Generating Alpha Through Data Science

Disruptive Forces in Investing

February 11, 2020

Anu Rajakumar: Back in July we sat down with Neuberger Berman's Chief Data Scientist, Michael Recce, to discuss how big data, the residue of information that we all leave behind as we buy things, sell things, browse, and live our modern lives can be used to evaluate the health of businesses. This, combined with advances in cloud computing, machine learning and artificial intelligence has become what is known as data science, allowing the ability to extract coherent, strategic investment insights that is deeper and more detailed, yet also broader and more comprehensive. To continue the conversation around how data is transforming the investment industry, I am pleased to be joined again by Michael Recce along with Richard Nackenson, Senior Portfolio Manager of the Nackenson Group, to discuss a real-life example of how thoughtful analysis of data can impact fundamental research. Michael and Rich, thank you for being here today.

Michael Recce: Great to be here, thanks.

Richard Nackenson: Thank you.

Anu: Now, Michael, in our last conversation, you explained that the alpha isn't in the raw data itself, but in how you use the data. Could you explain a little on how your work with data science aligns with traditional fundamental portfolio management teams like Rich's to enhance the conviction of investment views?

Michael: Super, yeah, thanks. I think it's important to start with the idea around how data is used in investing. And data is really an informational and analytical advantage. Michael Mauboussin, who's a professor at Columbia, has talked about this acronym of BAIT—Behavioral, Analytical, Informational, and Technical—as the four sources of alpha in the market. So behavioral and technical are much more related to market timing, and informational and analytical advantages are much more in the investing world. And data clearly is leaning towards a benefit for the longer term investing. So what we do is a deeper analysis of companies to leverage the analysis of this information so that we can actually make better decisions where we're prepared to wait for the market to discover the value.

Anu: And so now turning to you, Rich, in your view as an investor rooted in fundamental analysis, how has big data disrupted the traditional investment process in your opinion?

Richard: Disruption is an interesting choice of words because this is another critical tool that one has to use to succeed in a competitive marketplace. Creating alpha is not easy, and we attempt to do that, and, big data can be a powerful tool in achieving that. The key thing for us on the investment side, because we are fundamental investors, is we have an investment thesis. We start with that thesis and then we look at different ways to prove or disprove that thesis and then can we craft a specific question or series of questions that can be tested through big data analysis? And if we can achieve that and get clear, identifiable answers from that analysis, that's another way to win. So what big data has effectively done, it's created another way to analyze investments. And it's our job and the team, from a portfolio standpoint, to work with Michael and his team, to translate that data into actual insights. So it's very important to understand, because big data is out there, how do you utilize it, and are you utilizing it more effectively perhaps than a competitor out in the marketplace? What we at the team level try to figure out is what are the data sets that are relevant and important? And, just having it isn't what the critical issue is in terms of big data, but how do we utilize it? So, understand the data sets, set up a clear investment thesis or antithesis, and ask the right questions so we can figure out how to develop, validate, or disprove an investment idea.

Anu: And so maybe just on that theme, how do you go about ensuring that the data that you do have is indeed clean, useful, and actionable or that any pitfalls that you have to avoid or any examples you can share, maybe Michael?

Michael: Well, so I'll say that all data has bias. And so the most important thing you can do is to remove the bias, and the second thing you could do is to detect bias and understand how much bias you have in your data. So as soon as we get a data set, one of the things we do is we analyze that data in some dimension that we think is relevant and perhaps hasn't been looked at. So, in one of those examples at a theme park one of the parks with the largest gate receipts was not the largest park in our analysis, and it turned out that that theme park was next to the military's on-boarding base and all the G.I.'s were riding the roller coaster every day. That wasn't predictive of the nuclear family traveling to the theme park. So the best thing to do is to remove that park from the data. You get a much cleaner signal, and you can do better analysis. We also match the data against other sources of ground truth like census and other things in order to determine if there's bias in the data.

Anu: And so now your teams recently worked together on a deep, holistic evaluation of McDonald's. So, Rich, how did you come to the decision to use big data when evaluating McDonald's versus other companies that you were looking at?

