



Knowing where best to invest marketing spend requires context, precedent and common sense. Marketing mix modelling can be a vital tool for determining return on investment, but the models are not without their limitations. By **Tim Phillips**

The results of using this type of data – even if the models are far from perfect – often clearly outperform any other approach

Marketing mix modelling (MMM), in which historical data is fed into a model to try to optimise the return on marketing investment, has been a tool for marketers since the early 1990s. In the 2010s, it has created a flourishing sub-discipline: predicting the demise of marketing mix modelling. "To be provocative, I believe it will become obsolete," said Laura Desmond, then CEO of Starcom MediaVest Group, in 2011.

"If MMM is to survive, it is essential that it change and experience a rebirth," the editors of *GreenBook* claimed in September 2016. For this article, Michael Wolfe – one of the pioneers of the technique in the 1980s – told *Impact* that he believes the discipline is "in crisis" (see p30, A crisis in MMM?).

Like any statistical tool, MMM can't please everyone all of the time. There are marketers who can't learn much of any interest from the available data – for example, if there is a particularly long and complex path to purchase. For others, their market is simply too volatile, or too focused on one aspect of the marketing mix. For digital platforms, competition is increasingly for the market – rather than in the market – so this makes a sophisticated mix model less relevant.

Steven Levy's book *The Plex*, which tells Google's early story, reports that when Scott Epstein joined the company as vice-president in 1999, he proposed a marketing plan based on optimising the mix of the 'four Ps' – product, place, price and promotion. It was rejected by the founders, and Epstein left the company. A Google

insider dismissed the relevance of marketing mix for a start-up that had market dominance as its goal: "Do we want to put money into the technology, into the infrastructure, into hiring really great people? Or do we want to blow it on a marketing campaign we can't measure?"

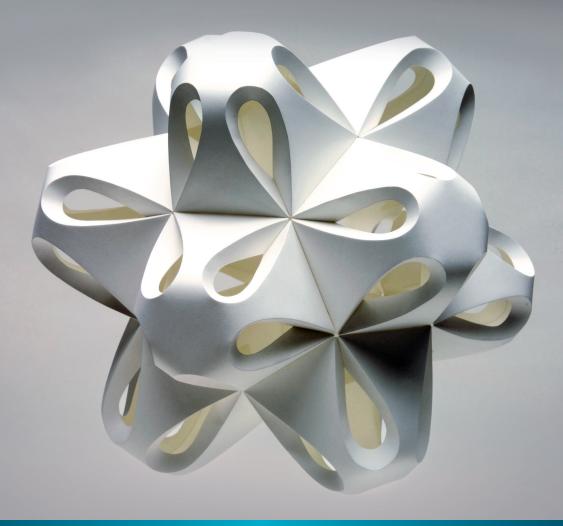
There are also econometric and data-quality limitations for all marketing mix models, which may frustrate clients when they are explained, and give false confidence if they aren't. Variables in the model affect each other, as well as the outcome that the modeller is trying to predict, making it hard to do accurate 'what if' predictions.

"It's no longer a journey where you see a TV ad and, when you're in Sainsbury's, you buy a can of beans. It's more like: you see a TV ad; you might search for it; you might end up on a brand site; you might have a friend who sent you a video; you might see an ad that you hate and you start complaining about it on social media. There are a lot more paths of communication and of interaction with brands," says Rob Sander, a principal consultant at Data2Decisions, who is working with digital media owners to build models that capture the role of influencers such as social media.

The data sources can also be hard to mash up into a single model, as they are measured at different levels of granularity – store level, household, individual or segment – and with different frequency and precision. Models compare many types of data that correlate poorly:







MONDELEZ HAS A 3P MIX

"My responsibility is to run marketing mix modelling for all of our chocolate brands in Europe," says Matt Stockbridge, growth analytics manager, Mondelez International. This means he models the marketing for some of our national treasures – not least, the Creme Egg.

The marketing mix models that Nielsen builds for him may take six months to create, and crunch three years of sales and spending data, from the store level upwards – as well as competitor and social media data. But Stockbridge warns that year-on-year consistency and clarity in the model-making process is vital.

This is simplified in his business, because the type of frequently purchased, inexpensive products he analyses respond to changes in a restricted set of marketing strategies: three of the 'Ps' of marketing. "As an FMCG brand, the contributions of distribution [place], price and promotion are between 90% and 95% of our effectiveness. The

contribution of media is still there, but it is much lower," says Stockbridge, who jokes: "I'm envious of people who sell things like cars sometimes."

