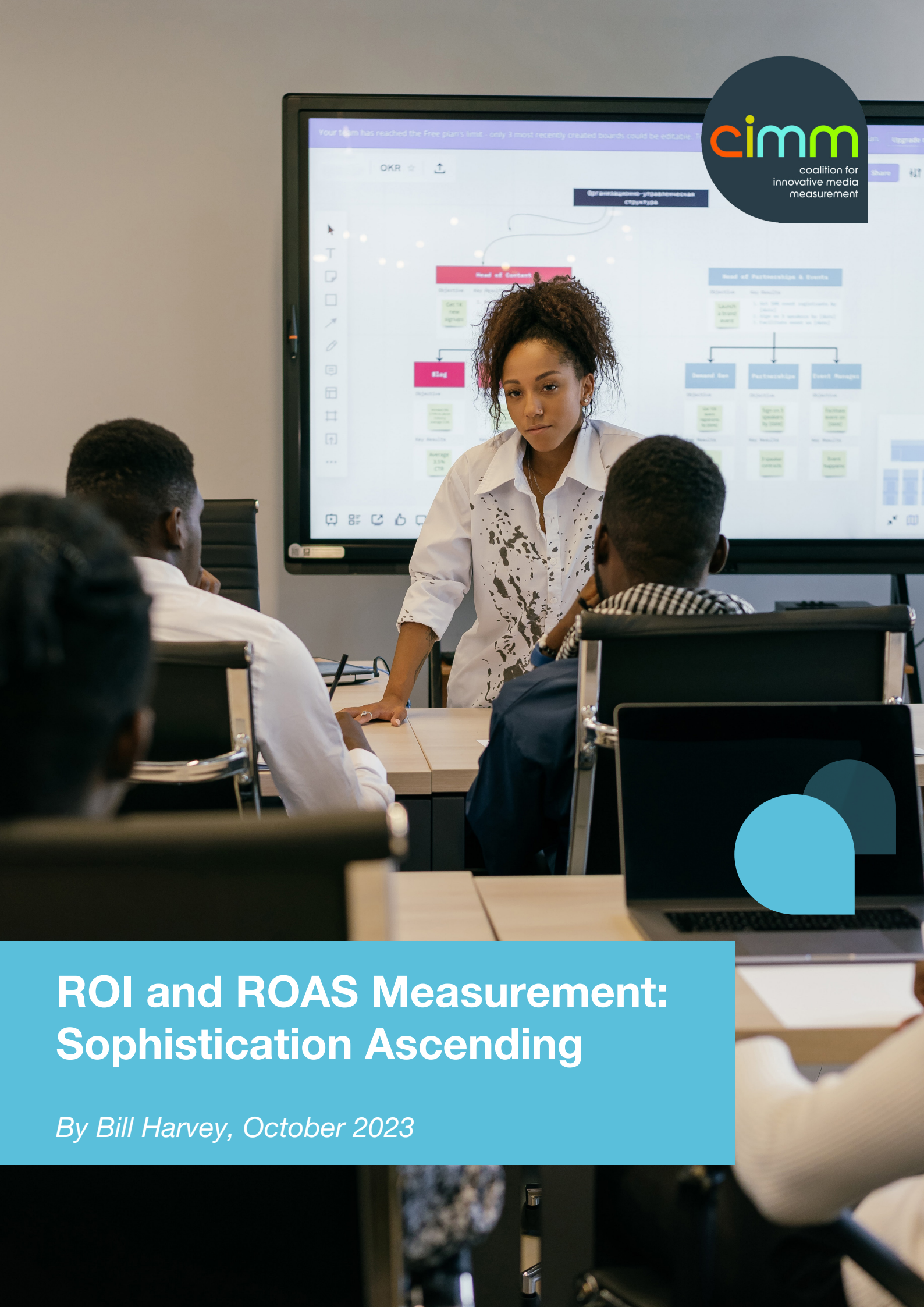


coalition for innovative media measurement

ROI and ROAS Measurement: Sophistication Ascending

By Bill Harvey, October 2023

ROI and ROAS Measurement: Sophistication Ascending

By Bill Harvey, October 2023

CONTENTS

Executive Summary

This report explores the most promising new developments in ROI and ROAS measurement, including MMM, singlesource, random control trials, MTA (a breed of singlesource which arose from digital), agent based modeling (ABM – MMM projected down to simulated household level), and other new methods.

The cross-analysis of media spend and sales data to deduce the contribution to incremental sales produced by advertising and other marketing stimuli began with what is today known as Marketing Mix Modeling (and as Media Mix Modeling in cases where non-advertising stimuli are generally excluded) in 1948.

By the turn of the century, MMM had become firmly established as basic to the toolkit of major advertisers. It was criticized for the degree of subjective analyst judgment required, weaknesses in the treatment of the raw data, slowness of results, and the fact that management could not understand how it worked, among other criticisms. Some advertisers tested alternative suppliers of MMM and found a good degree of disagreement especially for the newer media types.

Singlesource (actual longitudinal measurement of ad exposure and sales in the same households, not a feature of MMM) using small (5000 and below) panels briefly arose in the U.S. and a handful of other countries, but were not economically sustainable, and sample sizes were too small for most brands to see much statistically significant value.

The invention of big data singlesource (TRA, 2005) led to the use of digital ad tags, set top box data, smart TV data and other such data sets with sample sizes in the millions of households to be used with same-household purchase data by hundreds of companies, diminishing reliance on MMM to a degree. With the rise in privacy laws and announcement of the deprecation of third party cookies, advertisers refocused on improving MMM which would not be affected by these forces.

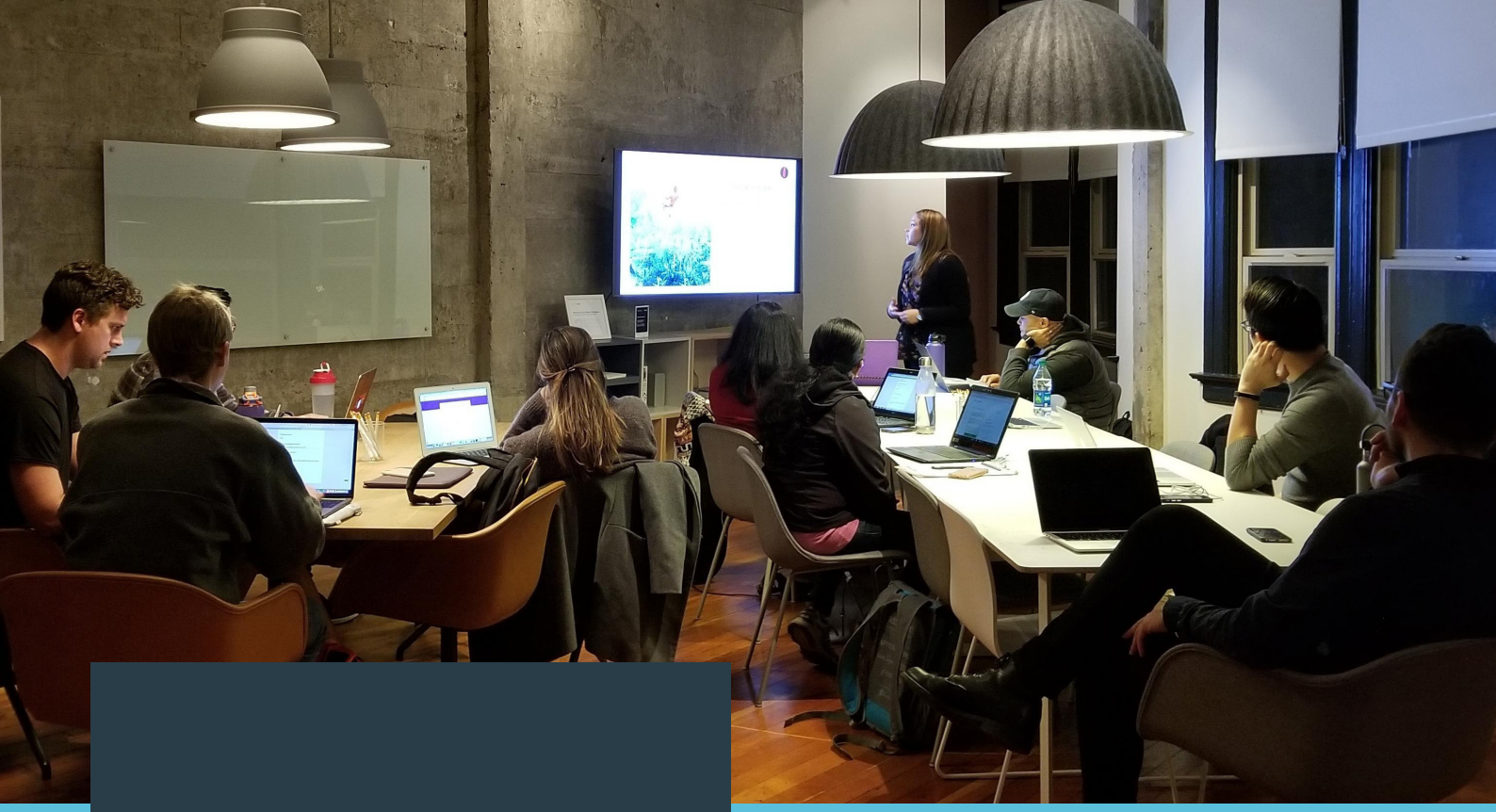
As might be expected, artificial intelligence (AI) is already playing an important role in many of these innovations. New evidence suggests that the most sophisticated practitioners will be using a combination of these methods, including ABM simulated populations on which *synthetic trials* may be performed, and that more subtle factors such as creative, context effects, media saturation effects, targeting, and synergies will play a larger role in the future of outcomes measurement and optimization of short and long term effects of marketing.

Whereas advanced forms of MMM will continue to help make better decisions about budget allocation to broad media types vs. other types of marketing stimuli, singlesource systems will continue to help make faster inflight decisions about phenomena not well reflected in most MMM today, such as creative, target groups, context effects, attention/resonance, frequency, recency, synergy effects and other leverageable variables which can make major contributions to overall ROI/ROAS.

Random Control Trials (RCTs) and other forms of testing/experimentation in market continue to be the gold standard and the truth standard for the modeling-based methods, and the practical difficulties of RCTs could be assuaged by the leadership of media companies, in which addressable commercials would be the essential enabler.

I. Introduction

CIMM invited Bill Harvey Consulting (BHC) to conduct interviews and produce a study white paper, and to present and discuss the topline findings at the 2023 CIMM Summit, to better understand why there is a resurgence in MMM, and how MMM is evolving along with other more granular ROI/ROAS measurement and optimization methods.



II. A Brief History

Professors James Culliton and Neil Borden at Harvard are credited with the coining of the term “marketing mix” in 1948[1].

Outside of academia, Interpublic’s Marplan leader Herb Krugman in 1958 carried out what he called a “Grand Scale Analysis” for a major beverage client. However it wasn’t until the 1980s that this writer automated an MMM tool of his own devising for General Foods. After applying multiple regression analysis to deduce the sales effects of each media type, the system automatically set up matched market tests to confirm the findings of the model based on this species of random control trials. A major finding shared with the ANA was the S-curve effect, i.e. different markets would be at different stages of the dose-response S-curve for any given brand at any given time, so that local TV ad weight could be added in markets where the brand’s potential for growth was highest (in the middle of the S curve).

During the period of the 1960s through the 1980s, the dominant belief in the advertising industry was that the sales effects of advertising could never be separated from all of the other stimuli affecting sales. This was popularized across the industry by Stuart Colley’s book *Defining Advertising Goals for Measured Advertising Results* (DAGMAR) which institutionalized brand lift questionnaire-based studies as the best way to assess the value of an advertising campaign.

