‘We think of ourselves as cyborgs’: Why the future of investing may be half-human, half-machine

June 19, 2019

By **Devin Banerjee, CFA**

Welcome to Human Capital, an open exploration of the ideas and people moving financial services forward. Every two weeks, we feature a leader or rising star who’s changing the game in his or her own way. “Finance is an apprentice business,” one often hears in this sector. Here are some of the teachers. Click Subscribe above to be notified of future editions.

Data has long been — arguably *always* been — the lifeblood of investing in public markets. Honing the practice of harnessing, interpreting, and acting on data has been the focus of some of the most successful hedge funds, such as Renaissance Technologies, D. E. Shaw, and Citadel.

But the application of data science and quantitative methods has been far less prevalent in the private markets, where it’s perceived that information liquidity is harder to come by.

The trio featured in this edition of *Human Capital* — Tingting Zhang of TerraCotta Group, and Ian Picache and Sajjad Jaffer of Two Six Capital — are out not only to bust that myth, but to show that data is the most potent source of “edge” in the future of investing.

Below are excerpts from my conversations with Zhang, Picache, and Jaffer.

Tingting Zhang, founder and president, TerraCotta Group

Your path to managing a real estate credit investment firm was quite unique. What was it?

It’s an illogical path. I did my PhD in international political economy in the mid-’90s, which was a time when the discipline was going through some changes. We started to look at using real-world data versus data that we modeled based on the older econometric models. So, it was a time when econometricians started to explore whether or not the discipline had any application toward solving real-world issues.

And my dissertation throughout my five years of grad school was really focused on that. I was interested in how the Chinese central government controlled provincial leaders through this whole

analysis of whether provincial leaders were promoted based on loyalty or performance. And it turned out to be a gigantic time-series regression and produced meaningful findings. That was the initial period that really made me interested in doing some complex work and being able to quantify variables that are seemingly not quantifiable.

Following grad school, I went to work for a think tank, and they pioneered in using quantitative methods in social analysis, political analysis, economic analysis. There, we had clients who were interested in looking at how democratic different countries were — variables that, at the time, based on the common perception, would not be part of the domain of the quantitative method.

When I was there, that's when I came across the commercial real estate credit investment area. When I was looking at how the property valuation and cash flow was assessed, it became very clear to me it was just a very simple analysis based on summary statistics. We would look at five comps in the nearby area in the last six to 12 months and, based on the observation of which comp is more comparable to the property, come to a conclusion. There was no underlying predictive analytics, such as what drove certain changes.

Without the underlying predictive factors, the whole investment area is unable to predict change. As a result, commercial real estate credit investments have been proven to do well during good times but not so well during bad times, because there is a gap between today's analysis and looking to the future.

I thought it was a classic problem set for quantitative methods, and that's when I started the TerraCotta Group in 2004, really with the eye toward building a systematic framework to understand what the drivers of commercial real estate value and cash flow are so that we can predict property performance if these values were to change.

Why do we see such a strong application of quantitative methodology in the public markets, but not so much in the private markets?

I'm a bit baffled that we still don't have attention toward the use of data in quantitative methods in this area. On the other hand, I do understand that this is complex, unlike the public market, where you have a more regular, more frequent occurrence of data.

In the private market, data does have its limitations. For instance, if a property is sold, the sales price is actually public information recorded on the grant deed because the buyer and the seller have to pay transfer taxes. So, that information is fairly reliable and that's similar to public market data.

But a lease transaction is actually a private market transaction — how much is transacted is private knowledge and is never published. The information collection really entails the data collectors calling brokers, trying to find out what the information was, and brokers oftentimes citing proprietary control of this very valuable resource that they use. So, the information, if not completely inaccurate, is hard to verify.

There also is an array of issues that make the commercial real estate credit market much more challenging than public markets. For example, how do you quantify location? How do you quantify a city's pro- or anti-development bias? These are critical factors that real estate investors consider. Oftentimes, if a city is pro-development, the real estate value actually doesn't hold as well as in anti-development cities, as the supply constraint is not present.