While his models show that some types of activity have vastly more influence than others, this doesn't mean that nothing changes – or can change – in the marketing mix. On one hand, Mondelez is keenly aware of the effects of shifts within retail distribution – such as the rise of the discounters. On the other, it knows short-term promotional activity can have a huge habit-forming effect, which is not just immediate.

"We work on the principle that you don't have a long term unless you have a short term," Stockbridge says, "and a £1 deal on a multipack is easy to model."

In some cases, he may build a smaller tactical model that can report effects to the relevant part of the business in a shorter timeframe – with the caveat that the price elasticities and

coefficients that this model spits out can't simply be scaled up to the overall marketing mix.

A lesson that Mondelez has learned in recent years is that – if your sales uplift targets are in six figures, your distribution is more or less everywhere, and your target market is the population of the country – it is easy to be too detailed in the nuances of your models. This was apparent in the early days of social media.

"Like everyone else, a few years ago we got very excited targeting specific groups of people on specific platforms," Stockbridge says. "But we need to create a response of millions of people each time we do something. We need to drive reach and penetration rather than frequency. The idea that we will create an amazing piece of viral content is attractive, but we don't plan for it. So online, our marketing is really about paying for eyeballs."

For that reason, his models tend to be fairly stable year-on-year – there is a trade-off between trying to measure the evolving complexity of the marketing mix by constant tweaks, and producing results that can be compared across countries and through time. Despite criticism of marketing mix modelling as a discipline, he has found that it has only gained in influence inside Mondelez.

"In the past, Cadbury modelled the effectiveness of its media, rather than what drove its business, but now the approach is to help us make better business decisions. Five years ago, I would not have had a senior marketing person coming to me to say, 'I have £2m to spend, I want £5m back, where should I put it?'

"Results from our models certainly help, internally, to decide debates. We used to face questions from people that doubted whether our numbers were correct, asking 'what about the art?' Now, our marketers understand there are literally millions of historical observations built into our model."

• Given these flaws, why do it? Because, for many brands, the results of using this type of data even if the models are far from perfect – often clearly outperform any other approach. In a 2013 study published in the International Journal of Research in Marketing, Frank Germann et al found that marketing analytics created an 8% higher return on assets for the companies that used it, rising to 21% in highly competitive industries, with all sizes of organisation benefiting. If the answer to doubts about accuracy in the modelling process is to increase trust in gut feel, or do the same thing as last year, that isn't palatable to most clients. Which is why, Sanders says, any dissatisfaction with large-scale marketing mix models is not a rejection of the principle of using data to try to optimise marketing.

"It's always exciting to say something's dead," he says.

Other vendors – all of whom offer alternative types of marketing analysis if clients require it – have similarly not noticed a reduction in demand for the models. "We build models for 60% of MediaCom's billings – and the other 40% are, more than likely, building MMMs with another third party. Our clients realise that it's an important way to measure the effectiveness of their advertising budgets," says Matthew Wragg, head of attribution at MediaCom. "Today, this isn't just about finding out how well TV is doing compared to press, but also how this fits into the wider marketing picture – the wider commercial perspective."

"I think marketing mix modelling still has legs," says Nat Greywoode, who works in the market insights and analytics team at Twitter, and who has spent his entire career building marketing mix models. "In my mind, there is no alternative."

While the challenges to MMM's effectiveness are getting more complex, Greywoode's argument is

that techniques to improve the models are known, even if they are not universally implemented yet. The problems of integrating new types of data – especially the social media data that Greywoode is working with modellers to incorporate into their MMMs – are the sort of knotty challenges that, in every industry, inspire econometricians to greatness or drive them to despair. But they are problems with solutions.

How the sausage is made

Statistical models – not just in marketing – create two significant and related problems. The first is that the map is not the territory; every model is a subjective simplification of reality that is built to solve specific problems. "My favourite analogy – the most complete map of the Earth is Earth itself, but a simpler map is much more useful," says Koen Pauwels, a professor of marketing at Özyeğin University, Istanbul, the author of *It's Not the Size of the Data – It's* How You Use It, and a specialist in the theory and practice of modelling the value of marketing.

