



Bridge Too Far

Maintaining perspective on the promise and potential pitfalls of big data and AI
April 2018

Welcome to the first of what is intended as a quarterly letter setting out the QMA—and my personal—perspective on topical industry issues. We hope you find them informative, and perhaps provocative, and of course we welcome debate with clients on these and other issues.

For much of my time in this industry, as both consultant and investment manager, investment managers were naturally seen as aligned with their clients. If managers didn't perform well, they didn't have a business, and clients' levels of trust reflected that alignment. Personally, it's a great sadness to me, but we have to accept that this trust has declined and our industry has a lot of work to do to get back to that place. I want QMA to be at the forefront of that process. That's a big part of the reason I have decided to start writing to you directly on a quarterly basis. But you should also expect the thoughts shared here to reflect our firm's objective and data-driven nature, as opposed to the hype of which our industry is so fond. Offering balance where others don't will be a recurring theme.

In landing on my first subject, I thought about the many press interviews and other conversations I have had over the past months. One subject leapt out: the attention being showered on big data and artificial intelligence, or AI. Our industry has always been quick to recognize a rolling bandwagon and quicker still to jump aboard, and these phenomena have some of the same feel. We have new, important-sounding titles, such as "chief data scientist." Quant strategies have been re-packaged as AI. The interest is understandable. Fundamental investing has been under unprecedented pressure as the quant sphere has grown. Inevitably, many managers have latched onto big data and AI as a way to shelter under the quant tent. It makes them seem modern, differentiates them from slower adopters, and, most importantly, restores some faded luster to active management. Unfortunately, real understanding of these technologies in our industry is still scarce, and their applications to investing are anything but straightforward.

Further complicating the issue is a related theme we have already been vocal about in the press—the growing confusion around the very use of the term "quant." It's easy to understand why many may want to keep definitions blurry. After all, the more one can stretch the quant tent, the more people can try and take shelter under it. However, from our clients' perspective, using the term indiscriminately does them a disservice and does nothing to address the trust issues I talked about earlier. See our recent [Pensions & Investments Op-Ed](#) for our perspective on a more sophisticated quant taxonomy. No pride of authorship here—I'm happy for others to improve upon it, but collectively those of us who care about the future of our field need to find a more precise way of categorizing our approaches and philosophies that allows our clients to understand us better and tell us apart.

Big Data and AI Are Not the Saviors for Active

Now, let us look at the facts one by one.

Big Data Our view is that big data is definitely valuable as a new data source but is clearly not a panacea. Let us take a stylized (but frequently cited) example to illustrate the limits of the hype, the apocryphal satellite photos of the parking

lots of a given retailer. The hype narrative is, “Once upon a time we used to estimate quarterly sales off econometric modeling, but now we can use the parking lot photos to tell much more precisely what the sales will be.”

So, let's analyze this narrative critically. If you're a fundamental manager, you probably do need to embrace these sorts of real-time (or near-real-time) data sources because if others are using them and you are not, then clearly you are handicapping yourself. However, that is not the same as saying that simply by doing so you have found a competitive advantage unless (a) you own the satellite (rather than buy the data), and (b) your satellite offers even more precise or timely data than the sources your competitors have found. Since those conditions are generally not met, it's clear that use of car lot photos is a great boon for the owners and sellers of satellite data but is not going to fundamentally change the dynamics of the active management industry.

AI Again, there has been massive hype around the investment potential of AI. And on the surface this seems eminently plausible given all the reports of the exponential progress of AI in winning strategy games such as chess or Go. But, again, not so fast.

Let's start with the elephant in the room: data mining. Anyone who's been in the industry more than a year knows to be wary of buying products purely on backtested data. Since 100% of products with bad backtests don't get launched, there is enormous pressure to search for a data set that makes the tests work. But a good backtest is not the same as a good product. Finding sets of historical data where a correlation can be found with stock market performance is extraordinarily easy, but that doesn't mean the data will have the slightest predictive power going forward.

Around a quarter century ago, I wrote a spoof paper for my colleagues at Mercer Investment Consulting highlighting how the English Premier League performance of my hometown soccer club (Leeds United—this was back when they were a leading team) near perfectly predicted the performance of the UK stock market in the following year. That some readers took a few seconds to recognize the satire is a perfect example of how rampant the issue of spurious correlations was even then. Moreover, the math is such that the more data sets you look at, the more you can overfit a model and find series that seem to correlate. More and bigger data sets increase the risk of data mining rather than reduce it.

The problem, as it relates to AI and its nearly synonymous cousin, machine learning, is that these are data mining machines. To understand this better let's contrast the growth of AI and machine learning in chess with their record in another of my historical hobbies, bridge. Chess is a perfect information game in which all of the possible moves are knowable. As a result, there's no doubt that the rise of machine learning techniques, in which the program continually trains itself on the data until it achieves total mastery, has accelerated the ability of computers to play chess far faster and more effectively than the best human players. However, bridge is not a perfect information game, with a couple of hands never known to any given player. So, computers have struggled to make progress, and, as a result have not mastered humans.

Stock markets are similarly not perfect information settings, complicating the value of AI technology. For more on this, see the book *Everybody Lies*¹ by big data expert Seth Stephens-Davidowitz, which highlights amusingly how different and less impactful these techniques are in a market setting. AI is fundamentally about pattern recognition in this context, but the basic truth is that for the great majority of the patterns it finds in the market, there is no cause and effect and therefore no underlying predictive rationale. Indeed, the more the programs train themselves on big data sets, the

¹ Stephens-Davidowitz, Seth, *Everybody Lies*, HarperCollins, 2017.

greater the risk of finding relationships with no logical explanation or causation. Machine learning can accelerate the ability to find patterns but doesn't address the core challenge because the information setting is not perfect or perfectly masterable. For posterity, I hereby dub this Leeds United Syndrome!

Takeaways

So, big data and AI are not the answer to all of active management's problems. They are tools that can be helpful in the investment process when they are used properly. We think the key is having a clear investment philosophy that guides what data sources a manager is using, what new signals are tested, how they are tested and how they fit in the manager's models to generate meaningful, sustained net outperformance. In our stylized example, while the satellite parking lot photos in and of themselves are not differentiators, *how* they are implemented, weighted, timed, etc., may be. Similarly, with rigorous and careful enough mapping, verification and out-of-sample testing, it's possible to use AI to uncover more complex relationships in the data that *may* have predictive value that leads to new sources of alpha. However, complexity simply for the sake of identifying *ex post* patterns is often just a more sophisticated version of data mining. With such analysis, it is probably even more important that there be a clear connection between the signal and stock returns—a trail of bread crumbs that can be followed back to fundamentals and a manager's deep understanding of investor behavior and the markets.

In the meantime, I'll conclude by noting that as the season comes to an end, Leeds United sits characteristically mired in the middle of the pack of the second tier of the English football league tables. I encourage you not to read too much into it!



Andrew

NOTES TO DISCLOSURE

The comments, opinions and estimates contained herein are based on and/or derived from publicly available information from sources that QMA believes to be reliable. We do not guarantee the accuracy of such sources of information and have no obligation to provide updates or changes to these materials. This material is for informational purposes and sets forth our views as of the date of this presentation. The underlying assumptions and our views are subject to change.

This is intended for Professional Investors only. Investing involves risk and the value of investments can fall as well as rise. Past performance is not a guarantee or reliable indicator of future results.