So, there is an array of complexity in the area of commercial real estate credit. It not just involves figuring out what the variables are, but how to measure them and how to make sure the data is clean before you use them. *Then* you consider the algorithms, which are very granular. That's probably one of the reasons why investors have been hesitant to adopt quantitative methods in this area.

Are there challenges to accessing this data?

Data itself is a commodity. It wasn't so much so 15 years ago. Certainly for any type of data, you have multiple accesses. And we have a lot of it — trillions of points of data at our fingertips.

So, at each of the millions of the commercial properties we target, we know when the space became vacant, when it was leased, to whom, at what pricing, according to multiple sources. And we know what the concessions are, and we know when the property is listed for sale, to whom, for how much, what the pricing deduction is, how the buyer financed the transaction. And we also have access to local demographics data, economic data, traffic counts, all of which yield business profitability within tenants.

The data sets are expensive, but the investment outcome it produces pales in comparison.

I think the challenge with data is, one, whether or not you have a system to clean the data such that it can be used. And, two, the quantification of variables. We all know location is important, and even today the typical real estate operator's or investor's perspective is that location cannot be quantified or it's more a gut — people will say, oh, west of the 405, that's close to the beach. All of these factors are absolutely quantifiable, and we use 60 variables to capture them and weigh the underlying variables to give a single location score for each property.

What are a few of the largest misconceptions about this process?

One misconception is that if you have some error in the data, then you can't use the data to produce accurate predictions. And I think the other one is probably more prevalent, which is that models are very stiff and don't provide the creativity needed to invest in real estate credit. It's actually to the contrary.

In the areas of not just commercial real estate credit investment, but for private markets in general, your algorithms will have to go through significant testing, with in-depth knowledge from an investor as to what works in what domain.

In this area of credit, we don't have "one model fits all." We understand there are differences across property types, there are differences across the market, there are differences across corners. We are really trying to understand: What are the joining forces of real estate values? So, those are the 200 variables that we have established over the last 15 years.

And once we have these variables, it's very dynamic. For instance, if we are looking at a shopping center in Phoenix, if the anchor tenant, Safeway, were to depart, the query would be: What is the quantitative impact on the center? So far in our field, we really don't have a consensus regarding the marginal impact of an anchor tenant departure. And many lenders or real estate investors stay away from lending against properties where the anchor departs because it's scary; does that mean a $1 million impact or is the whole center going down the drain?

With quantitative methods, you can include a predictor variable called 'presence of anchor.' If we were to run a regression with 200 observations, some of them would have the anchor at a center, some wouldn't. If we run the lease rates with a dummy variable called 'anchor present,' you can turn that on and off. You run a regression with centers with the anchor tenant, and you run a separate regression without the anchor. The delta will show you, if a center's anchor vacated, how much the adjacent non-anchor tenants would be willing to pay.

So, it's a tool, certainly. We are not completely replacing humans' judgment here, but it's a sharp and a dynamic tool.

Do you think the adoption of data science is progressing at such a rate that there will be fewer human investment professionals in the future?

I certainly do think it's going to impact the human-machine ratio in the investment area.

In TerraCotta's case, we can finish underwriting a very complex credit in two hours. It's a well-established technology framework. Many of the investment processes are semi-automated, and we are working on additional automation. I think we're a fairly lean shop given the loans we do, and we will probably stay a very lean shop as we scale.

Should an investment firm have data scientists *and* investment professionals, or should investment professionals *be* data scientists?

That is a really good question. We pondered that ourselves for a good part of 10 years. The initial thinking was that we would hire people who are real estate investors or credit experts and train them to use data, and that hasn't worked.

The other side of the experiment is to hire people who are quantitatively minded, primarily kids fresh out of college who are smart and driven. That doesn't work in every situation because it's actually two very different skillsets, two very different personalities oftentimes. But at least 25 percent of these kids turned out to be great investors.

I think that's another challenge for the private market's application of data and quantitative methods. These two professionals don't really mesh very easily, and that has been one of our biggest challenges as well.

Based on our experience, we don't believe that you can produce a meaningful, dynamic, and creative methodology if you separate research from the investment. At TerraCotta, research and investment are fully integrated.