However, use of MMM began to grow during the 1980s, driven in part by the availability of barcode scanner data in supermarkets. Many companies including Nielsen, IRI, MSA, MMA

1 Source: https://en.wikipedia.org/wiki/Marketing_mix

and others began to offer MMM, and today tens of thousands of MMM modelers operate in the field globally. By 2000, MMM global revenues exceeded $2bn according to the Honimichl Report, and have continued to grow.

Reporting history based on the writer's own experience, during the same period, IRI, founded by John Malec and Gerry Eskin and led by Gian Fulgoni, created Behaviorscan, a method of singlesource (household level – "singlesource" someday could mean "persons level") ROAS analysis. BScan was based on a checkerboard design using cable operators in small markets to create matched household panels whose purchases were tracked via credit cards issued by cooperating supermarkets. IRI also collected information on promotion and price and used these signals as predictors as well, thus transcending ROAS (Return on Ad Spend) and becoming a species of ROI (Return On all marketing Investments).

Max Ule's Milwaukee Ad Lab, Marc Portes' ERIM in France, and John Adler/ Bill McKenna's Adtel also pioneered improvements in the BScan methodology, which improved significantly.

In the 1990s, Ed Dittus and Gian Fulgoni analyzed hundreds of BScan studies in two waves to create a meta-analysis called AdWorks. AdWorks proved that advertising delivers on average twice the ROI of promotion (advertising heavy-ups were profitable 40% of the time vs. promotion 20%), and also determined that the optimal allocation of TV budget was 38% in primetime broadcast at the time. This was derived from an analysis of over 500 campaigns with a very wide range of variation of percent of dollars in primetime broadcast. ROI was maximized at 38% of dollars in that media type.

Ed Dittus writes a fascinating detail of how AdWorks was put together in **Section X, In-Depth Insights from Pioneering Innovators.**

Another method was conceived originally by Colin McDonald who called it "singlesource" meaning that both purchases and media usage would be captured from the same panelists, so that ad exposures could be estimated and related directly to changes in purchase behaviors, captured in diaries. In the mid 1980's, working with Time Inc.'s SAMI Sales Measurement unit, Bill McKenna, President of Burke's Test Marketing Group (which now included Adtel) conceived a singlesource TV measurement system named ScanAmerica that merged local TV market audience measurement and product purchases. In 1986, McKenna sold the ScanAmerica service concept to Arbitron.

As proposed by McKenna, ScanAmerica would be launched as a new National Single Source TV Measurement Service, competing directly with Nielsen's National TV Ratings monopoly. However, Arbitron elected to launch the ScanAmerica concept as an upgrade to its established local TV measurement service.

Continuing this history based on the writer's personal experience, after testing ScanAmerica in Denver during the late 1980's, Arbitron formally announced and launched ScanAmerica in 1991. In each of five established Arbitron local TV Markets, handheld barcode scanners and peoplemeters were installed in 1,000 probability sampled TV households. Arbitron committed to progressively increasing each local market sample size to 2,000 HHs.

U.S. national probability panels enabled singlesource (household level) modeling vs. MMM's typically market level aggregate data. By the end of the 1980s, Leslie Wood analyzed the data and found that two exposures within the 48 hours before a shopping trip had strong lift effects on incremental sales.

Singlesource

Singlesource panels varying in size from 1,000 Households or greater appeared around the world, but proved costly to use, especially when considering that sample sizes only allowed the largest brands to obtain statistically significant results. These panels appeared and shut down because the economics were not sustainable. However, some advertisers, notably Mars, made great use of the methods while they existed, the learnings increasing their ROAS from 70 cents to $2.00.

The last of these panel initiatives was the P&G and Unilever led Project Apollo, which appeared briefly in 2006 but again, could not be sustained, and the sample size of 5000 was again shown to be not useful for the majority of brands.

At Next Century Media in the 1990s, the idea of using big data matching to achieve singlesource at scale was spread across the industry, and NCM software was licensed by the startup TRA in 2005 to actually create a company to provide singlesource ROI and ROAS measurement on sample sizes in the millions. TRA was later sold to TiVo.

During the period 2005-2014, TRA grew to a client list of 77 top brands, and most of the major agencies and networks. Although it offered random control trials, most clients did not use experimental design but simply decided based on observed ROIs. These decisions focused on shifting to what is today called “advanced audiences”, changing targeting from solely sex/age based, to instead maximize reach of purchasers of one type or another.

A major finding that changed industry buying habits was that targeting heavy category users who bought your brand but not loyally (TRA named Heavy Swing Purchasers) provide the highest ROAS. This was later confirmed by Leslie Wood and by MMA consultant Joel Rubinson, former Chief Research Officer at the Advertising Research Foundation (ARF). Joel refers to this “Heavy Swing” phenomenon as the “Moveable Middle” and argues that simply pivoting targeting to this group is even more important than measuring ROI, because it is *producing* ROI.

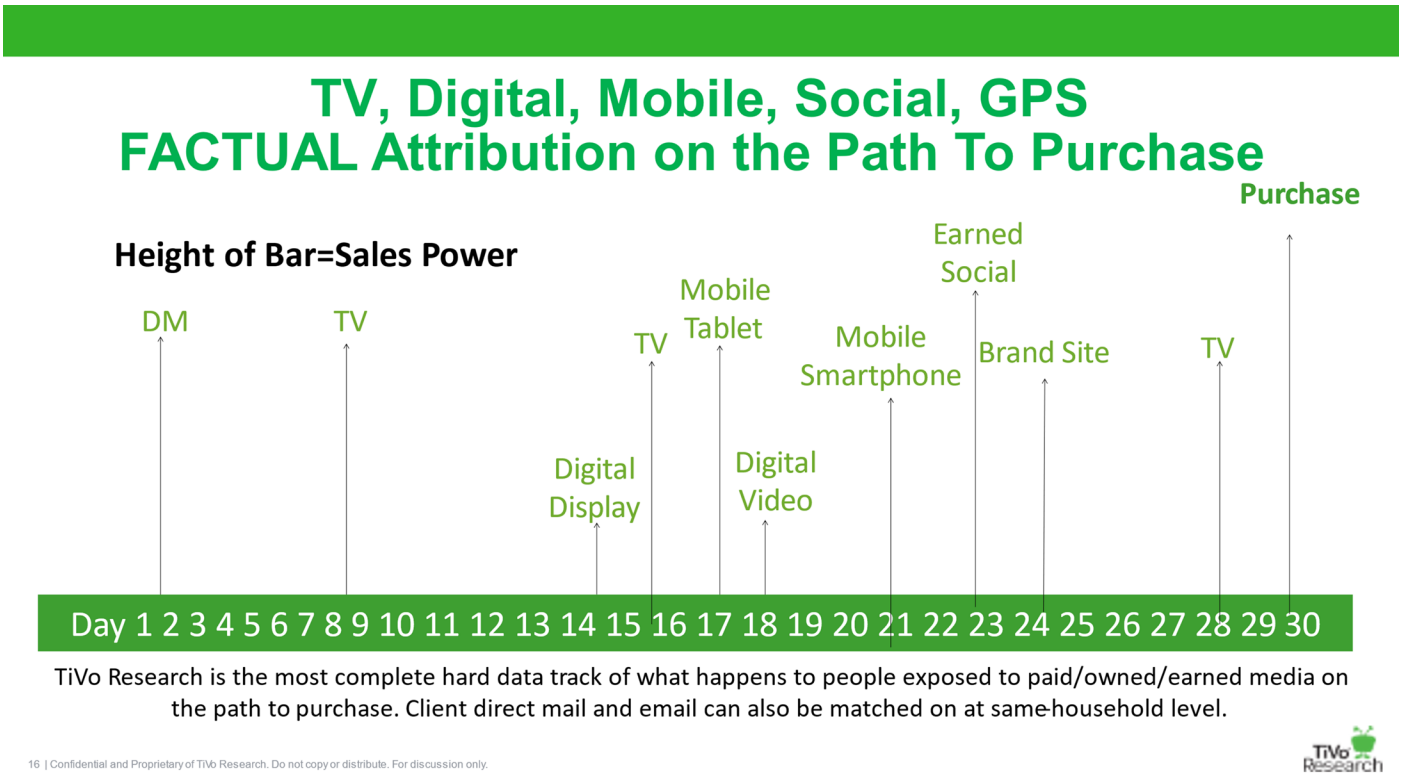
Ross Link, founder CEO of Marketing Attribution, an MMM company, argues that this remains the case: “We have also seen many times that targeting ‘fallen buyers’ is one of the highest ROI targeting strategies.”

Section X of this paper, **In-Depth Insights from Pioneering Innovators**, was created in order to provide ROI leaders space to expand upon their ideas. In that Section you will find more detail on Joel Rubinson’s thoughts and those of other movers and shakers.

Since TRA, as thousands of companies emulated the use of big data for outcome measurement, the growth rate of MMM was temporarily restrained. During the same time period, as digital caused marketers to want to be able to measure the ROAS of all of their media, many of the television networks began offering “value add” packages in which brand lift studies (typically questionnaire based) would be packaged in with working media spend.

MMM continued during the same period despite temporarily slowing its growth. Another factor during the same time frame as TRA was a digital version which came to be known as MultiTouch Attribution (MTA). MTA like singlesource captured the arc of the consumer's journey in detail. TRA (by this time renamed TiVo Research) responded by calling itself a form of MTA which included TV as well as digital:

Exhibit 1: Household purchase journey for auto sales[2]



During the heyday of TRA, in January and February 2011, the company went to Washington D.C. to share their privacy-centric methods with the FTC, FCC, and Congressional Privacy Committees, who agreed that these methods made it unnecessary to create laws to regulate privacy protection.

The TRA methods received four U.S. patents, and reflected the CASIE Privacy Principles issued in 1995 by ANA/4As/ARF and led by the author of this paper. However, one of the privacy tentpoles of the TRA method was that Experian and Acxiom would be the only entities to have Personally Identifiable Information (PII). Another important aspect of the method was to not add information to either Experian or Acxiom. These principles today are reflected in the new phenomenon of clean rooms. However, starting around 2012, the number of entities creating their own Identity Graphs containing PII grew quickly, resulting in significant privacy concerns as "Stalker Marketing" practices bloomed in digital, and the public became more sensitive to "creepy digital marketing".

With the creation and spread of privacy laws, singlesource and MTA were feared to have obstacles in their future, which has now caused a widespread move back to focus on how MMM can be improved, since it may be all we have for ROI and ROAS measurement at some point in the future. As a result, MMM is once again growing at an estimated +12% per annum (Marketing Science Institute).

2 Source: TiVo Research. Exhibit 1 shows a single household's journey in the last 30 days before buying a new car. Analyses combined such results across millions of households.

III. Roots of the New MMM

The roots of the latest advances in MMM can be traced back to the work of earliest pioneers in the 1980s and 1990s whose work centered around television and promotion, and companies such as Management Science Associates, Nielsen and IRI, and innovators such as Gerry Eskin and Magid Abraham while at IRI, Al Kuehn, Ross Link, Ed Dittus who founded MMA, and many others too numerous to mention.

A most interesting chapter in which MMM came to include digital and the first cross-platform analytics was led by Rex Briggs. His fascinating story is told in **Section X, In-Depth Insights from Pioneering Innovators** along with contributions by other major figures in the field today. Rex was also first to pair MMM analytics with survey data and executive dashboards. Wes Nichols and Jon Vein at MarketShare Partners, acquired by Neustar which is now part of TransUnion, continued to make new improvements to MMM, including increased automation and expanding on the types of stimuli variables which could be measured.

All these pioneer founders insisted that every variable that could possibly affect sales should be included in Marketing Mix Modeling, including price, promotion, sales force, innovations, weather, competition, creative, distribution, advertising spend allocation, brand content, brand ratings and comments in social media and word of mouth, influencer marketing, packaging, PR, recalls, *Consumer Reports*, economic conditions, wars, pandemics, and so on. Peter Drucker would

undoubtedly agree, because he was one of the first influential thinkers to include product innovation as a key component in his definition of marketing.

