4	3	1	2	3	1	3	No	10.8	5.8
IR	INGERSOLL-RAND PLC	79.36	3.3	3	3	3	1	2	2	No	5.8	20.6
NI	NISOURCE INC	23.91	3.3	4	5	2	2	5	5	Yes	8.8	7.7
CC	CHEMOURS CO	33.66	3.3	3	3	5	5	1	1	No	52.5	6.2
SQ	SQUARE INC	17.32	3.3	5	3	5	5	1	4	No	27.1	6.4
R	RYDER SYSTEM INC	76.15	3.3	1	4	4	4	3	2	No	2.9	4.1
AVY	AVERY DENNISON CORP	80.71	3.3	3	2	2	4	4	3	No	15.5	7.2
AEE	AMEREN CORP	54.69	3.3	2	5	2	1	3	2	No	4.3	13.3
IAC	IAC/INTERACTIVECORP	73.94	3.2	2	5	5	4	1	1	No	14.1	5.9
LLL	L-3 COMMUNICATIONS HOLDINGS INC	168.32	3.2	2	5	3	3	2	2	No	11.1	13.1
UGI	UGI CORP	48.23	3.2	2	4	2	2	3	2	No	4.7	8.3
HOLX	HOLOGIC INC	40.58	3.2	4	2	5	2	3	3	No	1.1	11.3
GXP	GREAT PLAINS ENERGY INC	29.06	3.2	2	4	1	1	5	4	No	7.3	6.3
FLEX	FLEX LTD	16.49	3.2	2	4	3	3	1	2	No	14.8	8.8
TSN	TYSON FOODS INC -CL A	62.56	3.2	1	4	1	3	5	1	No	1.8	22.3
AIZ	ASSURANT INC	99.00	3 2.	2	4	3	3	3	1	No	7.2	5.5

Source: 13F Filings, Empirical Research Partners Analysis.

[1] Based on holdings of the Big Five quant hedge funds, i.e., the five with the most assets in U.S. equities, as of the Q4 2016 13F filings.