This means that the modeller will always be trading off simplicity and transparency – against realism. Experience can suggest how to build a good model, but it cannot offer a definitive guide to every structural decision. Again, like all econometrics, modelling the marketing mix is both an art and a science – and the state of that art is constantly being tweaked and refined. "It's not just a mechanical process of chucking some data into a piece of software, which spits out some numbers that you provide back to the client," says Sanders. "What we try to focus on is building the first stage - then you can make recommendations and have discussions with your client about what the insights may be."

This often leads to a second misconception: that a market mix

model will give definitive answers to every question asked of it. In reality, it can only suggest evidence, based on historical information, that has been measured as accurately as was practical. "You're trying to assist decision-making. Because it is based on numbers, I can see how people jump to the conclusion that it will give a true answer – but there is no true answer," Sanders says.

This is why the teams doing the modelling have learned to be sceptical of surprising results from their prototypes, or radical innovations from newcomers. Wragg says he regrets that a few "cowboys" attempt to catch attention by creating models that seem to show exciting results. "That's a benefit of being in a big team," adds

Integrating new types of data is the sort of knotty challenge that inspires econometricians, or drives them to despair

Emma Whitehouse, associate director, MediaCom Business Science. "When you get surprising or unexpected results, there's always another client that has probably run something similar. You can look across a group of clients to get some benchmarks, to give you a better feel."

In a recent case, Whitehouse was surprised to find a cinema campaign performing unexpectedly well. She cross-checked with colleagues, and found that similar effects had been seen in other models – but with the proviso that the effects were mostly short-term. That gave her the contextual information to trust the result and present it appropriately.

If using context, precedent and common sense improve modelling – and data quality is fundamental – at least one of those factors may be in short supply. In which case,



• sometimes smaller, cheaper or more generic models are all that is possible. Marketing-mix modelling may not be easy, or even possible, for new entrants to a market or start-ups, which don't have enough historical data to feed the models. In those cases, the client can still rely on the breadth of data that the model builder can access.

Nielsen, which has more depth than most, now offers marketing mix modelling built to solve exactly this problem. Earlier this year, it launched Benchmark Media Optimizer: this is designed to use benchmarks for its custom models as proxies for marketers who want to improve their performance using data, but who don't have data to feed a model. Because MMM uses aggregates, Nielsen's aggregated data - on the returns on marketing investment of 1.600 brands in 30 categories, in more than 40 countries – may improve a client's decision-making, but, as the company admits, it "won't deliver

the breadth and depth of results" of a custom-built model. One constraint might be that a model for a new entrant – as in the Google anecdote mentioned earlier – does not account for the customer's perception of the quality of the product, which may be one of the dominant drivers of growth.

At the other end of the scale, as more marketing becomes data oriented, it may be tempting to throw every piece of marketing data into the mix. It's intuitive that, by doing this, you will end up with a model that captures everything useful – as close to Pauwels' globe-sized map of the Earth as it's possible to get. Unfortunately, econometrics doesn't work this way.

Apart from the time and cost of collecting data, the models are confounded if many of the inputs used are closely related to each other, and all related to the outcome – so it may give exciting, but wildly inaccurate, predictions. If you use the accepted threshold that you are

95% confident something is affecting sales, and you report 20 influences, one or more of them will randomly seem important.

"Just by chance," Sanders explains, "some things will correlate positively. Then, if you chose to interpret that as an actual causal relationship, you'll say 'oh, look –

MMM may not be easy for new entrants to a market, which don't have enough data to feed the models

this tiny thing drove 10% of sales'. You can end up in quite ridiculous situations if you're not careful."

This limitation – that too much data may sometimes be as big a problem as too little – is driving dissatisfaction in some clients, because their fastest-growing sales channel generates large amounts of

A CRISIS IN MMM?

Since the 1980s, Michael Wolfe has done marketing mix modelling for firms including the Kellogg Company and Kraft Foods. He introduced it to Coca-Cola in the 1990s, and set up Bottom Line Analytics in 2003. His consultancy is 70% based on MMM, but Wolfe has recently become disenchanted with the lack of innovation or rigour in the discipline.

Last year you wrote a blog titled 'The Death of Marketing Mix Modelling'. Why? Michael Wolfe (MW): In the US, the attitude is becoming increasingly negative towards this tool. They are running the models and not really making decisions from them, because they don't believe the results. I can name three large companies that have ceased using it - maybe not permanently, but at least for the time being. The biggest problem for them is digital attribution. I think there are a lot of larger issues than that, but that is why

companies such as Campbell's Soup have quit doing it.