This is in stark contrast to *Media* Mix Modeling, which only considers the media variables. (TRA differed from most practitioners of the big data matching method of singlesource by including price and promotion data licensed from IRI.) The problem with not including all possible causals is that if some stimuli are excluded, their contribution to sales will be falsely attributed to some of the media that are being modelled, inflating their performance. What this means is that Media Mix Modeling will be less accurate than Marketing Mix Modeling, although it is generally assumed that the relationships among media types will not be too distorted. We will return to this challenge shortly.

The exclusion of some causes from some types of ROI/ROAS measurement is in fact one of the problems, besides privacy, which has led to the present resurgence of all-inclusive Marketing Mix Modeling. The most pernicious version of this "exclusion of some causes" can be seen in the reporting of ROI/ROAS by individual platforms: all the incremental sales observed are attributed to the one platform producing the report, as if no other media were running concurrently. Here's an egregious example of how different the ROAS findings were between the media platform's report (attributing all incremental sales during the period to the platform) and the report of a third party (3P) ROI/ROAS company that was reflecting all media, promotion, price, and other marketing variables:

Exhibit 2: Anonymized excerpt from agency report to client[3]

	Spend	CPM	Platform eComm ROAS	Platform ROAS	3P ROAS
Platform Targeted	$50,770	$3.27	$2.04	$33.21	$0.21
External Supplier Targeted	$163,971	$6.44	$0.14	$2.49	$0.41

3 Major research company conducted random control trial for large national retail chain; agency reported side by side the ROAS reported by the platform vs. reported by the third party research company.

IV. Causes of the Resurgence in MMM

Why is MMM enjoying a resurgence? It has more to do with a downgrading of the perceived ongoing value of more granular methods of measuring ROI by media type, such as singlesource/MTA (which depend on matching of data across databases, today seen as being threatened by privacy regulations which are still in motion). The following appear to be the main contributing factors:

1. MMM has sped up in recent years and has been found to be a more useful tool. For decades, it had been done once a year and by the time you received the data you were 18 months out of step with the times. Now being done quarterly and even more frequently by the leading players, the attention paid to MMM across companies has become more widespread and frequent.
2. Within a given advertiser organization, MMM is now generally being used across silos, creating a **shared view of reality**, which is a positive mindset change for marketers whose silos have generally not shared information of that depth.
3. There is a plausibility to these MMM findings because the general practice is to have modelers use what is called “Bayesian Priors” but which used to be called “smoothing” in the 20th Century, i.e., this year’s MMM numbers are made to approximate the prior year’s MMM numbers wherever there is statistical latitude to justify this. In the MMM process, it is common to generate multiple models that best fit the data, and then to choose the one that has

the highest scores on statistical significance and other statistical "goodness" measures, and the lowest scores on p-values and other statistical "risk" measures. However, ties are common, and this provides the epistemological latitude to apply the rule to choose the new model that trends most smoothly with the prior report. Therefore, by the time results are shared, the numbers look right to everyone in the company. This is somewhat illusory but creates a more positive working environment. *Despite* biases (e.g. media types which do not vary in GRP or percent allocation across geographic markets get smaller coefficients and appear to be less effective than they might actually be), manual/system errors, lack of tactical granularity, and reduction of everything to money or GRPs, leaving out all creative/context/psychographic variables as if they are non-existent – these missing variables being responsible for about half of all incremental sales (Nielsen NCS). These are among the remaining MMM challenge areas where innovation is to come, as we shall soon be discussing.

4. The company's own top secret data, including profitability, are one source being tapped, so there is a belief in the numbers for that reason as well.
5. In stark contrast, there has been a decline in trust of survey-based brand lift studies, because they always show positive results and are now often paid for by the media as value add incentives to buy larger schedules. These questionnaire-based studies are sometimes lumped in the minds of advertiser executives with more rigorous methodologies such as singlesource.
6. This picture was also affected by a level of disappointment in the way that MTA had turned out. The expectation was that MTA plus programmatic were going to be one system, that learned over time through machine and human learning. To do that, TV would have had to be included or there would be non-inclusion bias and the guidance of the inaccurate ROAS data would be economically harmful to decisions and outcomes. But so far MTA including TV has not been generally implemented. Since in principle there is no conceptual distance between singlesource and MTA, the future is bright for all of these methods of empirically bolstering modeling at mass sample sizes of household level data as compared with MMM.

7. Probably the catalyst to re-escalating the status of MMM was the general trend toward privacy protection and especially the deprecation of third-party cookies.

   Because of the danger of over-regulation removing many positive improvements that have be brought to marketing without privacy problems, industry associations ought to collaborate to approach Washington with the economic justification for enabling a specific set of rules based on (a) the ANA/4As/ARF CASIE Privacy Principles (b) the TRA documents in public record in Washington spelling out how TRA used the CASIE rules and its own patented methods to never know the identity of the scientific samples they were drawing, and yet they could link the same household to itself across databases. These self-regulatory privacy protection methods could be accepted as sufficient, removing the need to add further government regulations that could set back marketing effectiveness.

   In the meantime, the effect of privacy overhang has led the marketplace to prepare to not be able to address specific anonymous IDs across separate media silos, at least not conveniently nor cheaply, and that singlesource and other such methods were going to be deprecated or extinguished. This is improbable and everyone hopes that is the case, but the times do not favor optimism, and so MMM has be repositioned as a privacy-safe method that can be counted on to remain available.

8. For the past year there has be mounting evidence – mostly from Truth{set} – suggesting based on participating ID graph suppliers, that the use of available ID Graphs is only about 50% accurate. This unwelcome news only added to the attractiveness of the idea of depending more on MMM.

9. This, then, was the context that preceded the leap to +12.6% MMM revenue growth per annum for MMM (according to Marketing Science Institute), the reasons why the renaissance is happening.

V. Media Companies Begin To Use MMM

As we have discussed, the major TV networks were already using BScan data in the 1990s with the ARF and many leading CPG brands. As MMM became a thing, the networks investigated it, and came to believe that MMM was harming them by grossly underestimating their positive impacts on sales. In the average MMM model, all advertising combined gets only 7% of the credit for the sales, with around half of the sales considered “baseline” that would have been sales even if all marketing disappeared. The rest of the pie is price and promotion.

TV always thought of itself as the stimuli most responsible for the success of so many brands - the timing of whose success tracked closely with their investments in TV advertising. Most any MMM modeler will tell you that it’s quite possible that 90% of the baseline came from TV in the first place, those old TV dollars were very high ROAS at the time, but that’s water under the bridge, the advertiser says.

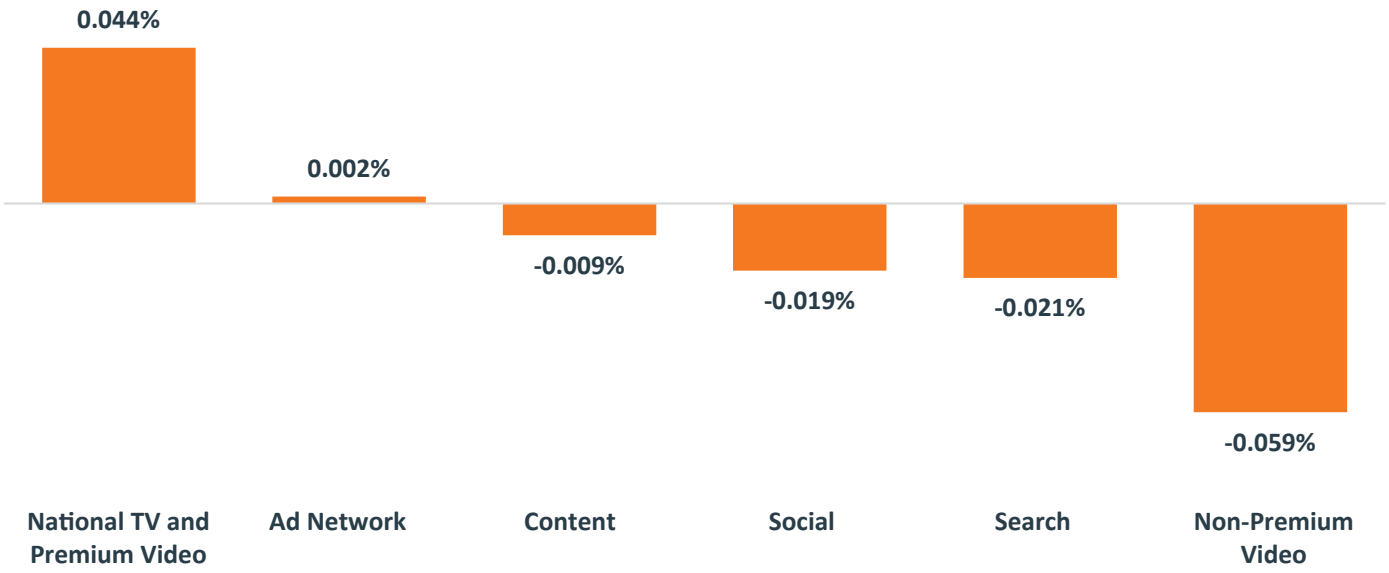
The TV networks similarly jumped on TRA and continue to be early adopters of all of the new cross-platform advanced audience and outcome measurers. Without these network investments we would not have what has evolved into “the multi-currency environment”. Attention. Resonance. Neuro.

In 2017, something changed. Historically, access to MMM data has been limited to the brands themselves and in some cases, some of their agencies. A large agency might have seen MMM data from clients across various verticals, but the sell side had not. There had not been a published set of MMM

numbers across brands and verticals freely distributed across the industry before. Bill Harvey Consulting, sponsored by FOX (and McKinsey that year), published an MMM report covering CPG, QSR and automotive, anonymized brands, all media types. FOX continues to sponsor the series and a third wave is in analysis now for delivery early 2024. The second series of reports helped cause TV network digital revenues to grow at 30% across the whole pandemic 2020-2022, as they had shown that this was the only media type to beat linear TV in ROAS (the upward impact on spending was reported both by Jack Myers in Mediavillage and separately by Standard Media Index). In the second wave, $3 trillion of sales was regressed against $48 billion in ad spend. In 2024 the third series of reports will add consumer electronics/technology.

The reports also utilize marginal utility analysis to consider media saturation effects. Charts like this one are about not the ROAS the brand achieved last year, but what an investment right now would return for you, based on saturation effects. Marginal utility analysis forecasts the future using the same statistical data base that had been used to report what the looking-backward ROAS has been in the recent past.

Exhibit 3: Table from FOX BHC ROAS Study Series[4]



That's across all car, CPG and QSR brands studied, a total of 22 brands, the biggest brands in each category. Between 2014 and 2017, they had shifted so much mass of money to digital that they had topped out the leading digital media. Smaller advertisers might never get to saturate a media type, but it happens all too often with big advertisers, and that's why MMM should always include marginal utility analysis, or some other way of reflecting saturation effects.

In this complex and fast moving marketplace, what is the optimal opportunity for the networks, with regard to ROI/ROAS? We shall address this shortly.

4 Source: BHC FOX ROAS Wave 2, U.S. 2014-2017

VI. The Latest Innovations in MMM – a selective tour

Let's take a look around us now to see who is doing what that we find exciting. These findings derive from about two dozen interviews with leading ROI/ROAS experts from major advertisers, agencies, and ROI/ROAS measurement companies.

