Risk.net March 2024

The quants who kicked the hornets' nest – to champion causality

The quants who kicked the hornets' nest – to champion causality

A small but influential cadre says the multi-trillion-dollar factor investing industry is based on flawed science. By *Rob Mannix*

Since 2013, factor investing performance has stagnated. And Marcos Lopez de Prado thinks he knows why. In January, he released a 35-page working paper that practically declares the sector's systematic investing approach to be built on bad science.

"Factor investing has failed to perform as expected … because the econometric canon used to make and peer-review factor claims is flawed," states Lopez de Prado – global head, quantitative research and development at the Abu Dhabi Investment Authority (ADIA) and a founding board member of its independent research spin off, ADIA Lab – writing with his colleague Vincent Zoonekynd.

Need to know

- Quants from ADIA Lab say the factor investing industry is built on models that are wrongly blind to causal effects.
- A "substantial portion" of the trillions in factor strategies could be invested in flawed strategies, they argue.
- They are calling for the wider use of techniques from causal inference as part of the model-building process.
- "The causal revolution has arrived," say the quants.
- Others doubt their ideas will catch on, however, saying that to establish causality is a gold standard that may be too hard to attain.

What's missing from the way quants and economists work, they say, is an appreciation of causality – or what's causing what – and why.

Factor investing is grounded in conventional statistics – and conventional statistics look for patterns in data regardless of causality, say the two quants. The approach leads to models that are liable to breakage, they argue. Or, at worst, are fundamentally wrong.

Investors should instead be using the computational tools of the fast-growing field of causal inference, says Lopez de Prado, who has been setting out his case in this and other papers – and in a short book – since January last year.

"Our findings challenge the scientific soundness and long-term profitability of the current multi-trillion-dollar factor investing industry," the authors told me in February, when I corresponded with them in detail about their paper.

"The tenets of the factor investing literature need to be revisited. At this point we must assume that the majority of factor models are misspecified, hence investors are potentially exposed to systematic losses."

The issue is potentially huge. And hugely controversial.

Take systematic value investing, a long-established and enormously popular strategy with investors. The biggest systematic large cap value ETF [exchange-traded fund] – Vanguard's VTV – has more than $150 billion in assets.

Value strategies like this one are based on the observation that buying stocks with low price-to-book ratios has delivered a premium over time. But if something other than the relative cheapness of the stocks led to that performance, and the unknown variable changed, the strategy might wane.

And value investing has performed poorly in recent years, losing more than 2% annually from 2010 to 2019, on average. The factor is down more than 11.5% over the past 12 months (see figure 1).

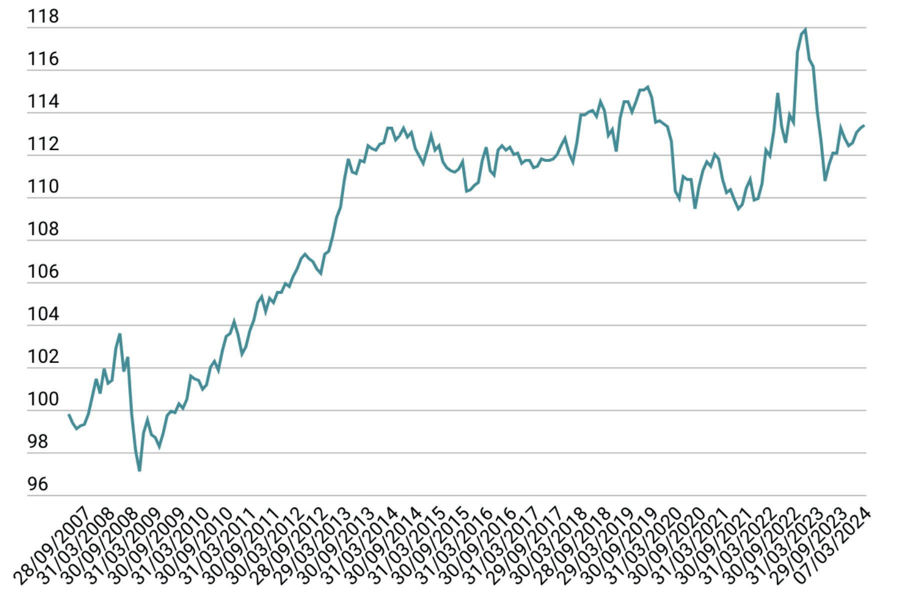
It's an argument that's likely to get heard. Lopez de Prado holds two PhDs, founded and led Guggenheim Partners' quant investing business, then headed machine learning at AQR, before joining ADIA. He is the top-rated author in economics on SSRN, the online repository of preprint academic papers. He is a former *Risk.net* buy-side quant of the year and *Journal of Portfolio Management* quant of the year.