You said digital attribution isn't the most misleading feature of many models - what is? MW: That the measure of the short-term effect of negative. If you know Coca-Cola, for example, the long-term effect of advertising for that product in its better years was five or six times the short-term effect. I even did one health and beauty brand last year where the long-term effect of advertising was 20% of total revenue. If you don't measure this effect, why are you running the model?

Have we reached a point where problems such as integrating new data streams mean advanced MMM will only be for a few large clients?

MW: I hope not. When there

were only three TV networks and not a lot of promotion being done, simple models seemed to work OK. But, increasingly, complexity is part of the issue.

You say the voice of the customer is absent from many models. Have you managed to integrate social data?

MW: I've used some linguistic algorithms to turn social media conversations into a group of metrics, which I can sum up in one metric. I've done it 28 times and this metric has an average correlation to sales of about 80%. When I put this into a model, it is a very large contribution to overall revenue. The brand experience in real time is a major driver of almost every business, and it's simply missing from most marketing mix models.

You also claim models do not accurately capture the impact of quality in advertising.

MW: We have partnered with an advertising copy-testing

company, which tests ads that are in market – up to 1,000 every week. I've put these copy tests as a scaler inside my models, and it works. I'm able to decompose the creative portion of the advertising from the investment – and we find about 60% to 80% of the advertising effect is the creative!

Are you optimistic or pessimistic about the future of these models?

MW: It's increasingly becoming a crisis, at least in the US. There's a lot of talk and dissatisfaction from the advertiser world, but it's going to take some fortitude for them to change. It's going to happen when a large advertiser says: 'I'm tired of this. This is the direction we want to go; here are the answers that we need to build into our models.'My optimism is that we know what the issues are. I think I've some of the answers, but there are probably other versions somewhere. I'm not saying I've got a monopoly on the truth, but there are solutions.

The digital challenge
"In the past, networks such as ours



identify that an online search led to

not just how much they contribute



social media often works by catalysing other types of marketing spend. This means that there are collinearity problems – when the return on offline goes up, so does the return on online – and temporal issues; not all the sales effects of digital happen at the same time, so may not be captured in a slice of data. This means that naively adding spend on Facebook, Google search or Twitter to the model without careful thought may decrease that model's ability to predict.

"We've seen not only a direct impact of Twitter exposure on sales, but also an indirect impact. So Twitter kind of influences your customer on to the next step of the journey, which is intuitively understandable," Greywoode explains. "This is the No 1 area where brands are missing a step, in that they don't tend to look at the stages along the way to a sale. They look at how Twitter exposure equals sales, whereas there are other impacts that would maybe ladder-up to that."

Neustar MarketShare is another model builder attempting to improve how it treats digital in the marketing mix. It recently partnered with Google, and can offer its clients slightly more granular search data in its marketing mix models.

A recent example of how this data could be used was a benchmark analysis on the respective contributions of television, digital, video, social media, search and other forms of advertising as an influencer of sales for action films. It found that TV was the most effective influencer, but it swallows 82% of marketing budgets, so is over-invested. On the other hand, increasing digital investment from 10% to 34% of marketing budget would improve RoI for distributors.

In other areas, Neustar MarketShare predicts where the influence of digital on sales has been hyped – and then its models may also dampen expectations, it admits. Dirk Beyer, vice-president of data science research, doesn't try to pretend that modelling digital's influence is easy. "Paid search, for example, is highly correlated to who is in the market, or how many people are in the market for your product at a given time," he says. "It's highly correlated to seasonality, to the hype that's around the product, and so on. If you just throw in paid search as a single variable in a monolithic model, it will look like everything you do has been driven by paid search.

"We're now in a position where we get that data directly from Google, at a level of granularity that is exactly what our brands need. That allows us to do a much better job teasing out the actual impact of search versus – for example – the upstream advertising that got people to the point where

The difficulties of working with digital data should not cause clients to abandon their models

they started searching. That's a big deal for our clients."

The difficulties of working with digital data should not cause clients to abandon their models, Sanders says – just to question the providers about how digital has been integrated. He also argues that it's not a reason to abandon this type of modelling, because these are the type of econometric challenges that model builders such as IRI (see p34, More than one tool) have been facing for years.

It has similarities to the way models face the challenge of splitting the immediate sales effect of advertising from its long-term brand-building effect (see p30, A crisis in MMM?), or the interaction between advertising and promotion.

Circleresearch

We're the B2B experts

It's all we do and we're good at it.