- Arima up in Canada provides a self-directed always-on MMM which is connected to their national synthetic populations for US and Canada. This fuses the advertiser-defined audience's media behavior with the MMM's ROAS metrics. The MMM ROAS is used as a "cross media conversion pixel" for all media planning. The synthetic data are generated from aggregate sources, organized by census block and downscaled to the Virtual ID. From there, virtual households are created. Driving the use of synthetic data are the privacy regulations and consent management requirements that have existed in Canada for decades whereas US regulators are just beginning.

  Arima is a member of the species called "agent based modeling" (ABM) which is, to quote Rex Briggs, "simulated agents that attempt to fit aggregate patterns of human behavior by adjusting internal values," whereas singlesource and MTA are "human actions observed at a person or household level".

Arimadata enables user to see how the forecasted sales changes as user adds or subtracts from a stimulus type using sliders

- Michael Cohen, Chief Data and Measurement Officer of PlusCompany, a global network of full funnel creative marketing agencies, is a former marketing professor at NYU where he published several papers on econometric methods for measuring marketing impact, and taught quantitative marketing. He is a serial innovator in the ROI/ROAS field for invention of new mathematical and computational approaches for companies such as mProductivity, AOL/Yahoo, and Marketing Evolution. His completely new predictive intelligence produces all of the same types of data as MMM, MTA, agent based modeling (ABM), and singlesource. He advocates experimentation to further establish truthsets for model precision. This use of MMM + agent level + experiments is one of our main recommendations to all companies going forward. More on this below.

Michael agrees that MMM is inclusive of non-advertising stimuli beyond where media techniques like ABM, MTA, and singlesource can go, but he points out that MTA, ABM and singlesource are inclusive of important things that are not in MMM, such as creative, the non-random differences among households/people in both exposure and response to stimuli, sequence effects, frequency… suggesting that the road ahead will emphasize the combination of techniques far beyond what has been done in the past. He feels that MMM users too often fail to reflect nuances in media saturation effects, and campaign saturation effects – and ignoring the latter leads them to keep spending heavily on campaigns which have accumulated enough residual effects ("Adstock") to not need all that spending to achieve the same results. In the present environment when almost all large marketers want to reduce media spend, this is a very important learning.

- Ross Link, founder CEO of Marketing Attribution, is a pioneer in MMM and in its automation, creating systems for Nielsen and others, and today offering patented state of the art MMM, random control trials, matched panel trials, store group models, price optimization, ROI optimization incorporating saturation effects, and analyzing effective frequency. He too has long known that the shape of the long term dose-response curve of sales to advertising is a sigmoid (S-shaped) curve, and this underpins all of the work of Marketing Attribution. His results also confirm that lapsed users are one of the most profitable

targets, as found on similar targets (some purchase of brand but not loyal) by the author's work at TRA, Leslie Wood formerly at NCS, and Joel Rubinson.

Ross advises that brands should vary their plans more by market, both to reflect the higher responsiveness potential of certain markets (middle of S curve), and to create natural experiments which can be more easily read by MMM than the relatively "flat across markets" data that has resulted from the shift by national advertisers, for CPM reasons, away from spot TV and other local media investments.

He also advises that networks should themselves use MMM as part of the value add deal concept for purchase of larger schedules, where clients would be able to hide their actual data via clean rooms which enable modelers to apply their math without seeing the input data. Ross feels that a major win/win opportunity would be for networks to combine this use of MMM with clients with making it easier and more affordable for clients to try more new ideas by random control trials and other types of experiments in specific markets, which conjures up a vision of what TV networks can do with their addressable TV inventory as it grows, and in the short term especially through deals involving specific station affiliates and/or cable system carriage deals.

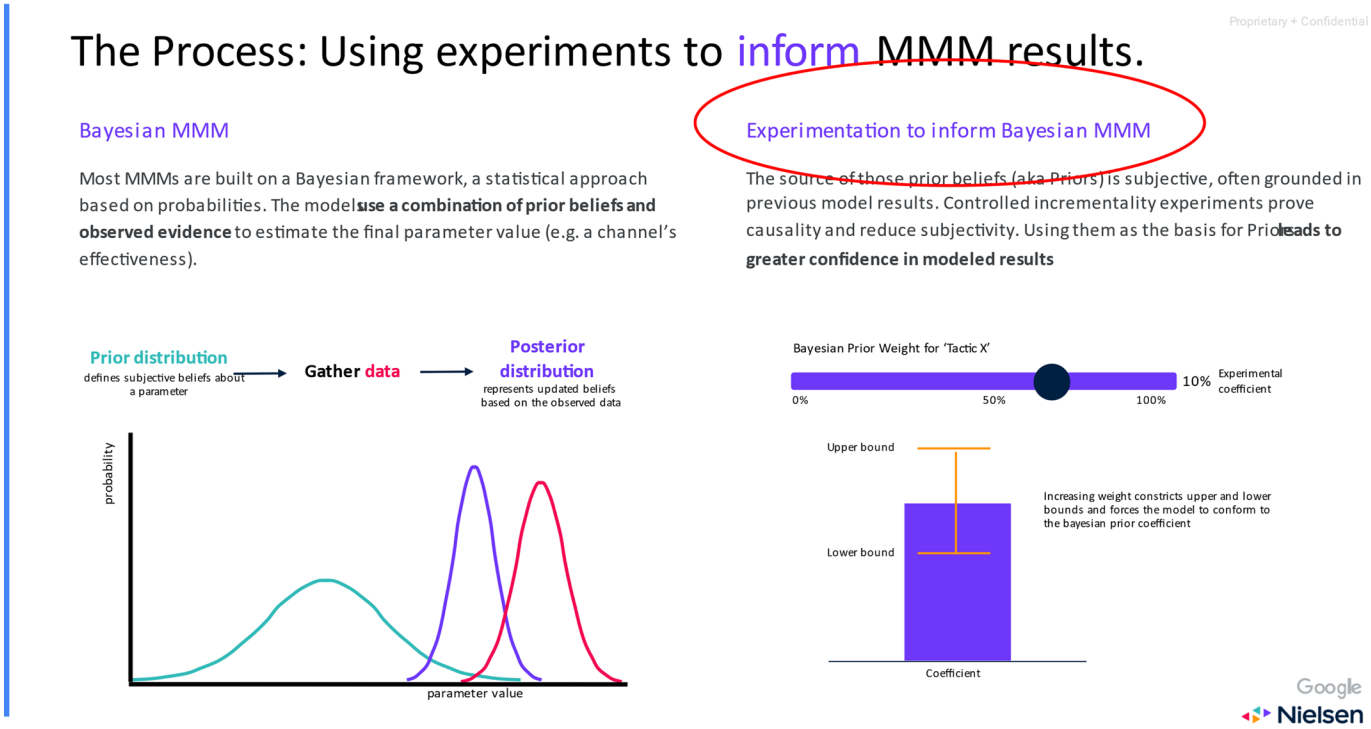
The singlesource supplier 605 (now part of iSpot) has been a pioneer in the use of random control trials based on addressable TV, and in 2020 partnered with ARF in a campaign to encourage more use of random control trials in the industry. Ross believes that the TV networks could take a more leading role in making it easier to set up RCTs by TV network-MVPD collaboration, an ideal opportunity for CIMM and VAB to catalyze and coordinate.

As Joel Rubinson says in **Section X**, "Testing [i.e. in-market experimentation] is not a gold standard without a hell of a lot of work."

This underscores the opportunity for media companies to ease the pain of buy side ability to use the scientific truth source to maximize confidence in ROI/ROAS results.

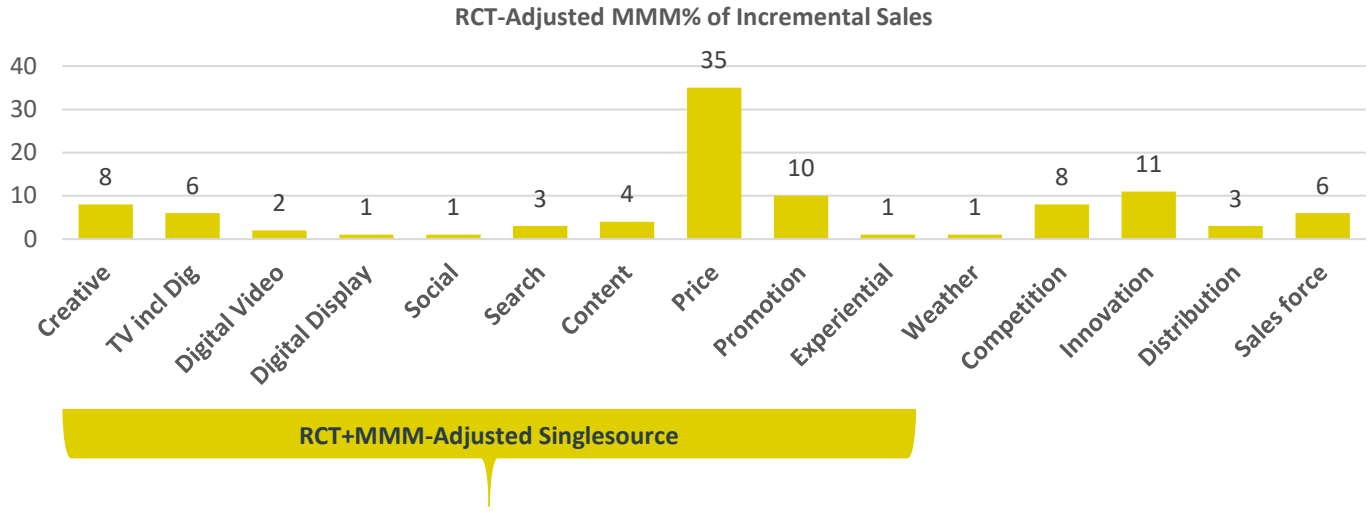
When asked about AI, Ross opined that AI could probably be leveraged the most as a means of assembling and cleaning up the raw input data, which has been a source of much noise in MMM in the past.

- One large agency holding company has created a proprietary AI MMM system which automatically tests thousands of models to determine which test model best fits the real data. The system updates itself on a daily basis. The exact techniques are proprietary. This turns MMM into a fast response ROAS methodology for the first time.
- An analytics lead from another large agency holding company, Hui Wang (see **Section X**) emphasized the importance of focusing on the way of working as an integral part of MMM deployment, the need for a human in the loop when incorporating AIML in MMM. Furthermore, given the constantly changing world, it's essential to supplement MMM with other measurement solutions.
- Nielsen has two separate divisions, one using MMM and the other, known as NCS, using singlesource. In one publicly-released study, sponsored by Google, the two methods were used together. In this way, MMM is able to cover all of the additional marketing stimuli not covered by NCS, and the media covered by both NCS and Nielsen MMM can be lined up so that the household level granularity and content (ads and contexts) level granularity can be provided by singlesource, whose aggregate results in this combined technique were integrated in the study for Google as Bayesian Priors into MMM.



Another future possibility discussed with Nielsen is the use of conforming the singlesource results to the MMM results, as in the meter/diary integration of the past.

How RCT-Grounded MMM Can Serve as a Framework in which RCT-Grounded Singlesource is Conformed to fit within MMM Frame



All creative executions within granular media subtypes add up to their respective **MMM** parent: for example, Facebook, Instagram, et al add to the **MMM** coefficient for Social Media

Possibly both techniques, in the Priors and in the final stage of conforming MMM and singlesource, will become the way of the future, with experiments to adjust models to proven causal realities only verifiable via true scientific experiments.

In **Section X** the reader will find a multifaceted interview of NCS CEO Alan Miles, delving into the bevy of directions in which ROI is evolving there.

These are but a sample of the kinds of innovations that companies are already using to make MMM and singlesource even more useful than they have always been.

VII. Future Innovation Opportunities for ROI/ROAS Measurement

During the course of our research, industry participants identified various innovations and exciting developments. As we have seen, the innovation leaders in this field have already been working on making MMM faster, delivering much more frequent results, with the data received in at least one case only a day or two after the period measured. In stark contrast, for these many decades MMM had mostly been done only once a year, each study covering the last 104 weeks, with further delays of typically several weeks before the results were reported to practitioners.