Richard: Let's start with why we were looking at McDonald's in the first place. This is a company we were familiar with. In our investment process, McDonald's always had a superior free cash flow profile and also very high return on invested capital. And what we saw in our screening process was that the future looking free cash flow was inflecting negative for the next 12 to 18 months, which was quite unusual. When we're looking at free cash flow, we're looking at it in terms of what a company could actually do with that free cash flow, either to the balance sheet side; increasing cash, reducing debt, paying a dividend or buying back shares, or on the asset side; reinvesting in the company, doing a select acquisition, investing in an expansion project. So we disaggregated our analysis of free cash flow and realized that a great deal of the discrepancy had to do with the growth capex portion of it. So effectively, McDonald's was taking their cash flow and using it to reinvest in the business, and specifically in something called Experience of the Future, which is effectively these large touch screen kiosks that help in the ordering process, but also feed into McDonald's data set. And what we were trying to determine: would this be helpful and effective and a good use of growth capex in the North American theater because we'd seen it be very successful in the international regions. If we could somehow evaluate on a same-store sales basis, how could you track the effectiveness and/or the inflection point on a store-by-store basis of the investment of Experience of the Future? So we used this opportunity to work with our centralized research department and Michael's team and big data, and say, hey, can we figure out an analysis that works? Can we get a large data set to better understand the impact on same-store sales of the Experience of the Future growth capex investment? And could we see that this would actually work from a return on invested capital, and would it move the needle in same-store sales growth? And that was the concept behind our collaboration.

Anu: Sure, absolutely. So now at this point, Michael, your big data team gets looped in. What types of data did you look at and what did you find?

Michael: Great, yeah. So, when we're looking at this, data, what we want to do is we want to identify in, credit card data, so we have 10 percent of U.S. credit card at transaction level, so we want to identify when a McDonald's is being modernized for this experience of the future. So the way we would do that is we would take, the subset of those credit card, panel that actually are the most active visitors to McDonald's. Now, if all of the people in the panel stopped going to some particular McDonald's on some same date, then the chances are they didn't all just change their habits. Chances are that store's been closed. If it reopens a month later then that's the period of time it usually takes for them to modernize it. So then we're able to find the subset of McDonald's, which actually had been modernized, and then we were able to look at the customer engagement before the modernization and after the modernization to determine what impact it's actually had on commerce. And so this same-store sales is taking the established stores and looking at the average growth year over year of those stores. And so we're able to estimate that parameter that Richard was interested in.

Anu: And so now back to you, Rich. How did your teams monitor the result of what you're finding after getting this information from Michael's team?

Richard: The initial results were quite positive, so that enables us to take action from a portfolio standpoint. But the nice thing about this particular big-data-oriented analysis is that you could repeat it on an ongoing basis. Companies aren't going to tell you the answer. We have to figure it out ahead of time, and we're longer term investors. So, one, you can get ahead of the curve; two, you can, answer those questions about the effectiveness of free cash flow, alongside management. So you could actually see, does what we're finding dovetail with what they're saying on the quarterly conference calls, public releases, etcetera, and importantly, monitor this for effectiveness going forward. And one last thing, sort of testing this effectiveness, versus their international experience, because then it's really extrapolating it through a much larger store base.

Anu: Have you found anything interesting between U.S. versus international stores with McDonald's specifically?

Richard: Well, a lot of it has to do with the ongoing time frame for which the U.S. has been implementing. The U.S. was a little bit behind select regions. It's really a region-by-region analysis. I think the big takeaway, the concern upfront, was that it might work for selective regions around the globe, but if you could get into the North American base, you would really move the needle. So what we were pleasantly surprised to see that in fact the U.S. was amenable, for a whole series of reasons for this to be effective. And you could actually look at this, from an age standpoint, from a cohort standpoint. So not surprisingly, people the age 35 and below were early adopters and more interestingly was that for McDonald's this was really important because this giant touch screen kiosk would allow them to offset the cost of labor increases, reduce any of the mistakes of ordering when you go to a counter, but the surprise was the upsell. The surprise was, if let's say you like the cheeseburger with small fries and you normally ordered a coffee, maybe this time around you might want to buy, let's say a cappuccino, which has a higher price point and a higher gross margin contribution. The results were pretty fascinating.