Zoonekynd, a former director at Deutsche Bank, also worked as a quant at Goldman Sachs for four years and joined ADIA as quantitative research and development lead in 2021.

It's "conceivable" that a "substantial portion" of the more than $3 trillion invested in factor strategies has been allocated to structurally flawed strategies, Lopez de Prado and Zoonekynd tell me.

"It may take a few years. But eventually the industry will evolve to no longer be content with correlations of unknown origin. The causal revolution has arrived to investing."

1 Factors have stagnated



Source: Bloomberg

Nevertheless, many doubt the idea will catch on. Experts, including Kenneth French, who with Eugene Fama created the Fama-French factor models used throughout finance, say efforts to establish causality may be a distraction.

Speaking to me for this article, French, a professor of finance at Dartmouth College in the US, says asset pricing models aim only to identify differences in expected returns and not why the differences exist. “The regression just captures this linear relationship. That’s all.”

Lopez de Prado is nonetheless happy to court controversy. One quant describes his recent papers as “kicking the hornets’ nest”. He keeps kicking. ADIA Lab is now offering a $100,000 prize for the best three research papers about causality in investing.

The causal revolution will not be televised

Right now, Lopez de Prado is the one putting causality in the spotlight. But the debate about its importance goes back years.

One early advocate was Riccardo Rebonato, professor of finance at Edhec Business School and previously a senior executive at Pimco and RBS. In February, I spoke to Rebonato by Zoom from his study at home in London. He explained that his interest in causality started when he was carrying out scenario analysis at Pimco, where he was the firm’s head of rates and FX analytics from 2011 to 2016.

As we talk, he takes down books from his shelves – including Lopez de Prado’s, which Rebonato edited – and several of his own publications that make the case for a causal understanding of markets.

Some of the titles go back more than a decade. It seems the progress of the causalists has been slow – at least until now.

Rebonato gives an example from risk management of why causality matters. When central banks hike interest rates it’s important why they hike interest rates, he says; risk managers can say little that is meaningful about the consequences without knowing the cause.

Much of this may seem obvious. Students of statistics know that correlation is not causation. Students of causal inference, though, want to address this problem head on.

“I can find the correlation between the number of ice creams sold and the number of people drowning,” Rebonato explains. “But … there’s a hidden variable – the hot weather – that causes people both to eat more ice creams and to go swimming.”

The mathematics of causal inference, most notably pioneered and championed by UCLA professor Judea Pearl, centres on the construction of DAGs – directed acyclic graphs – which are essentially maps of the interactions between variables.

Rebonato chuckles at the quirky language of the graphs and the growing number of causal discovery computer algorithms used to work them out: “parent” variables, “child” variables, “moralities” and “immoralities”, “roots” and “leaves”.

Factor models that ignore such graphs, can be flawed in two ways, the causal experts say: by leaving out of the model confounders that influence the variables that go in – the hot weather in Rebonato’s example – or by mishandling so-called colliders.

“A collider – two causes converging on an effect – implies backward-moving causality, which can be counterintuitive, but is real,” says Rebonato. The name reflects the fact that causal arrows in a graph collide at the variable in question. He gives an example.

“Suppose my car does not start in the morning – most likely because the tank is empty, or the battery is flat. If I check there is petrol in the car, the probability of the battery being flat goes up.”