**Appendix 1: Select Recent Academic Papers on Big Data
Sorted by Subject and Date[1]
2013 Through Early-March 2017**

Authors	Year	Title	Summary	Time-Horizon
Natural Language Processing (NLP)				
Gandhi, P., Lougran, T., and Bill McDonald	2017	Using Annual Report Sentiment as a Proxy for Financial Distress in U.S. Banks	Negative sentiment in the annual reports of banks predicts higher probability of distressed delisting in next three years, lower probability of paying dividends next year, and higher year-ahead loan losses and lower ROA.	Year-ahead financial performance; three-year ahead probability of delisting
Dexin Zhou	2017	Good News in Numbers	Uses NLP to measure the ratio of qualitative-to-quantitative information in company conference calls. Finds those with more quantitative information have positive announcement-day returns and post-announcement drift. The theory is that managers use qualitative content to try to spin or disguise negative news.	Three-day announcement-day returns; 60 day post-announcement returns; quarter-ahead earnings surprise
Filzen, J., McBrayer, G., and Kyle Shannon	2016	Risk Factor Disclosures: Do Managers and Markets Speak the Same Language?	Tracks changes in risk factor language in 10-Qs from quarter-to-quarter. Finds those that update their language have lower future returns.	Three-month post-announcement returns
Palmon, D., Xhu, K., and Ari Yezegel	2016	What Does 'But' Really Mean? Evidence from Managers' Answers to Analysts' Questions During Conference Calls	Use of constrastive words like 'but' in the analyst Q&A section of earnings calls is associated with a greater reaction to earnings news. Theory is that contrastive words preceed value-relevant information.	Three-day announcement-day returns
Amel-Zadeh, A., and Jonathan Faassee	2016	The Information Content of 10-K Narratives: Comparing MD&A and Footnotes Disclosures	Investors react more strongly in the short-term to the content of the MD&A section of 10-Ks, as opposed to the footnotes. Changes in the textual content of both MD&A and footnotes are associated with lower future returns and ROA. Doesn't provide a strong theory for why that should be the case.	Four-day announcement-day returns; one-year post-announcement returns; year-ahead ROA
Hwang, B.-H. and Hugh Hoikwang Kim	2016	It Pays to Write Well	Uses NLP to assess the readability of close-ended investment companies' annual reports. Finds that those with less readable reports trade at a larger discount to NAV.	Spot discount to NAV
Jiang, F., Lee, J., Martin, X., and Guofu Zhou	2016	Manager Sentiment and Stock Returns	Construct a company manager sentiment index based on the text of conference calls and earnings releases. Find it negatively predicts future market returns and earnings growth, i.e., high optimism is a warning sign. Predictive power is stronger than that of *investor* sentiment.	Three-year market returns with peak predictability at nine months; 12-month forward aggregate S&P 500 earnings growth; one-month cross-sectional stock returns
Fishe, R., North, D., and Aaron Smith	2014	Words that Matter for Asset Pricing: The Case of IPOs	Words in an IPO prospectus that are associated with negativity and uncertainty predict first-day IPO returns	One-day IPO returns
Networks				
Parsons, C., Sabbatucci, R., and Sheridan Titman,	2017	Geographic Momentum	A firm's future returns can be predicted by the lagged returns of peer firms that are headquartered in the same metropolitan area. The effect is stronger than industry or same-firm momentum.	One-month returns
Alldredge, D. and Andy Puckett	2016	The Performance of Institutional Investor Trades Across the Supply Chain	Show that institutional investors tend to invest along supply chains, e.g., a manager who owns a customer firm is five times more likely to also hold a supplier of that firm. Find that institutions tend to generate abnormal profits in the supplier firms, paticularly when they are small.	Quarter-ahead profitability of institutional trades
Esad Smajlbegovic	2016	Regional Economic Activity and Stock Returns	Uses NLP on company 10-K filings to identify mentions of state-level economic activity. Then constructs a weighted leading indicator for each company based on the leading indicators of the states it is exposed to. Shows that signal predicts future stock returns.	One-month to three-year returns
Cao, J., Chordia, T., and Chen Lin,	2016	Alliances and Return Predictability	The lagged returns of strategic partners predicts future stock returns.	One-month returns
Binying Liu	2016	Circle of Competence and the Gradual Diffusion of Information in Prices	Uses the social accounting matrix from the Bureau of Economic Analysis (which measures flow of commodities through industries) to determine connected firms. Shows that past returns of connected firms predicts the future returns of the other firm.	One-month returns
Hoberg, G. and Gordon Phillips	2015	Product Market Momentum	Uses NLP on 10-K filings to identify the business segments that a company operates in. Shows that past returns of business segment peers predict future returns to a much greater degree than industry or same-firm momentum.	One-month returns
Wu, J. and John Birge	2015	Supply Chain Network Structure and Firm Returns	The past returns of a firm's suppliers predicts its future returns. Also manufacturing firms that are located more centrally in a supply chain have lower future returns while logistic firms that are more central have higher future returns.	One-month returns
Scherbina, A. and Bernd Schlusche	2015	Economic Linkages Inferred from News Stories and the Predictability of Stock Returns	Uses co-mentions in news stories to identify firms that might be economically linked to each other. Finds that past returns of a linked firm can predict future returns for the other firm, even after controlling for customer-supplier links.	One-month returns

Source: Empirical Research Partners Analysis.

[1] Date of last revision used for working papers and publication date used for papers published in peer-reviewed academic journals.