If an analyst has a difficult file, or the outcome that the methodology produced from a real-world test is peculiar, they would raise the issue of, how do we improve the methodology, how do we improve the algorithm, can we find a different way of quantifying the variable or adding a different variable. They are the one who raises questions first, and once we go through the collective thinking process and find the answers, they buy in. And they apply this new finding, this new addition, to the transaction at hand. Not just transactions that we're looking at, but also transactions that happen in the real estate world. There are about 100,000 to 300,000 transactions every quarter in the market that we monitor, and every time there's a transaction, we want to test our methodology.

So, I really believe that the fast iteration, the buy-ins from the investment team who are sophisticated data analysts, that makes this whole process work.

Talk more about the challenge you mentioned in 'meshing' the two types of professionals. Do you mean that culturally, it's difficult for investment professionals and people with data backgrounds to work together? Or do you mean it's hard to mesh those skills into one person?

I think both. We saw during a certain period of TerraCotta's history the real estate guys thinking, I don't think data is going to help me, and the data people desperate for feedback from the real estate guys.

And we observe the college graduates wanting to work here. When we bring on an analyst, it's one out of 100 to 200 kids. They're all smart, mathematically inclined, and the vast majority of them are introverted. And the skill that makes a data scientist click is oftentimes that they have the patience, they have the curiosity, to go through a large data sample and not rest until they get a solution. The real estate side usually is a bit impatient, a little bit restless, a little bit of just wanting to make moves; they're action-oriented. So, the traits don't really mesh into one person.

They also have very different inclinations; for example, when they get a file, do they want to get back to the sponsor immediately, or do they want to figure everything out and put a bow on it before they call back.

So, there are a lot of very subtle, intangible factors that make a quantitative person and a successful real estate investor a really rare find in the real world.

I think that's probably the biggest challenge in TerraCotta's journey, and I think it would be the biggest challenge for any firm in the private market taking up quantitative methods.

Ian Picache and Sajjad Jaffer, founders and managing partners, Two Six Capital

Two Six Capital applies data to the private equity diligence and investment process. Why did you start the firm?

Jaffer: One, we recognized that there's been no technological innovation in private equity since the invention of the Excel

spreadsheet. And two, focusing on revenue as the basis for driving returns was a fresh approach to alpha generation. The traditional approach in private equity was around cutting costs and levering up the assets, and our view was that there was an opportunity to focus on the top line. If you can understand the top line really well, then there are a lot of interesting things you could apply in diligence and value creation.

Picache: There hasn't been innovation in private equity largely because they haven't really needed to. The returns have been really good, and a lot of the firms haven't needed to innovate. I think now, with increased competition, innovation is going to need to occur.

How do you obtain data?

Picache: I would say you actually get more information from the private markets than the public markets.

We get data from the companies we are analyzing.

To that, we then append all of the other stuff that public equity investors are looking at. So, there's no hindrance in taking all of that public data and making much more educated private equity investments.

The good news about the data that you get from the companies themselves is that it's a complete set. A lot of the data that you see in the public world are samples and not complete.

Most of what we do is in diligence and value creation. So, it's taking data from a company that's for sale, understanding what's going on in the business, and looking at what the levers are for revenue in this business. Are the levers for revenue all going in the right direction? Are some going in different directions? How can we optimize revenue? Once you understand revenue and can project that out going forward, the rest of the stuff is relatively mechanical: You know how to hire against that revenue; you know how to build your cost structure against that revenue. That's why revenue is a big thing to understand.

Take us through an example.

Jaffer: We've worked with large retail companies on corporate transformations where they're trying to understand: How do I think about who my best customers are, my worst customers, what services I should launch, how I maximize the share of wallet?

In this particular example, when we took the entire transaction database of every customer that had ever transacted with this business, we found out that there was a large 'one and done' problem — very few repeat customers. The management team hadn't paid attention to that or weren't aware that issue was real.

Once we quantified the problem, we could turn around the sales and marketing plan, what merchandise to sell, what new services to add, which stores to expand, which stores to consolidate, how to think about the lift from various elasticities of pricing.