We have an intimate understanding of B2B environments, we know how to intelligently engage the people that matter and we bring our experience with hundreds of other B2B businesses like yours to bear.

The result is smart research which makes a real difference.













• "It is interesting to compare how something like Twitter or Facebook works compared with TV, press or out of home. Actually, we have found they are quite similar," Sanders says.

"There would be a danger if you obsess about measuring Facebook or Twitter's RoI, and you change your model in other ways to make that happen. Then you've probably just built a bad model."

A better model

The solution to building the best possible marketing mix models may sometimes lie outside the models themselves. With new data sources that interact with each other, markets that evolve, and the need for rapid decision-making, creating a monolithic piece of software that answers every marketing budget question (until machine learning and artificial intelligence can replace

econometricians) may not always suit the commercial imperative that created MMM in the first place.

In that case, MMM may become a toolbox of methods, rather than one model, designed to answer a range of questions in a coherent way.

The first challenge is how to model smaller customer segments using MMM. While large FMCG brands such as Mondelēz (see p28, Mondelēz has a 3P mix) do not use MMM for this, other clients do. IRI, for example, is creating models that can capture the returns in increasingly small segments – but granularity of data is a constraint.

"MMM is a top-down, aggregated viewpoint," says Wragg. "And with only so many data points, there are small channels that we have, historically, not spent lots of money on, or don't reach millions of people – so we sometimes struggle,

statistically, to pull those factors out of the models." For such models, he suggests, attribution – enhanced to capture the impact of non-digital spend – may sometimes produce more actionable insight.

At Data2Decisions, a strand of research focuses on how to combine the top-down and bottom-up

MMM may become a toolbox of methods, rather than one model, designed to answer a range of questions

approaches, in what Sanders calls an "ecosystem" model. "It's the same kind of metrics, but we just broaden out the scope of our analysis. Rather than simply building a model of

MORE THAN ONE TOOL

Rather than build a marketing-mix model that tries to solve all problems, some providers are focusing on how to complement it with other tools. We asked Carl Carter, head of media and connected consumer in Client Growth Delivered, part of IRI Worldwide, how this works.

Do you recognise the broad criticisms of MMM?

Carl Carter (CC): Marketing-mix models give a very complete and comprehensive picture of marketing performance, but they do have shortcomings. The breadth of measurement across all demand drivers creates an high-frequency, rapid cycles of digital campaigns, clients often need insights in days, rather than weeks and months. Digital is typically measured at aggregate platform level, without measuring executional mix, including publishers, creative and content genres. And campaigns that have a niche targeting – for example affluent millennials - are not measured accurately; their impacts are washed out in the aggregation process.

How have you approached these problems?

CC: We have developed methodologies around A/B testing using stores and consumers, as well as variations of attribution models – including reattribution models. It lets us go faster and deeper, and address the detail in targeted campaigns.

How does A/B testing work in this context?

CC: Using the simple approach of exposed versus non-exposed, we can apply this methodology to our store data, treating each store as the proxy for sales purchase data. We can use this to test creative, format, placement and publisher, and insights are delivered in less than two weeks. This analysis is flexible too: some clients use this for testing before large-scale media investments; others employ it as part of a continual measurement of digital campaigns. We have also begun to integrate the findings alongside marketing mix. This provides calibration for the marketing-mix analysis.

What are reattribution models?

cc: They use a secondary regression to understand the relationships between various marketing touchpoints measured within the marketing-mix analysis Using this approach, we are delivering some very complex insights for digital – using core marketing-mix principles – but

Digital advertising has made attribution models more attractive, leading to criticism that MMM's aggregated data throws away useful data.

CC: MMM is often criticised for being too macro, but this is often self-imposed. For instance, IRI builds store-level mix models and pushes marketing data as far down as possible to match this level of consumption. Not matching down to individual households or consumers is what permits MMM to measure across all the marketing levers, including offline. Marketing mix can have a bad rap purely because the data that feeds the model isn't at the level of granularity to push the model to its limitations.

How do the restrictions on the data that platforms such as Facebook provide affect the power of MMM?

CC: Often, you can find that digital content owners and ad networks operate in 'walled integration of data with other understandable because of of the data. Walled gardens are not an issue for MMM, as consumer segment is the lowest it goes down to. Combining store-level data, causal effects and the most granular media data, marketing mix is still very powerful. We are also seeing media owners offer increased granularity of data for more detailed analysis.