Anu: Terrific. So now this is for both of you, how do your teams monitor the results after these initial findings? Is this something you'll continue to keep track of and look at going forward as well?

Michael: Well, so from our point of view, the nice thing about a computer is that once you've written the code, it's very easy to continue to actually see future reports on the same subject. And that's one of the benefits the data provide. I just want to quickly follow up on, what Richard said about

millennials, because one of the things that we do, lots of folks that are looking at credit card data, they're simply just looking at the dollar amount. But we infer the demographics by looking at credit card data because men and women transact differently, and so we can detect that. Millennials and baby boomers transact differently. Immigrants and people born here transact differently. And so we can actually infer demographics, and that actually also gives us a perspective on who the subset is and whether there's penetration into other groups.

Richard: And specifically to your question about how do you monitor this. If we can actually glean data that the company hasn't specifically stated, we can go back and say, look, this is what we're finding. Can you please confirm or deny this? And/or if they say something publicly, we have data sets to see is this really accurate. It's very powerful from a monitoring standpoint.

Anu: Yep, super interesting. Now, Michael, credit/debit card transaction data seem to be just one tool in your toolbox. What are some of the other types of data that you're out evaluating to generate insights?

Michael: Yeah, it's great question. So, there are many forms of data I think besides credit card data. Online activity, search terms, online purchases is a very useful data set. Also, job postings is a very useful data set. So just to give a couple of quick examples: So on the search terms, what you search for is what you're interested in and search has always been a leading indicator of intent to purchase. And in particular, if you search generically, let's say for aspirin, you don't care about the brand, or for batteries, you don't care about the brand. But if you search for the brand, then you do care about it. And after you search for the brand and if you go online and you see the same generic battery, for example, it's 40 percent cheaper, maybe you buy the generic battery. And that means the brand is eroding in its value. And so we're able to monitor that, and look at that in search term data. In job postings, job postings is immensely informative because people will signal in their job posting what they're interested in. So, for example, as companies move to the cloud, they signal in the job posting which cloud provider they're most interested in, so we can use that to monitor the relative growth in this competition of different providers of software and other things.

Anu: And with job posting specifically, do you think that you have a large enough data set of total job postings? What's the number that you're looking at?

Michael: Yeah, so, again, most of the jobs are posted online. There are lots of of job boards. And the nice thing about job posting also is no one complains about you having the data. It's a tear sheet. Please take my job posting. We have seventy percent of all online job postings going back to about 2013.

Anu: Ok, great. So now as we wrap up here, Rich, I'd love to get your take on this. How do you see this transforming information for the longer term? And over time, do you think there will be any kind of alpha decay, of these identified anomalies?

Richard: It's interesting you talk about alpha decay. I actually believe it might be alpha creating for those organizations, hopefully ours, that can do it the right way. And so why is that the case? It takes money, time, and expertise to actually crank through a big data analysis. You also have to start by asking the right questions and focus that analysis on things that lend itself specifically to that analysis. If you do all of that right, and there are firms out there doing it, a huge amount of firms are focused on the very, very short term with their big data analysis. We tend to be longer term investors. And so I believe that if implemented correctly, this could actually be an alpha-creating opportunity and an opportunity both in terms of the answer and the time frame that it's applied, one could use it to differentiate versus the competition.

Anu: Great. Thank you so much. Michael and Rich, really appreciate you joining me today to help investors think through how to truly get the most out of big data when it comes to making investment decisions. Again, thank you both for being on the show.

Michael: Great being here. Thanks very much.

Richard: Thank you.

Anu: If you'd like to learn more, I'd encourage you to read our recent publication, 'The Data Science Revolution,' available at www.nb.com/bigdata. Please subscribe to our podcast via Apple podcast or Google podcast and visit www.nb.com/disruptiveforces for more information about our firm and offerings. Thanks for tuning in.