Appendix 1 (Cont.): Select Recent Academic Papers on Big Data
Sorted by Subject and Date[1]
2013 Through Early-March 2017

Authors	Year	Title	Summary	Time-Horizon
Social Media and Internet Search				
Lawrence, A., Ryans, J., Sun, E., and Nikolay Laptev, 2016	2016	Yahoo Finance Search and Earnings Announcements	Finds that the volume of Yahoo Finance searches for a ticker predicts future stock returns. Also shows that Yahoo Finance searches subsume the predictive power of Google and EDGAR (the corporate filing website of the SEC) searches.	Two-day announcement-day returns; one-year post-announcement returns
Bartov, E., Faurel, L., and Partha Mohanram	2016	Can Twitter Help Predict Firm-Level Earnings and Stock Returns?	Aggregate sentiment in tweets about a company predict future earnings surprises and announcement-day returns.	10-day ahead earnings surprise; three-day announcement-day returns
Vicki Wei Tang	2016	Wisdom of Crowds: Is Nonfinancial Information Disseminated on Twitter Informative About Future Fundamentals?	Aggregate sentiment and volume of tweets about a company predicts upcoming fiscal year sales growth, particularly for firms that sell to consumers.	Quarter-ahead fiscal year sales growth
Storms, K., Kapraun, J., and Markus Rudolf	2015	In Search of Alpha - Trading on Limited Investor Attention	Shows that stocks with abnormally low Google Search Volume scores outperform in the following week. The theory is that such stocks are underpriced because investors haven't been focused on them.	One-week returns
Chen, H., De, P., Hu, Y., and Byong-Hyoun Hwang	2014	Wisdom of Crowds: The Value of Stock Opinions Transmitted Through Social Media	Uses NLP to parse posts and comments on SeekingAlpha.com. Finds that the fraction of negative words in the original posts and the follow-up comments both predict negative future stock returns. These returns do not reverse so the authors conclude that SeekingAlpha content does indeed contain value-relevant information.	Three-month returns
Sprenger, T., Tumasjan, A., Sandner, P., and Isabell Welpe	2014	Tweets and Trades: The Information Content of Stock Microblogs	Tweet sentiment predicts future stock returns and tweet volume predicts future trading volume. Also show that users who provide better investment advice tend to be retweeted more, amplifying their impact.	Two-day returns
Securities Lending Data				
Stratman, T. and John Welborn	2016	Informed Short Selling in High Fail-to-Deliver Stocks	Find stocks with high fail-to-deliver rates tend to underperform in the future.	Three-day returns
Huszar, Z., Tan, R.S.K., and Weina Zhang	2016	Stock Lending from Lenders' Perspective: Are Lenders Price Takers?	Uses detailed securities lending data to show that stock lenders are not passive players. Rather, post the financial crisis they have become more proactive in pricing their lendable inventory. Raising stock lending fees in anticipation of bad news can hinder price discovery.	20-day returns
Beneish, M., Lee, C., and Craig Nichols	2015	In Short Supply: Equity Overvaluation and Short Selling	Use detailed securities lending data to model supply and demand of lendable securities. Show that when supply of lendable securities is binding then supply is the main predictor of future returns. Also show that some well known anomalies disappear when cost of borrow is considered.	One-month returns
Drechsler, I. and Qingyi Song Drechsler	2014	The Shorting Premium and Asset Pricing Anomalies	Uses detailed securities lending data to show the cost of shorting a stock negatively predicts future returns. Eight of the most well known asset pricing anomalies disappear within the 80% of stocks that are easy to borrow, suggesting transaction costs are a big part of the apparent profitability of these anomalies.	One-month returns
Lynch, A., Nikolic, B., Yan, X.S., and Han Yu	2014	Aggregate Short Selling, Commonality, and Stock Market Returns	Uses detailed securities lending data to show that aggregrate short selling activity predicts future market returns.	20-day market returns
News Sentiment and Media Coverage				
Engelberg, J., McLean, R.D., and Jeffrey Pontiff	2016	Anomalies and News	Studies 97 return anomalies identified in the academic literature and finds that anomaly returns are seven times higher on earnings announcement days and two times higher on corporate news days. Argues this proves that many anomalies are the result of biases expectations that are corrected when new information comes out.	One-day returns
Akansu, A., Cicon, J., Ferris, S., and Yanjia Sun	2016	Firm Performance in the Face of Fear	Uses facial recognition software to study CEO interviews. Finds that firms where the CEO exhibits fear or disgust tend to show improved profitability in the next quarter. Argues that such emotions motivate the CEO to work harder to improve his/her plight.	Quarter-ahead profitability
Hillert, A. and Michael Ungeheuer	2016	The Value of Visibility	Uses an archive of New York Times coverage of companies since 1924 and finds that companies that are persistently mentioned outperform in the future and also deliver improved fundamentals.	One-, two-, and three-year returns
Erik Mayer	2016	Investor Attention and Stock Prices: Evidence from a Natural Experiment	Studies the stock price reaction of companies that sponsor College football bowl games around the date of each game. Bowl games deliver the sponsor considerable airtime, particularly major games like the Rose Bowl. Finds that sponsoring firms' share prices do outperform around bowl games.	10-day returns
Hadzic, M., Weinbaum, D., and Nir Yehuda	2015	News Content, Investor Misreaction, and Stock Return Predictability	Uses an archive of Reuters news stories to study instances where the release-day stock price reaction was in the *opposite* direction to the textual sentiment of the story. Find that in cases of initial misreaction the stock price eventually drifts in the direction of the textual sentiment.	One-month returns
Hafez, P. and Junqiang Xie	2014	Web News Analytics Enhance Stock Portfolio Returns	Compares news sentiment coverage based on the Dow Jones news wire to that derived from web content. Shows that web content is additive to a one-week reversal strategy.	One-week returns
Y. Han Kim and Felix Meshke	2013	CEO Interviews on CNBC	Studies the reaction of stocks when their CEOs are interviewed on CNBC. Find the stocks outperform in the run-up to the interview but then mean revert in the following two weeks. The pre-interview run-up tends to be mostly driven by retail investors.	10-day returns