We built technology dashboards, and these dashboards give you an almost real-time view of how the different levers are moving in the business. If you're training for a marathon, you might look at your body weight, your BMI, and all the different performances. You will be watching your Apple Watch pretty closely on how that is tracking over time. So, we have built the equivalent of that kind of watch discipline so that you can now watch the different KPIs and then drive different changes to the business.

Where a lot of the industry has focused in the past is around, let me just go put a bunch of the right people in, the right ex-management consultants. Or, I can put a little bit of process in — I'll draw a playbook and I'll require management to follow these rules as part of the post-investment plan. We've taken the best of those two and we've added technology, and using technology now gives you an undisputed view on how the metrics are trending over time. So, when you go to the board meeting, there's no ambiguity between how the board sees the company and how the management has presented it.

What if it looks like a great investment but the quality of data from the company isn't high? You will pass?

Jaffer: Yes, we will pass.

On diligence, when we started requesting the transaction data, almost always people would scratch their heads and say, no one has ever asked for it, why would you even want it? And so we had to educate the market on, well, actually if you get this data, these are all the key hypotheses that you can answer.

We've now analyzed $134 billion worth of transaction-level data, so these are receipts that businesses generate. It's a lot of information, and we think that the data is just going to continue to increase.

If you don't get high-quality data, then obviously you've got one of two things going: One, you are going to bail on the deal, or two, the management team has got some questions to explain themselves.

When else do you pass on a deal based on data?

Picache: We've killed a lot of deals because we're saying, OK, the data is telling us one thing, but the price for that business is too high relative to the amount of work we need to fix it.

Jaffer: Also, when management doesn't acknowledge or accept what the data is saying. During diligence, we're observing the resistance or the openness to the data-driven transformation.

We'll kiss a few frogs and have to move on. And then we'll find a few who will embrace it, and they're the ones you want to work with.

I noticed on your website that you call everyone here data scientists and data engineers, not investment professionals.

Jaffer: When we stood up the firm, we realized that there was no one in the industry that had ever tried this approach, so there was no point in trying to hire people from the industry to build this out.

We wrote our own talent management playbook, we developed roles and responsibilities, we have our own recruiting and filtering process. We borrowed a good number of pages from Silicon Valley tech companies. And the market has responded very well to it, especially the millennial generation, which we think are the future of this industry.

Last year, we had over 700 applicants for a couple of spots. So, we are fortunate that the talent is available, provided you have a good mission statement that can appeal to the millennial generation.

What resonates from Two Six Capital with data-oriented young people who might be considering working in the tech sector instead?

Jaffer: When we describe the ideal Two Six person, we describe somebody who is highly amphibious and bicultural.

When I'm recruiting, I'll wear, like, a tie and a hoodie and a pair of sneakers.

The point is your front game will be very similar to your traditional private equity investment banker, but your back game is very Silicon Valley, very tech-driven. And that actually tickles a lot of people's curiosity, to say, wait a minute, that's me, no one had created a job that would play to both sides of my personality until now.

What's the future of Two Six Capital?

Picache: The future of the firm is to really continue to ride the technology wave. We are able to prosecute many more deals in parallel than we were before. When we first got going, it took everyone here to do one deal, and it took us months longer. We have since gone through and productized a lot of the stuff, so we can run probably six deals in parallel now with a fewer number of people on any given deal.

Jaffer: We have quantified processes that were FTE-intensive, and we can now do that same process with 80 percent less head count. Which is why the team size has remained pretty flat even though the volume of deals we've worked on is multiples of what we used to handle. So, we want to drive that scalability by investing in technology.

I think the future belongs to folks who can really bring the best of what humans can do and couple them with machines. We think of ourselves as cyborgs.

Humans are great at asking the right questions, and our firm is very hypotheses-driven; we will never outsource that to a machine. Similarly, humans are great at drawing inferences and using deductive reasoning to say, hey, this is a pattern I saw in another industry when I look at company X. Machines can't do that. But then, once you get 5 billion rows of data and you want to crunch it in seconds and generate 10,000 iterations, well, that's a machine's job. So, you have to really reimagine the whole private equity process.