What's the future for these marketing mix models at IRI?

CC: Marketing mix requires some support to improve granularity and speed of turnaround. We think the only effective way to deliver this is an integrated measurement strategy, in which it is complemented and supported by a range of methodologies.





 influences of those touchpoints, the output can then be fed back into the marketing-mix model to improve its recommendations.

Increasingly, marketing-mix models may also take advantage of natural, or planned, experiments. Correlation between an input and a key performance indicator (KPI) of a model – such as social media

Ultimately, the test of a model is how effective it is at helping the client to make better decisions

sharing and sales – might be causal; the more people share, the higher your sales will be. Or both might be caused by a third factor – such as TV advertising. Finding two sets of data for which the main difference is that one group was exposed to TV advertising and the other was not would settle the argument.

Brave clients may even decide to run an experiment to test model predictions – an approach recommended by Pauwels when he builds models for his clients.

"Individuals and organisations need to experiment to figure out what really works for them.

Unfortunately, experiments on a large scale are often considered costly, risky for one's career, and tough to sell, politically, in the organisation," he says. "Who wants to be in the 'control group' for a potentially very successful new marketing initiative?"

Pauwels built a model for an office-furniture company, which then decided to settle the question of causality by running a field experiment, which divided the market into two groups. One group doubled the spend on what was predicted to be the most effective channel, funded by cutting the least effective one in half.

"Profit, net of marketing, increased 14 fold in the experimental group versus the control. After the experiment, we again modelled the data and showed the company should stop reducing the least-effective action, but it still had room to increase the more effective one," he explains. Pauwels calls the approach MEME – model, experiment, model, experiment.

Ultimately, though, the test of a model isn't just how accurate it is – it's how effective it is at helping the client to make better decisions.

Increasing complexity isn't always the way to go; creating the structure to answer precise questions will achieve better results than simply adding more data, says Pauwels, who might use the output of one model to decide how to build the next stage.

"Senior managers often ask me to evaluate the marketing allocation among communication channels. For this purpose, I would not include information on the specific campaigns within each channel. Next, I may build a model that distinguishes the separate effect of TV campaigns, treating the other marketing channels as control variables. Then these smaller models can fit into bigger models."

"At the end of the day," Greywoode adds, "MMM is a bet. It's hopefully as good as you possibly can get. Clients should challenge their vendors to make the models as representative as possible, but – once they are satisfied with that – focus more on the uses and measuring the outcomes." I

MEASURING TWITTER: FOUR TOP TIPS

Twitter is committed to helping marketers measure the impact of their advertising on the platform.

Our measurement tools sit within three pillars – Reach, Influence and Action – which allow marketers to understand whether a message has been delivered to the right audience, whether it has changed the audience's perception of a brand, and whether the audience took an action – or made a purchase – based on the message.

Twitter's marketing-mix modelling (MMM) programme is an established part of the Action pillar, supporting our largest global brands in their Rol measurement efforts and supplying tactical insights to help improve their activity on Twitter.

To date, we have partnered with more than 30 brands, in five markets and sectors, achieving impressive returns of around 2x investment on average.

In all these partnerships, we have worked with the client's preferred MMM agency and some of the best practitioners in the industry, who have helped guide and evolve our point of view on how Twitter – and social channels at large – should be measured within MMM.

Here are the top four things we have learned from measuring Twitter through marketing mix models:

1. Data is king – MMM agencies that collect impressions, views and engagement from Twitter sources, and include them as their data inputs, were able to achieve more accurate and insightful results on Twitter performance for their clients.

2. Product and platform changes – Twitter's ad offering has changed significantly over the past two years – particularly in the advent of new video products – which is often overlooked by reporting a single Twitter Rol figure.

Best-practice models included multiple Twitter input variables to capture different executions and changes over time.

3. Consider the power of organic tweets – For some brands, driving incremental consumer tweets and engagement as a result of their advertising activities is a significant outcome in itself, and can be a

meaningful amplifier of their bottom-line KPI. Some agencies took steps to identify the importance of this for their client and included additional inputs to reflect this in models, alongside paid Twitter activity.

4. Keep in mind the objective – Agencies that clearly understood what their client's Twitter activity was trying to achieve were able to set up the analysis to extract the most value for their models. In these cases, models were modified to include Twitter advertising's influence on brand health metrics, as well as intermediary outcomes, such as search volumes and website visits.

By **Nat Greywoode**, market insights & analytics, Twitter