Another pattern we see is the combining of techniques, with different practitioners favoring different methods to combine with MMM.

Several experts expressed ongoing interest in MTA, suggesting that more MTA-like consumer journeys across touchpoints will probably show up in future hybrid ROI/ROAS measurement solutions. These MTA analytics will undoubtedly be cross-platform, not just digital-only as in the past.

VIII. Recommendations

These are recommendations made by the author, and in many cases respond to needs stated by experts interviewed.

1. **Add Creative To MMM.** This is a recommendation for one method as to how to achieve the inclusion of the creative factor in MMM, which was a goal expressed by a number of experts.

The NCS 5 Keys of Advertising Effectiveness Study, which is updated every few years, is a meta-analysis of all of the NCS studies done for the many clients, aimed at tracking the contribution of five main advertising forces: creative, reach, targeting, recency and context. The latest NCS 5 Keys Study shows that the creative accounts for about half of the total effect across television and digital. In classic MMM, there is no input information regarding the creative. Common sense tells us that this is a disconnect: something that accounts for half of advertising's sales effect ought not be left out when analyzing ROAS or ROI.

Some researchers, including Meg Blair in one study, and Chuck Young in another such study, have attempted to add copy test scores in as variables in MMM. The results were inconclusive.

The Cognition Council of the ARF conducted a 2022 study in which a different approach was taken, not using copy test scores, but instead using a new form of empirically-derived content coding developed by RMT using machine learning to reduce over 10,000 words and phrases to the 265 words or phrases with greatest predictive power. In the original development phase, the predictions related to

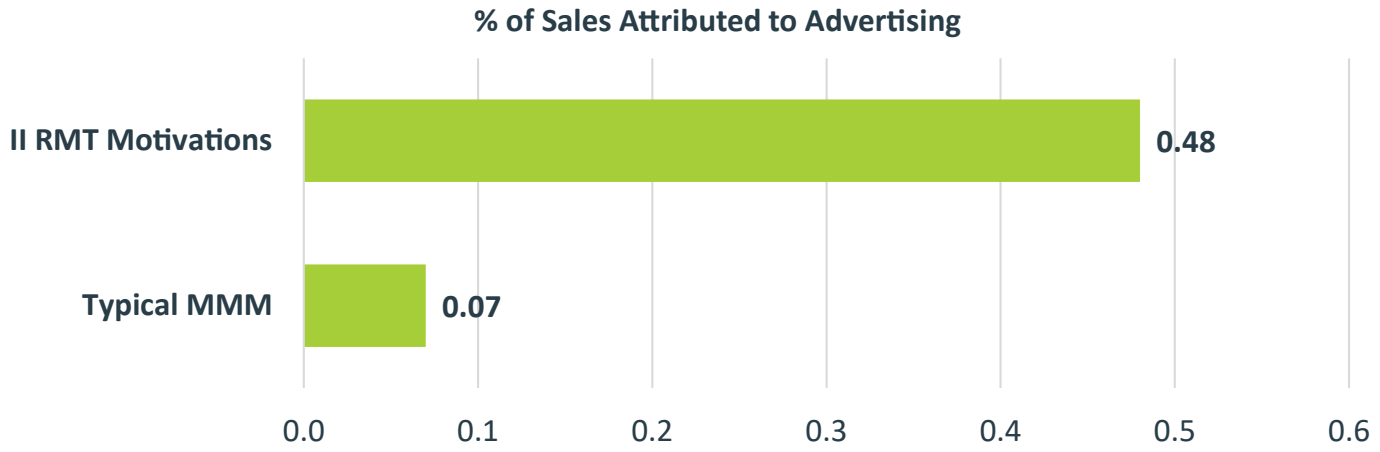
being able to predict which television series a household would be most likely to watch, with measurements taken by set top box data. The 265 DriverTags which came through as most predictive, in later validation studies were also found to be predictive of other outcomes, including Nielsen ratings, brand adoption (Simmons), the increased sales effect associated with an ad having more DriverTag overlap with a program context (NCS), the increased branding effect of the same thing (605, now part of iSpot), and the increased sales effect associated with an ad having more DriverTag overlap with a specific person being reached addressably (Neustar). In the latter case the same 265 variables used to code ad and program content were associated with individual digital IDs based upon having DriverTag coded all major advertiser used websites and apps, and then scoring people who visited based on the content they consume. In the persons level application, the 265 DriverTags were rolled up into 15 Motivational types for operational practicality. RMT global partner Semasio carried out all digital operations.

The ARF Cognition Council chose three CPG product categories and a six year period ending in 2022. Kantar supplied all of the ads used by these brands and the start and end dates when the ads were in market. IRI supplied sales data for each of the brands. 11 of the 15 Motivations were found by ARF Cognition Council staff to be statistically significant correlates of the sales trends for these brands over the six year period, and together these Motivations accounted for 48% of the variance in sales. This was startling because over past decades, all of advertising spend by media types on average was shown to only be able to predict or explain about 7% of sales. Without using any ad spend or media type data, using only the essence of what the creative was communicating, the ARF Cognition Council explained almost half of sales – just about identical with what NCS, using entirely different methods, found as the power of the creative. Of course, where we see this heading is to use all of the data that has always been used for MMM, *plus* represent the creative by one means or another.

In the slide below note the extremely low p-values, which are further signs of strong statistical relationships rather than accidental outcomes. P-values are the probability that a specific correlation might have occurred purely by chance.

Promising Method for Inclusion of Creative in MMM

- ARF Cognition Council study using content coding found that Creative accounts for 48% of brand sales
- Interesting in light of MMM typically finding media accounting for only 7%



Variable Name	Pearson Correlation Coefficient	p-value
Wealth/Success	0.521	0
Laudatory Deference	0.516	0
Status/Prestige	0.51	0
Support Group	0.471	0.001
Service To Humanity/Self Transcendence/Enlightenment/ Spiritual Awakening	0.468	0.001
Positive/Ethical Command/Boss/Ruler	0.445	0.001
Helpful	0.438	0.002
Good Role Model	0.43	0.002
Aspiration	0.428	0.002

SOURCES: ARF, CIRCANA IRI, Kantar, RMT; 2016-2021 19 brands in 3 CPG product categories

https://www.mediavillage.com/article/arf-cognition-council-analysis-finds-rmt-creative-codes-explain-48-of-sales/

In short, we can now see multiple pathways to the inclusion of creative in MMM, and practitioners will come up with their own solutions.

2. **Add competition to the variables used in MMM.** The effects of competition tend to be strong effects so leaving them out is not a good idea. It used to be unfeasible due to lack of data, however, competitive data for paid media advertising are available with greatly increased accuracy in the last ten years especially due to Standard Media Index, and with a share of credit to Pathmatics, BI Science, and others. Other marketing stimuli have some catching up to do in order that a marketer can round out the picture of what each key competitor is doing at any given point in time for inclusion in the MMM-plus system.
3. **Foment Hybrids.** Whatever MMM-plus system a company decides it wants to use, we recommend use of some form of singlesource to make highly granular results available, conformed to the MMM. This might be based on empirical singlesource as it was in the case of TRA, and is today in use by NCS, 605 and others, or it might be based on a simulated population, as in the case of Arimadata, or by new cross-platform MTA, or by new methods. Some may decide to combine multiple methods with MMM all in a single system.
4. **Adjust Models based on Consilience.** Rather than assume that MMM should always be the truth proxy, and all other approaches conformed to MMM, experiment with using AI to study patterns of agreement among multiple sources to decide on a case by case basis which ROI estimate appears to have the greatest degree of consiliency (agreement) across methods/sources. And adjust the weights of the different data sources such that MMM might not always be the truthset proxy, if a number of other methods agree in certain types of cases.
5. **Bank Account Validation.** TRA introduced the simple concept of “bank account validation” in which an advertiser compares the actual amounts of money being deposited into a brand bank account during

a specific period of time, with the MMM, singlesource, MTA and other estimates of brand sales lifts. An alternative similar method is to compare only the percent sales changes from week to week, rather than absolute dollar amounts. This is a common sense practical approach to checking the plausibility of ROAS measurements regardless of their method or source. It asks and answers the question of "could these results possibly be true?"

6. **Validate Models based on Experiments.** Make a practice of verifying important MMM-plus findings using in-market experimental design, in the form of random control trials, AB tests, holdout groups, "ghost ad" tests, matched market trials (can also be performed at finer geographic levels e.g., cable zones, zip codes, etc.), matched store panels, and other designs which enable scientific certainty about causality and not just correlation or association. Between experimental verification and bank account validation, improved decisions can be made about how to set systems with the most accurate truthset proxies. See **Section X** for Rick Bruner's contribution to easing the pain of executing Random Control Trials and other types of experiments.

7. **Adjust Models based on Experiments.** Use random control trials (RCTs) or other forms of scientific experimentation to establish adjustments to the models at both MM and singlesource levels. This slide illustrates at a high level how this might work:

Why Use More Than One ROI/ROAS Method?

The future: integration of MMM, singlesource, and experiments

MMM % of Incremental Sales

- Creative
- TV incl Dig
- Digital Video
- Digital Display
- Social
- Search
- Content
- Price
- Promotion
- Experiential
- Weather
- Competition
- Innovation
- Distribution

Singlesource

Adds granularity, branding, targeting, journey/sequence, timing, frequency, recency, "WHY" insights, tactical actionability; conforms to MMM

Experiments

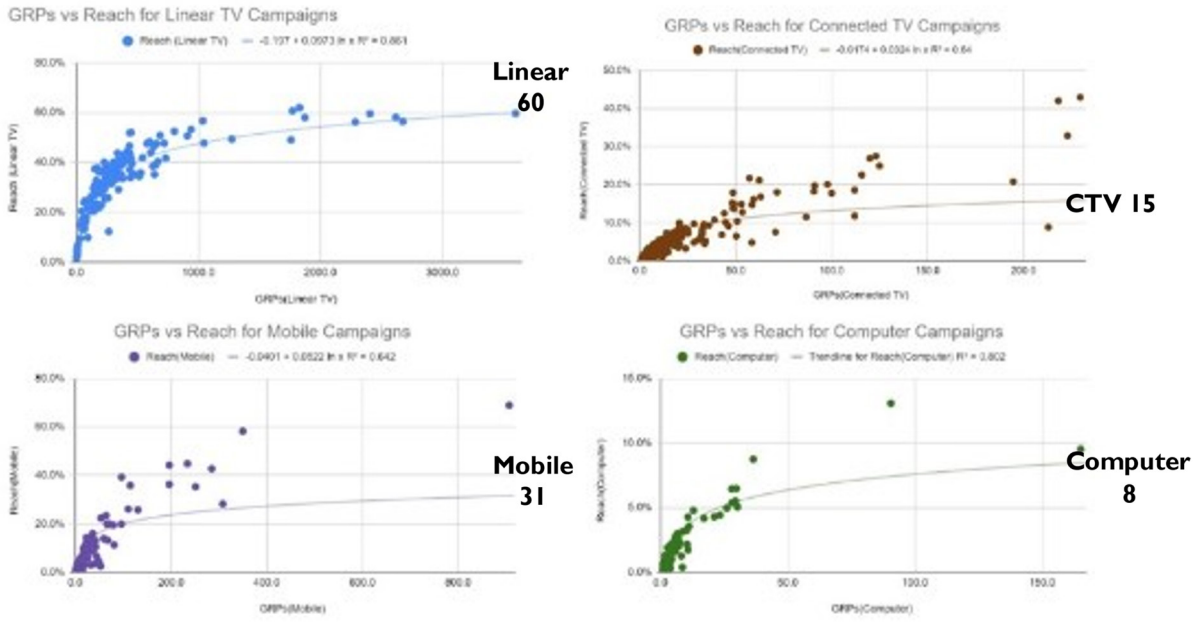
Provides greater certainty for investment decisions, may adjust MMM and Singlesource; includes Random Control Trials, Matched Market Trials, Matched Store Panels, etc.