Source: Empirical Research Partners Analysis.

[1] Date of last revision used for working papers and publication date used for papers published in peer-reviewed academic journals.

Appendix 1 (Cont.): Select Recent Academic Papers on Big Data
Sorted by Subject and Date[1]
2013 Through Early-March 2017

Authors	Year	Title	Summary	Time-Horizon
Options Data				
Zhenping Wang	2017	Option Trading Leverage and Stock Returns	Shows that volume-weighted options leverage is predictive of future stock returns. Also finds this metric subsumes previously proposed metrics like the option volume-to-stock volume ratio.	One-month returns
Ge, L., Lin, T.-C., and Neil Pearson	2015	Why Does the Option to Stock Volume Ratio Predict Stock Returns?	Digs deeper into past research that has shown the option volume-to-stock volume ratio is predictive of future returns by decomposing the source of the options volume. Finds that volume related to opening new call positions is most predictive.	One-week returns
Christoffersen, P. and Xuhui Pan	2014	Equity Portfolio Management Using Option Price Information	A good literature review of the recent research showing that options data can predict future stock returns.	One- to four-week returns
Ownership Data				
James Bulsiewicz	2016	Predicted Institutional Trades and the Cross-Section of Returns	Finds that lagged financial variables explain around 9% of the variation in future net institutional trades, in other words the net trading activity of institutions is somewhat predictable. Shows that investors can use predicted trades as a momentum signal.	One- to 20-quarter returns
Crowdsourcing				
Johnson, R., Kang, T., and Michael Wolfe	2016	Crowdsourcing forecasts: Competition for sell-side analysts?	Uses data from Estimize, a website that crowdsources earnings estimates, to show that these estimates predict quarter-ahead earnings surprises and post-announcement returns. They find they are complementary to IBES forecasts and tend to be more accurate in the short-term but less accurate in the long-term.	Quarter-ahead earnings surprises, two-week post-announcement returns

Source: Empirical Research Partners Analysis.

[1] Date of last revision used for working papers and publication date used for papers published in peer-reviewed academic journals.