Colliders in a model are not a flaw, in the way that leaving out confounders is, he explains. “They just make life tricky.”

For their fans, causal graphs have to be the starting point for originating new models; quants should identify confounders and colliders in their data in order to control for the first and avoid the trap of controlling for the second.

“Consider the canonical Fama-French three-factor model,” say Lopez de Prado and Zoonekynd, when I ask them about how this might apply in finance. It explains the average excess returns of different stocks in terms of the market’s overall excess returns and the exposures of the stocks to the value and size factors. “The model treats these explanatory variables as if they were all independent causes of excess stock returns. However, empirical evidence seems to indicate that reality is much more complex.”

Another factor, such as momentum, might influence the value factor and market returns, both of which might influence the size factor, they suggest.

“In that case, momentum is a confounder and size is a collider. Failing to control for momentum can severely bias the estimated risk premium for value. And controlling for the size factor will mislead researchers into believing that gaining exposure to small caps is a rewarded risk.”

This misspecification scenario is hypothetical, say the two quants, though it is consistent with their results. Their point is that Fama and French should have explained carefully why this scenario is unlikely, they add.

The factor zoo

It was through Lopez de Prado's work at the Lawrence Berkeley National Laboratory, where he has held a position as a research fellow in computational science since 2011 that he formed the view that finance might take a more scientific approach.

That's far from what's happening, he says – instead, with its associational practices and by selecting models for their explanatory power, the industry is in fact amplifying the risk of choosing models that are wrong.

What's more, this is occurring on top of an already known problem. "Until now, authors have explained the proliferation of factors known as 'the factor zoo' as the result of brute-force p-hacking via multiple testing," say the co-authors in their paper.

They are referring to the work of Duke University professor Campbell Harvey, who showed that testing dozens of possible investing strategies would inevitably lead to spurious ones meeting the p-value thresholds that quants apply. P-values measure the likelihood that a statical pattern is a fluke.

Harvey's work triggered much scepticism about the 400-plus factors identified in academic journals, many of which form the foundation of commercial investing strategies. Lopez de Prado and Zoonekynd are saying that causal ignorance, for want of a better term, exacerbates the p-hacking problem.

Simply put, non-causal models are likely to include specification errors that make them appear more predictive. And such an appearance could be enough to lift a misspecified model above the test thresholds that quants apply.

The way those taking a causal view see things, these practices will lead to models that generalise less well, meaning they break down quickly when superficial conditions change.

"Our results show that the econometric canon enables p-hacking within a few trials, by favouring over-controlled, underperforming – including money-losing – models with a higher R-squared and lower p-values than the correctly specified money-making model," say the two quants. R-squared is another measure of a model's explanatory power.

Over time the effect will undermine – arguably, has undermined – the spectrum of strategies on offer, they add. "It is natural for over-controlled models to crowd out correctly specified models, making it more likely for underperforming and money-losing models to be selected."

> "At this point we must assume that the majority of factor models are misspecified, hence investors are potentially exposed to systematic losses"
> Marcos Lopez de Prado, ADIA

The authors stress that they have published their paper under their affiliation with ADIA Lab – a research unit set up and funded by ADIA – and that the paper does not necessarily represent the sovereign wealth fund's institutional view.

Past imperfect

To the causal experts, the dangers of relying on associational statistics seem obvious. "If the past doesn't repeat itself, you're completely screwed," says Darko Matovski, a former quant at Man Group who today runs causaLens. "You haven't really learned how the world works." His firm provides a platform that accelerates the build of causal machine learning models to hedge funds, banks and asset managers.

Matovski draws a comparison to Newton's inverse-square law, which supplanted with a simple equation more complex pre-Copernican models of planetary motion and generalised better to other problems including apples falling to the ground.

Causal discovery algorithms offer a fix, the experts contend. The algorithms can derive from raw data the causal graphs the quants require.

Broadly, the algorithms work by assuming everything potentially causes everything, then stripping away links between variables that can be seen to be statistically independent; and then examining how different pairs or combinations of variables change in the data, depending on each other.