The success of such a program will depend on making RCTs easier – TV networks and other sell side players can do this. The network cut-ins which have always been available are difficult to use and tend to be imperfectly executed and expensive. The networks could invest in making network cut-ins/holdout groups easy and their costs sometimes covered as value add if buys are large and consistent enough at the right CPMs. This will be valuable even in linear, and all the more value in addressable whether MVPD addressable or streaming – because in addressable true RCTs are possible, as 605 has proven for Walmart (presented at ARF Conference June 2018) and others.

8. **Dashboards for Top Executives.** Marketers (including some interviewed for this study) want a dashboard system through which they can manage all the marketing levers for a brand. This means being able to obtain fast feedback loops of sales and other outcomes, enabling inflight reoptimization. CMOs and brand managers want to be able to do this not just for digital but also for TV, which is another area in which TV networks could profit by leading the way. Such systems would present one view of reality across advertiser organizations and would be shared to whatever degree the advertiser wishes with their agencies. They could be interfaced through Large Language Model Chatbots using voice or text, or slider bars, and soon. We have seen above how Arimadata (and others) offer easy to use interfaces where even top executives can "play with" the sliders to see what becomes of their forecasted sales. In the past this use of sliders in this context have been focused on marketing stimuli such as specific media types. In the future there could also be sliders for how much to target various segments, how much weight to give to short term sales versus branding effects, how much weight to give to resonance vs. attention, and so on.

9. **Optimize the Future.** As marketers become more sophisticated, they gravitate away from just looking backward with ROI/ROAS models, they also project them into the future, so that seasonality, and other predictable future changes can be considered in optimization. In MMM, these forecasts have been a standard practice. In media optimizers, one of the first to use projection into the future were the media optimizers built for each of their largest clients in the 1960s by the American Research Bureau (later called Arbitron and today part of Nielsen). But then the practice of media optimizing into the future fell into disuse until being brought back to life by Howard Shimmel and Dan Aversano then at Turner and now at DatafuelX. This forward-optimization should be a bedrock feature of the new ROI/ROAS measurement and optimization systems. As we saw earlier, MMM itself can project forward (marginal utility analysis) and used that way can optimize considering the empirically measured saturation levels of each media type. These will tend to follow the reach saturation curves of different media types as illustrated here by Nielsen ONE Ads:

Nielsen ONE Ads shows Reach Saturations by Media Type



10. **Include Missing Variables.** There are a number of other variables such as branding metrics https://thetomroach.com/2020/06/22/the-greatest-hits-of-binet-amp-field/, timing, sequence, frequency, continuity, recency, synergy, and others, which should be reflected in optimizing marketing investments and tactical selections. These are comprehensively listed and considered in the ARF paper "Cross-Platform ROI Optimization Must Include Creative" https://www.rmt.solutions/crossmedia-roi/.
11. **High Performing Media Stand To Gain The Most.** TV Networks and other media who are certain that their impression is worth more should remain at the forefront in helping to support the development of new MMM-plus systems and also in making it easier for advertisers to vary their tactics by geographic areas. Clean rooms can protect the confidential advertiser information (and any confidential media company information) while enabling the necessary calculations to generate ROAS and ROI measurements. Media companies should shift research investments to proving their media type's sales and branding efficacy, using the highest quality data and methodologies possible.
12. **Use Addressable Targeting based on the 100% Accurate Broadband Internet Service Providers.** There is an opportunity for the networks and MVPDs in the recent realization that ID Graphs on average are wrong half the time. Because MVPDs not only sell cable/satellite subscription TV packages – they also sell broadband Internet access. In order to subscribe to broadband you have to disclose your accurate postal address. In the U.S. MVPDs have this for two-thirds of all U.S. households. Even if their TV subscriptions went away, they would still have this 100% addressable accuracy advantage.

The author tried through Next Century Media (which won an Emmy Award for the pioneering technology development of set top box data) and failed in the 1990s to get the **networks and MVPDs to partner in making all linear TV network inventory addressable – but now is surely the time to jump on that idea.** The juncture of linear TV plus addressable TV plus streaming TV, all coming essentially from the traditional television companies, with their higher ROAS (which more than overcomes their higher CPMs), and their 100% accurate ability to match to purchaser targeting fact-based data (i.e. not lookalikes - see ARF study https://www.rmt.solutions/advanced-audiences/), and willingness to make experiments easier, and 100% addressability if TV networks and MVPDs partner to make it so – that is the Manhattan Project to restore these players to the top of the ladder where they have earned their place.

IX. Conclusion

MMM is on the rise again and this time will evolve to incorporate AI, Machine Learning (ML), multimodal and more granular methods, and to interface with interoperable activation systems.

Everything will eventually be programmatic, always on, and fast, reoptimizing continuously from learning through systems that deliver more insight with less marketer effort.

Instead of 90% of the cost of doing this kind of work paying for manual labor, the costs of the fully automated marketing nexus once the original engineering work is paid for will mostly be for the highest quality data.

X. In-Depth Insights from Pioneering Innovators

Rex Briggs, Founder, Marketing Evolution: The Story of the First Cross-Platform Analytics, and Predictions Going Forward

ARF and IAB partnered on first round of lift studies (Scott McDonald was there in 1997) after the 1996 HotWired study (see publication in JAR from Briggs & Hollis) which then led to a desire to compare digital to TV and Magazine. Microsoft, then IAB did over 24 “cross media optimization studies.” (Greg Stuart and I published most of these in our 2006 book).

More specifically, there were three approaches: the connection to offline sales in a single-source approach (see Ford study below in “Related articles” link), via randomized control tests to measure incrementality that could be added to a mix model (see Coffee Mate Nestle study), and third, was converting digital impressions into regional daily (or weekly) impressing into a mix model (see P&G and Marketing Evolution patent). All of these were in play by 2004. Academic published articles and patents will give you a better guide. IAB integration of digital for P&G, Kraft, Nestle and J&J was 2002 and Ford was 2004. Here is a 2005 publication:

Integrated multichannel communication strategies: Evaluating the return on marketing objectives—the case of the 2004 Ford F-150 launch

R Briggs, R Krishnan, N Borin - Journal of Interactive Marketing, 2005 - Wiley Online Library

Companies have made major improvements in improving the ROI in areas such as production, logistics, and services. However, examining the productivity of marketing has long been [inclusive].

Cited by 82 Related articles

Market Share Partners was founded in 2006, and it was Hansen from UCLA who gave it the IP for their approach to mix models.

Ultimately, the important advance from digital was more granular measurement at a message level (see P&G xmos study for Olay, Ford, VW, Colgate, McDonald's, etc.) and, in some cases, measurement of the person profile & message interaction (see iCom's award video for the Warner Bros movie "Creed"). Mix models have struggled with person level and message level interactions, and I don't see that changing with the MMM structure. The bet I'm placing is on AI personalization (see Management & Business Review, page 37. https://mbrjournal.com/wp-content/uploads/2023/07/MBR-Winter-Spring-2023.pdf).

MMA Global has published results of Kroger, ADT, GM, monday.com and average lift is over 2x compared to randomized control getting same creative but without AI decisioning.

With Claritas acquiring ArtsAI last month, the people that brought you NCS (Nazzaro and team) seem to be making that bet too. if they get PRIZM to feed into real time AI personalization they will have more than person and message interaction, they will be working at the message feature level and feeding GenAI.

Rex Briggs

Rick Bruner, Founder/Chairman of Central Control: The Future of Advertising ROI: "MPE": Models Plus Experiments

Marketing ROI measurement is going through a generational transformation right now. The practice is improving all the time, and a new best practice is emerging known as Models Plus Experiments (MPE).

First, here is a recap of the major advances in advertising ROI analysis in the past 50 years:

1980s: **MMM: Marketing Mix Models**

1990s: **Clicks**

2000s: **MTA: Multi-touch Attribution**

2010s: **Quasi-Experiments**

2020s: **MPE: Models Plus Experiments**

The Gold Standard Is Not a Silver Bullet

Randomized controlled trails (RCT), equivalent to "clinical trials" for proving efficacy in medicine – where outcomes of a test group are compared against those of a control group, which, critically, was assigned by a random process before the experiment – may be "the gold standard" for measuring cause-and-effect, according to science, but according to many advertising practitioners, it is not a "silver bullet."

In part, that is due to a common misperception among advertisers that running high quality experiments to measure a campaign's true impact on sales and other outcomes is terribly difficult. I am here to tell you that is not the case.

Below I elaborate on one new method, Rolling Thunder, which most advertisers can use to set up a high-quality experiment very quickly that can run in most major advertising media (cable TV, CTV, search, out of home and so on).

Running a good experiment is certainly a far cry easier than building a complex statistical model to try to explain what is driving the best ROI in the mix, and advertisers spend time and money on those all the time.

But another criticism of experiments is legitimate, which is that it is hard to generalize from a single experiment. True enough. It's a snapshot of the effect of one campaign, at a point in time, in select media, with a given creative, a particular product offer, specific campaign targeting, and so forth. But, which of those factors mattered most to driving that lift?

To generalize, you need to do more experiments. That sounds glib from a guy who sells tools to run good experiments, but it's true.

According to the "hierarchy of evidence," the ranking of different methods for measuring causal lift, from which comes the idea that RCT is the "gold standard," the only practice that regularly outranks an RCT is a meta-analysis of the results from lots of RCT studies. Think of a large benchmark of experiments, scored by the various factors within the control of advertisers, such as ad format type, publisher partners, media channels, and so on. Such a system of analysis is easily within reach of any large advertiser (or publisher or agency) that routinely practices lots of experiments. For Bayesian analysts, these are your ideal "priors."

The shrewdest advertisers are increasingly adopting this practice, dubbed "always-on experiments."

The Best Models Are Wrong But Useful: Experiments Make Them Better

But even such an RCT benchmark doesn't take the place of a good model. As they say, all models are wrong, but some are useful. Models are good for the big picture, zooming in and out to different degrees of granularity about how the mix is understood to work. They provide scenario-planning capabilities, simple summaries for strategic planning, and other merits that won't be supplanted by practicing regular ROI experiments.

Regular experiments are, however, the missing factor in the ROI analysis for too many advertisers. Experiments enable analysts to make the models better by recalibrating assumptions in their models with better evidence.

That is what I mean by "Models Plus Experiments": honing coefficients in MMM and MTA models through the practice of always-on experimentation. (And, when I say "experiments," I specifically mean RCT.)

Rolling Thunder: An Experiment Design for All Occasions

For advertisers to adopt experiments more frequently in their measurement practice, one requirement is that the experiments be easy to implement. One option comes from many digital media companies themselves, those that offer good quality experiments, usually based on the technique known as "ghost ads," often for free. Advertisers should take advantage of those regularly.

But in order for advertisers to have measurement independence from the media companies ("grading their own homework"), and for experiments that work across many different media companies and media channels, a new approach is needed. All the more so where "simple" is a requirement.

Central Control has come up with just that, a novel method we call Rolling Thunder. We have productized this technique, but any advertiser could replicate the method once they understand it.

At its heart, it uses large geographic areas, namely DMAs, or cable zones for cable campaigns, as units in a randomized controlled experiment. There is, of course, a long history of marketers using metro areas for advertising tests, often known as "pair matching," where pairs or sets of DMAs are assigned to test and control groups. Rolling Thunder is not that.

First, those classic old DMA tests almost always involve the marketers making assumptions about which pairs of markets would make a good matches. The answer is none do. Everyone knows that. They've got different weather, competitive mixes, supply chains, etc., so comparing small groups of them is futile for certainty.