"In the past, researchers had to guess their model specifications," say Lopez de Prado and Zoonekynd. "Today, we can do better than guessing."

A few researchers have made initial forays into testing the idea. Quants from Bloomberg and the Stevens Institute of Technology applied a causal discovery algorithm to Apple's returns and found, for example, that the value factor seemed to have no causal influence on returns during the period studied and that causal factor relationships changed through time.

Late last year, students at Cornell University working on a project that Lopez de Prado helped co-supervise, which applied causal discovery algorithms to the Fama-French five-factor model. They concluded that a definitively causal version of the model would be stripped to just two factors, relating to value and company investment practices.

Just cause

To some, though, focussing so much on causalitymay at times be unnecessary or unproductive.

To French's mind, questions of causality should come after identifying the patterns at play. "We're not trying to attribute causation, at least not through the model," he says.

"Trying to figure out what's causing the investor behaviour that is causing these differences in expected returns – that is the \$64 million question. But the asset pricing model has nothing to do with it. It's just trying to say, what are the patterns?"

"When we try to interpret these models, causality is interesting. But causality doesn't inform me about whether the model is right or wrong."

Rob Arnott, sometimes described as the godfather of smart beta investing, and founder and chairman of Research Affiliates, has written plenty on the shortcomings of factor investing research. In a 2016 paper entitled *How can smart beta go horribly wrong?* Arnott points to what he calls the "alpha mirage" of positive performance, derived from rising factor valuations, and warns about a possible smart beta crash.

He is also pragmatic about how a hypothesis might be formulated. "The approach advocated by Marcos is more rigorous and less vulnerable to false positives than what you might call hand-waving theory creation. But hand-waving theory creation is not without its merits," Arnott told me by video call from his home in California.

"It's fine to have some hypotheses that don't come out of statistical metrics but come out of simple human thought." He highlights the quality factor as one that was identified from the data, with the rationale identified after the fact, because there ought to be a risk premium for lower, not higher, quality companies.

Campbell Harvey also gives the impression he would be comfortable exercising judgment on factors without the algorithmic pursuit of causal discovery.

"To me, a factor is a source of risk and there should be a reward for it," he tells me by email. "If there is an asset class that is illiquid, prices are naturally depressed so that expected returns are higher. It is hard to sell an illiquid asset in a downturn and investors know that. So the risk premium – expected return – is high."

"There are many other factors that seem to me to be trading strategies. They provide an impressive premium in the backtest," he adds. "However, there is no obvious reason that premium will persist in real time and in the future. Investors will crowd the trade and the premium will vanish. This is not a factor."

"A collider – two causes converging on an effect – implies backward-moving causality, which can be counterintuitive, but is real"

Riccardo Rebonato, Edhec Business School

Riccardo Rebonato: "I can find the correlation between the number of ice creams sold and the number of people drowning. But … there's a hidden variable – the hot weather – that causes people both to eat more ice creams and to go swimming."

Sanity clause

Whether a causal approach becomes more widely adopted, then, remains to be seen. Practitioners who have explored the area report mixed experiences so far.

A team at Robeco used a causal discovery algorithm to formulate alternative groupings of stocks versus MSCI's Global Industry Classification Standard.

The researchers were able to identify changes in GICS ahead of time, such as the 2016 separation of real estate from financials, says Clint Howard, a quant researcher with the firm. The algorithm picked up in the data that real estate stocks were starting to break their ties with broader financials as early as 2013. Such insights could help quants in neutralising sector exposures in their strategies, Howard reckons.

Yet the findings of Howard and his colleagues were "reasonably aligned" with what they would expect based on past correlation-based work, he says.

Proving causality may not be worth the effort, he suggests. "The process of applying causal discovery algorithms can bring additional opacity compared to simpler, associational relationships."

Rebonato says causal factor investing is a gold standard. To work up a scenario with a step-by-step walk-through of causation would take a team an afternoon, he told me. Drawing causal graphs for factor investing strategies is similarly taxing.