Moreover, as soon as the researcher makes assumptions about what markets should go into which groups, it's clearly not a randomized test. It is what Bill Harvey calls a "judgement-based control," which is inherently subject to biases, which is exactly what randomization is meant to control for, the hidden biases that would otherwise explain the variation between test and control.

Instead, Rolling Thunder relies on using not just a few metro areas for the test, but lots of them, all 210 nationwide in the U.S., if that's possible. Fewer is also possible, but more markets give the experiment more statistical measurement power, so the more the better.

Whereas realistically you cannot compare sales results in small sets of cities, because of natural wide variations from other factors, as noted, Rolling Thunder uses multiple baskets of dozens of cities, large and small, assigned by random algorithm. The "law of large numbers" and the randomization itself effectively controls for the noise of all those other exogenous factors, leaving only the variation in the media to explain the difference in sales, exactly the point why RCT has its elevated status in the scientific method.

The markets are assigned at random not just to two arms of the experiment (test and control) but rather into one control arm and multiple test arms. The number of test groups in the design can vary but two to five test groups is typical.

And those test groups do not run concurrently but are staggard over a period of time, as illustrated in the figure below.

Rolling Thunder Experiment Design Example

Before assigning a particular design schema, ideally the researchers can examine counts of sales (or other KPIs of interest), ordered by zip code, by week, for a long period prior to the experiment (e.g., two years' worth). From that simple data set (readily extracted from a CRM database or other systems), our pre-test "power analysis" process runs hundreds of simulations of different randomized sets of DMAs, across a variety of configurations of the parameters numbers of arms and weeks "in test" for each group. That allows us to arrive at an optimal framework (e.g., the one illustrated here) with the most statistical "power" to successfully be able measure the effect size (aka "lift") that the advertiser is expecting.

Experiment Framework: 5/4/8: Five Test Cells / Four-Week Test Flights / Eight Weeks Total

	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8
A								
B								
C								
D								
E								
Control								

The cascading feature of the multiple test groups has several advantages, including controlling for seasonality or special events during the test (e.g., Black Friday). It also gives an early-warning signal if the experiment drastically impacts sales negatively before too much media has been subjected to the test intervention.

Many Advertisers Already Have the Needed KPI Signal In-House

Reading the signal for the sales (or other KPI) of the experiment can also be managed directly by the advertiser, in many cases, eliminating external dependencies on the media company, agency or technology partners. No clean rooms, tracking pixels, user IDs or any PII required.

Many large and medium-sized advertisers have sales or other key signals in in-house CRM systems, which typically already include zip codes. Those advertisers that don't have such signal handy in-house could use this technique with various vendors of single-source signal, such as IRI, NCSolutions, Polk, foot traffic measurement firms, TV tune-in providers, brand lift vendors and more – all of whom have legitimate consumer KPI signals but generally lack true randomized control groups.

Zip Codes Will Save Advertising

That humble zip code is a key component of Rolling Thunder. It's fairly universal in customer databases. It is anonymous. And it ladders up to DMAs (and also cable zones), which, in the Rolling Thunder design, themselves ladder up to the test and control groups. Therefore, advertisers can read the lift themselves: did sales patterns increase proportionally when each test group was in-test, compared to the control group?

That is the gist. Recap of its key value propositions:

- **Scientifically sound:** true RCT
- **Media universal:** works for cable TV, CTV, digital video, programmatic, social, search and more
- **Advertiser friendly:** works for many kinds of advertisers, large and smaller
- **Extensible:** Works for various KPIs including sales, foot traffic, brand lift, TV tune-in and more
- **Independent:** Requires nothing but plan compliance from media firms, agencies or other partners
- **Low Tech:** No special tech required (no clean rooms, user IDs, cookies, etc.)
- Privacy assured: No PII, only zip codes
- **Fraud proof:** Performance cannot be gamed: we defy anyone to hack it
- **Fast results:** final report in minutes of last data upload in Central Control tool
- **Simple:** Any media partner or planner that has a problem with a DMA-level on/off schedule is not fit for the modern media

Ed Dittus, Founder of MMA: The Start of Commercialization of Aggregate Data Analysis

AdWorks was not an ARF project though Gian Fulgoni and I presented results at the ARF Infoplex in March of 1998. It was a joint project of MMA, IRI and Nielsen Media. Absolutely unique. MMA analyzed hundreds and hundreds of brands and then sliced and diced the results.

Thing to remember is that IRI had panel data and it was in their interest to use it rather than aggregate scanner data for Mix Analyses.

The first commercial application of aggregate data analysis (i.e., not a one-off academic study) was performed for Maxwell House by me about 1995.

The key to the intellectual underpinning of aggregate modeling was the realization and implementation of von Helmholtz's insights into the nature of learning (learning and decay) into a programmatic framework. That was the start of MMA.

Actually, you can go back to the work that George Williams did (with me and Jim Spaeth) at Y&R. He was hung up on "making" a panel using the old RD R&F algorithms then "fitting" a learning and decay to that population. It was horribly cumbersome and missed the essential truth that a GRP combines both reach and frequency. A simplifying assumption that created an industry.

Anyway, that is the start of the commercialization in a nutshell.

Alan Miles, CEO, NCSolutions: The Interview

Q&A with Bill Harvey and Alan Miles, CEO, NCSolutions

BILL: WHAT ARE THE MAJOR USE CASES TODAY FOR NCS, AND ANONYMOUS SUCCESS STORIES OF THOSE USE CASES?

Alan: When NCS first started in 2010, its primary solution was measuring campaign sales lift, helping CPG brands link up the ads consumers see with the products they buy. We call that service our **Sales Effect report**, and it remains one of our key solutions.

We now have a closed-loop set of services built around our Sales Effect measurement solutions and the purchase graph that sits behind it. We have **custom and off-the-shelf audience segments** and have honed our

expertise for in-flight campaign optimization. Our primary solution for brands happens right in the bid stream.

With the changes happening in consumer privacy and more and more companies bringing their data in-house, we have launched a solution named **CPG Insights Stream**. It allows our customers to do all sorts of advanced analytics solutions using NCS-proven insights. Given that our teams have been in the business of advertising effectiveness for 13 years, we have some deep knowledge of how to build the right set of insights to achieve the advanced analytics brands are seeking today.

We are particularly proud of this innovation because as we meet with marketers, we hear how expensive and time-consuming it can be to stand up a clean room and begin gaining value quickly. With NCS's **CPG Insights Stream**, marketers and data scientists can dive right in fairly quickly and get value immediately because the data is turnkey.

Alan: As for case studies, we have a whole library!

Three more recent ones are from KIND Bar, Walmart Connect and an unnamed cat food brand. These are public-- both brands are named in public-facing marketing materials. Happy to send you the links to the materials if it's useful.

A quick rundown of each.

Recently, **Kind Snacks** used our bidstream optimization solution to increase return ad spend (ROAS) during the campaign using real-time purchase insights. The result was that KIND increased its ROAS by 6% during the campaign—using real-time purchase insights for the campaign earned Kind a Shorty Award nomination.

As for **Walmart Connect**, the retail media arm of Walmart. We enable rest-of-market Sales Effect measurement for Walmart. We were recently on stage with them to discuss our work for their retail media network. It is called a REST OF MARKET MEASUREMENT.

The rest of the market measurement gives advertisers a complete view of their retail investment. Purchases happen as a result of seeing the ad on the Walmart network. These purchases could happen at Walmart, but the rest of market piece brings in purchases that resulted from the Walmart campaign being seen, but the purchase occurred at a different retailer. Walmart gains a third-party authentication of the incremental sales created from their platform to share with their advertisers.

As for audience targeting, we worked with an anonymous **cat food brand** on its data-driven linear and addressable TV campaigns. They used NCS audiences, which are purchase-based, to define the most frequent category of cat food buyers.

And yes, it worked! When they used this combo of data-driven TV, they reached over 50 million households. The heavy category buyer households that got these targeted ads spent a whopping 31% more on average compared to the brand buyers who were reached just on linear TV.

BILL: THE 5 KEYS – NCS DOES A FAR GREATER VOLUME OF ROAS STUDIES THAN THE REST OF THE FIELD AND – THEREFORE, CAN ROLL UP BENCHMARK NORMS USEFUL FOR MEDIA PLANNING – CAN YOU EXPAND UPON THAT THEME?

Alan: We have normative data; however, we do not have a formalized campaign planning solution. Clients use results from previous Sales Effect reports to plan future campaign strategies.

Our normative data comes from thousands of CPG digital and TV campaigns we have measured. It is a valuable asset for us. Our client teams use that benchmark data to help our clients improve advertising effectiveness. It also provides a way for clients to know how their campaigns are doing compared to the averages of all others.

The Norms benchmark campaign success across various industries and metrics like incremental sales, return on ad spend, and

more. In fact, many of our publisher clients have run enough campaigns with NCS measurement to develop their own normative data - which they use to share with advertisers about the results on their platform.

As for the five keys of advertising effectiveness (which we know you are quite familiar with), that is a piece of thought leadership we originally published in 2017 and conducted an updated analysis this year.

From our analysis of campaigns we measured, we can tell which aspects of a campaign contribute to sales (creative, brand, targeting, reach, recency). The most significant change is the impact of brand factors, including consumer loyalty, market share and brand penetration. This key element is responsible for 21% of incremental sales — an increase of 6% since the original meta-study.

It's also worth noting the increase in the impact of brand factors on sales comes at the expense of audience reach, which now contributes 14% to incremental sales — down 8% over the same period.

Our new analysis underscores how much the balance in the advertising ecosystem is shifting. Targeting now influences 11% of incremental sales in 2023, a slight increase over 9% in 2017, while recency impacts 5% of incremental sales, unchanged from the 2017 study. The advertising creative, or what was called creative quality in the original study, drives nearly half (49%) of incremental sales and remains the most critical driver of advertising effectiveness by a wide margin. Its impact is unchanged from 2017.

These shifts in balance mean brands will need to adjust their advertising strategies. What worked in 2017 won't be as effective in 2023. They'll need to pull different levers to get the most value from their advertising dollars.

BILL: LESLIE WOOD, WHEN SHE WAS AT NCS, TOLD ME THAT NCS USES THE NATIONAL PEOPLE METER PANEL PLUS A COUPLE OF HUNDRED THOUSAND SET-TOP BOX HOMES – THAT WAS A LONG TIME AGO – WHAT IS IT TODAY?

Alan: ~17.7MM is the number today.

BILL: WHAT ARE THE REASONS TO BELIEVE NCS IS MORE ACCURATE THAN ANY OTHER ROAS METHOD?

Alan:

1. **NCS Leads the Industry:** We're at the forefront with an extensive array of big data sources and more integrations than any other player in the field.
2. **Extensive Data:** Our data pool consists of over 2 trillion CPG retail purchases spanning over 100 million households. This vast dataset, combined with our proprietary models, allows us to accurately project the CPG purchasing for the entire landscape of US households and retail outlets.
3. **Pioneering AI and Machine Learning:** While AI and machine learning have been making headlines as groundbreaking technologies, NCS has been harnessing their power for almost ten years.
4. **Continuous Innovation:** We are deeply committed to innovation and continually refine our models and methods. This constant fine-tuning of our approach has enabled us to stay ahead of the curve and innovate consistently.

As for ROAS, initially, we launched a highly successful ROAS platform based on an ANCOVA methodology. However, recognizing the potential for improvement through AI and machine learning, we launched our Next Gen platform several years ago. This next-generation platform, driven by AI and machine learning, provides more precise, detailed and consistent insights into campaign effectiveness, all delivered faster.

BILL: CALIBRATED WITH NIELSEN MMM?

Alan: NCS is not calibrated with MMM studies. Instead, think of it as an input to MMM studies.

MMM is gaining popularity as MTA solutions become obsolete due to the evolution of consumer privacy regulations. There is also increased budget pressure, driving a desire to optimize and understand the long and short-term impacts of advertising that MMM can provide.

Marketing Mix Modeling is great at understanding how advertising works at scale but is traditionally less effective when the goal is to understand the precise contribution of incremental sales at a publisher or channel level.

We did an experiment with Sales Effect inputs that began because Google wanted to understand better how individual campaigns contributed to incremental sales at scale. And Sales Effect is considered the gold standard for understanding the causality of a campaign. Also, they wanted to look at how YouTube performed by leveraging more recent results from Sales Effect measurement. Google wanted the best of both worlds - the breadth of MMM and the precision of causal campaign sales impact.

The study details are that Google teamed up with us to analyze 10 different CPG campaigns that ran on YouTube that overlapped between Nielsen MMM and NCS Sales Effect. Upon re-running all 10 models with new sales lift outcomes included, Nielsen saw an 84% increase in YouTube ROAS. For the average brand, this translated to $570K more in YouTube-attributed sales. You can also watch the full ARF presentation for more details. Incidentally, only 30% of this YouTube increase cut into other publishers.

The outcome is that NCS Sales Effect studies can be a meaningful input to multi-mix marketing (MMM) studies for brands to increase the precision of individual partners or channels. We have had a lot of client inquiries about how this might work.

BILL: WHAT IS NEW AT NCS, AND WHERE DOES NCS GO NEXT?

Alan: We are focused on three key areas to support the industry and our mission of improving advertising effectiveness for all media.

1. **Leveraging AI, ML and our Next Gen Platform further.**
   - Cross Media, RMN strategies
   - Feature ups for current products with constant innovation.
   - Partnerships to bring insights together for a more comprehensive understanding of advertising's impact.
2. **Bridging the gap between long-term brand growth and short-term sales impacts.**
   - Understanding the long-term value of advertising
   - Brand loyalty
   - MMM studies
3. **Supporting in-housing of data.**
   - Clean Room - Brands and publishers can access our purchase insights via a clean room. We have relationships with SnowFlake and LiveRamp. We offer much more than a data feed. We provide turnkey purchase insights designed for immediate use without requiring a data scientist to unravel unruly data sets.
   - CPG Insights Stream and applications
   - Helping brands see ROI on their clean room investment
   - Growing partner collaborations for accessibility

BILL: ADVERTISERS AND AGENCIES ARE TELLING ME THAT THEY WANT TO BE ABLE TO REOPTIMIZE TV IN-FLIGHT, JUST LIKE THEY DO WITH DIGITAL. HOW DOES NCS HELP THEM DO THAT NOW, OR WHEN IS IT COMING?

Alan: It's possible to do it now. We can help agencies, advertisers, and media companies reoptimize TV inflight using real-time purchase insights. Network, daypart, program, genre, and time period are all potential optimized variables on TV. NCS audience segments are matched to the Nielsen People Meter (NPM) and continually updated in Nielsen Tools (e.g., NPOWER, etc.).

Most clients buy TV based on historical trends and re-optimize a future campaign after the current campaign is complete. Other clients are willing to optimize, but they have technical limitations right now. This is a progression, and we are ready to support this application as the industry evolves.

We do have clients optimize their programmatic campaigns in the bidstream with their DSP.

Another opportunity for in-campaign improvement is to access our real-time purchase insights in a cleanroom for campaign optimization.

BILL: ADVANCED AUDIENCES FOR MAKING OPTIMAL MEDIA SELECTION DECISIONS BEFORE THE BUY IS MADE?

Alan: NCS does have advanced audience capabilities and we already work on custom audiences for our clients that are inclusive of first-party data. As a reminder, we don't have a media planning tool.

BILL: CAN NCS DO STUDIES WHERE THE CLIENT HAS HIS/HER OWN SALES DATA AT THE HOUSEHOLD LEVEL E.G. AMAZON, CITIBANK, GEICO?

Alan: Yes, we can integrate client data into our Sales Effect studies within the CPG sector. Beyond CPG, we're confident that our data modeling and intellectual property (IP) can effectively generate valuable insights for various industries.

BILL: NCS INTEROPERABLE WITH NIELSEN ONE ADS – WHEN?

Alan: This makes so much sense since Nielsen Media is a majority owner of NCS. NCS and Nielsen are actively working on how NCS metrics might integrate within the Nielsen ONE Ads platform. While we don't have an official launch date, we expect this to happen in the near future.

Joel Rubinson, President, Rubinson Partners: A Constructive Critique

- I am bullish on MTA; we could not have done Brand as Performance research without it. MMM would have fallen way short.
- Testing is not a gold standard without a hell of a lot of work.
- Agent based approaches have to prove to me that they capture certain interactions that are at the heart of Movable Middle theory before I could accept them. Is there a relationship between brand probability of choice and response to media? We used agent based modeling in the original Movable Middle work but they were given rules that followed beta distributions and a lot of media response stuff. Does the ABM replicate purchasing patterns described by beta and NBD distributions? If not, the model is woefully inadequate.
- Bayesian. I am not sure how I feel about Bayesian regression models to tackle multicollinearity and parameter stability issues. Whenever you use priors you are purposely introducing biased estimators.
- I wish there was more on the power of targeting the Movable Middle. It is better than ROI measurement. It CAUSES ROI to happen and isn't that what marketers really want?
- Finally, I have created something that we call Franken-modeling to integrate disparate signals of ad effectiveness into a unified model. It does NOT use ABM. Think of it as a logistic function like MTA uses without needing millions of data points to estimate the parameters. It is based on Taylor power approximations that use Jacobian and Hessan matrices which linearize the estimation and integration problem. It is not too dissimilar from Euler's method that was used to get John Glenn back to earth.

Hui Wang, Global Data Intelligence, Analytics Service at Publicis Media: Best Practices Today and Tomorrow

Way of working: Publicis has provided consultation to clients regarding MMM deployment. Some advertisers struggle to see the benefits of their MMM due to a lack of organizational alignment. When stakeholders within an organization have differing expectations, decision-making can become limited or non-existent, ultimately weakening the overall impact of the MMM.

In order to ensure the success of MMM projects, it is crucial to align key stakeholders on the learning agenda set for the project. It is essential to examine whether the MMM providers have a good understanding of media buying and can comprehend the implications of new metrics and data available from media tech platforms. It is also important to develop a process that addresses challenges in data wrangling, as well as to consider the change management aspect of the process.

In my opinion, the way of working is just as important, if not more so, than the math, AI, or automated dashboards that deliver impactful MMM insights for business decisions.

AIML: AIML can significantly enhance the efficiency of data processing, reduce data wrangling, and automate and streamline model development processes. However, in-depth knowledge of campaign or media execution is usually necessary to identify anomalies and apply the appropriate fixes. While it may sound ideal to use AIML to select the best model, there is a risk of overfitting without proper validation resulting in a model that fits the data well but lacks robustness. Furthermore, a model that fails to reflect and explain the nuances of media dynamics could lead to inaccurate results. Maybe someday, with the advancement of Gen AI, things will change, but at this stage, we still need a human in the loop - someone who understands the media, the math, and can translate the results to drive decision-making.

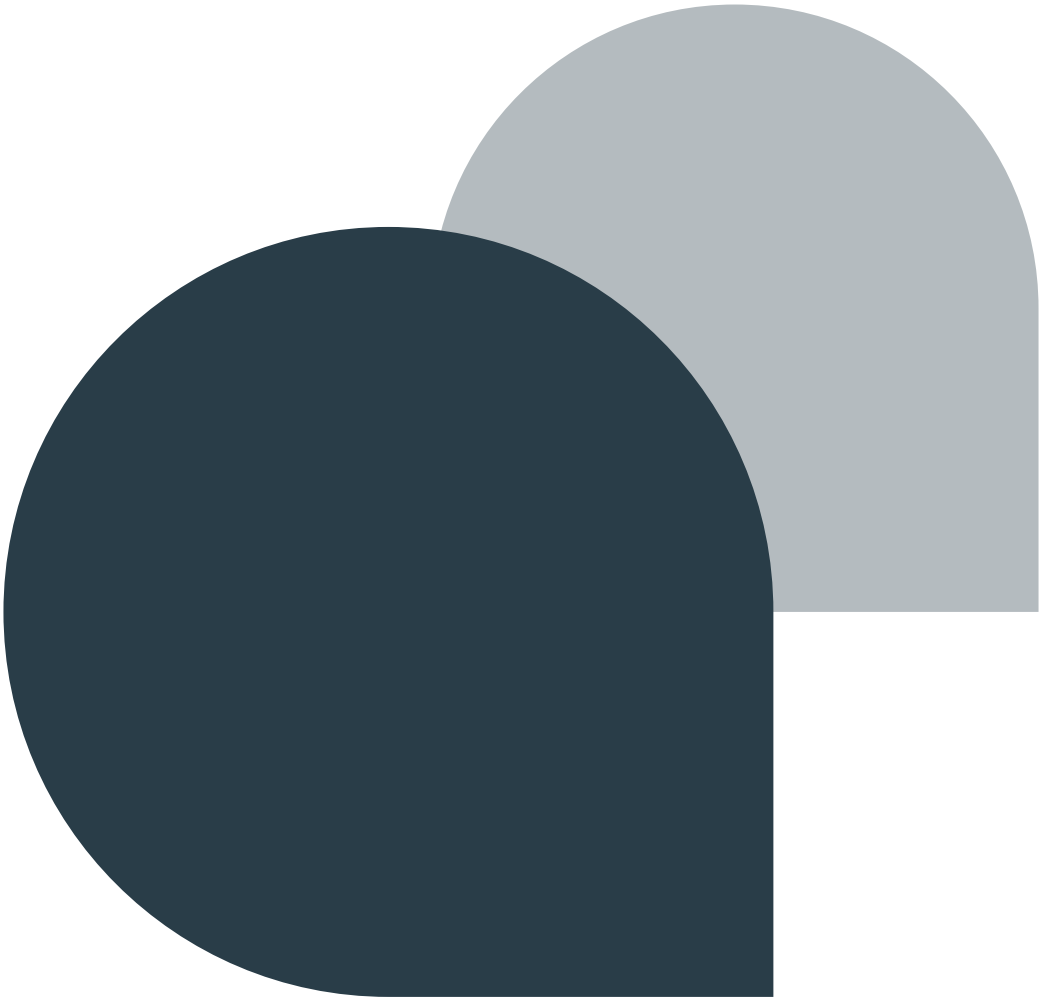
On other methodologies:

MMM provides top-down strategic recommendations for budget allocation and helps determine the optimal channel mix. Experimental Lift offers always-on validations to assess the incrementality of multiple tactical factors. Meanwhile, MTA serves as the foundation for tactical optimization, enabling marketers to continuously improve their campaign performance.

MTA has faced many challenges in recent years due to cookie deprecation. There are a few things advertisers are doing to future-proof. Some companies lean on attributions from marketing/media platforms. Some invest in clean rooms and first-party data, and some are deploying regression-based modeling that provides more granular tactical insights. Each with its own limitations and should be assessed and prioritized based on the needs of the organizations.

Publicis agencies have been increasingly tasked with deploying Experimental Lift for clients. Setting up experimentation in a truly Randomized Control Trial manner may take a lot of work, especially when clients want to measure both digital and offline sales. When RCT is not feasible, we apply quasi /natural experimentation to measure incrementality. In a world of constant change, we need methodologies that provide agility - easier to implement but still maintain some statistical rigor.

Recently, our analytics teams, especially those in Europe, have started working with clients on Agent-Based Modeling (ABM) in conjunction with MMM for portfolio management and simulating the impact of marketing activities on brand perceptions. There is a growing desire and need to educate clients on how MMM and ABM complement each other.



https://cimm-us.org/