"Marcos has said, unless you give him a causal mechanism – it could be behavioural, it could be due to finance theory, it could be due to anything – he will not accept an association. An association is prone to 'ice cream syndrome'."

But it's a high bar, and commercial forces are pushing the other way.

The industry needs to generate ideas, says Rebonato. "A salesperson needs to come up with a trade idea every month. If I build a rigorous process, I'd be happy to find a causal relationship every three to five years. This is excellent discipline. But where is my trade idea of the month? Everybody says: 'Yes, yes, yes'. But when it comes to doing it, I fear it will not be so widely used."

At the same time, causal discovery brings complexity and is computationally expensive. At Robeco, the research project was limited to computing causal graphs for just 500 stocks, based only on returns.

The discovery process is also imperfect. The computational complexity of the algorithms may lead to some "inherent noisiness", says Howard. This arises from the optimisation process and parameter estimation underlying score-based causal discovery algorithms, he says.

The value of causal inference may, then, be in providing a "sanity check" for research rather than replacing existing methods entirely. "It's just a different way of seeing things. We can use it as part of the research process," says Howard. "But I don't see it as an either/or."

"It's fine to have some hypotheses that don't come out of statistical metrics but come out of simple human thought"

Rob Arnott, Research Affiliates

"We know causality is the way to go, but we do not know yet clearly how to get there"

Alexander Denev, Turnleaf Analytics

Once seen

Even early advocates of causal thinking acknowledge difficulties in its application.

Alexander Denev is a former head of quantitative research at IHS Markit and a lecturer in mathematical finance at the University of Oxford. He is the co-founder of Turnleaf Analytics, a macro and inflation forecaster that uses machine learning and alternative data. He also studied at Oxford under Rebonato, who was his thesis supervisor. The two wrote a book together on causality and asset allocation in the 2010s.

Finding causal laws is notoriously tough, says Denev. He cites the Greek philosopher Democritus: "'I would rather discover one true cause than gain the kingdom of Persia.'"

Causal models have proven successful in stable environments, such as genetics and language translation. "But language doesn't change every month." Twenty years ago, oil price rises might prompt inflation in the US. Now the US is a net exporter.

Denev also talks of the labour involved in compiling causal views. In 2014, he put together for Pension Insurance Corporation a causal analysis of what might happen after the UK's Scottish referendum – which assets would perform in the case of a Scottish exit and so on. The analysis ran to 50 pages, presenting the final output as a causal graph. It took two months to write the narrative and build the causal model.

At the same time, constructing the models risks missing causes they cannot directly observe, Denev points out. In Turnleaf's inflation model, the firm uses Nasa pollution data as a proxy for industrial production. "Does pollution cause inflation?" he asks. "No. But it's a good proxy for a latent variable that is causal for inflation that we can't observe in real time."

"I cannot give you a definitive answer as to whether causality will make a difference," concludes Denev. "Running a purely associational model when it comes to prediction – you can argue it doesn't really matter." His own firm uses an inflation prediction model with 6,000 indicators that is purely associational but still does a good job, he says.

Denev nonetheless remains positive about the future. "We know causality is the way to go, but we do not know yet clearly how to get there," he says. "It's still research in progress. It hasn't yet been proven to work in practical settings. But I think we're already starting to see some interesting developments in this area."

And pragmatism seems key to the progress of the causal view. At causaLens, Matovski says his aim is not to replace existing models with something perfect. "The idea is: can we do better? Can we eliminate spurious correlations? Can we find colliders?

"If you're trying to discover how the market works, you will probably fail. What we're trying to do with causal models is to do better than traditional models. And in the world of alpha discovery, small improvements make a big difference to the P&L."

Lopez de Prado and Zoonekynd are also realistic about how fast they might advance their side of the debate. "There is still a lot of inertia and resistance among some seasoned academics to any form of change – particularly when it challenges tenets held as sacred for over 60 years," they tell me. But then, they add: "Once you see causality, you cannot unsee it." ■

Editing by Louise Marshall