Measuring digital advertising in a post-cookie era: A study of marketing-mix models, attribution and incrementality

Received (in revised form): 23rd July, 2024

Shashank Hosahally

SHASHANK HOSAHALLY

Business Engagement Manager and Doctoral Researcher, Birmingham City University, UK

Shashank Hosahally is a business engagement manager and doctoral researcher at Birmingham City University, developing knowledge transfer partnerships within the Faculty of Computing, Engineering and the Built Environment. He is an experienced professional in consulting technology firms with extensive knowledge of the global digital media landscape. He specialises in the field of marketing attribution, digital media measurement, decision-making frameworks, press release distribution, news media coverage, and incrementality-based media spending optimisation.

Birmingham City University, Birmingham, B4 7XG, UK
Tel: +44 7405 251583; E-mail: shashank.hosahally@bcu.ac.uk

Madan Bharadwaj

MADAN BHARADWAJ

Chief Executive Officer, M^2, USA

Madan Bharadwaj is a leading expert in advanced marketing measurement systems. Over 15 years and four attribution startups, he has brought innovations like multi-touch attribution and incrementality testing to market, helping marketers navigate away from last-click attribution. Most recently, Madan founded M^2 to shape a methodology-agnostic framework to solve for attribution. He teaches the framework in the industry's first advanced attribution masterclass and helps brands access high-quality attribution talent and services through the M^2 attribution services marketplace. Prior to M^2, Madan co-founded Measured, pioneering industrial-scale incrementality testing for marketing measurement. Madan has also held senior product and technology roles at various ad tech and loyalty tech companies, driving key innovations in algorithmic cross-channel programmatic buying and digital promotions technology, and working with both large and emerging brands across various sectors. Madan holds a master's degree in computer engineering from University of Central Florida and a bachelor's degree in electronics and communication engineering from University of Madras.

M^2, 25 Olde Village Drive, Winchester, Boston, MA 01890, USA
E-mail: madan@msquared.club

Arkadiusz Zaremba

ARKADIUSZ ZAREMBA

Lecturer, University of Warsaw, Poland

Arkadiusz Zaremba is a lecturer at University of Warsaw and Managing Director of Otomoto PAY. He has 12 years' experience in the digital advertising industry, most notably in the field of media buying. He holds a PhD in management from the University of Warsaw, having defended his thesis, 'Customer journey in the online environment', in 2020. His research interests focus on marketing attribution and its theoretical and practical gap.

University of Warsaw, Krakowskie Przedmieście 26/28, 00-927 Warszawa, Poland
Tel: +48 668 366 735; E-mail: azaremba@wz.uw.edu.pl

Olena Volkova

OLENA VOLKOVA
Chief Operating Officer, Research.ua, UK

Olena Volkova is a UX researcher and Chief Operating Officer at Research.ua LLC. Her research specialises in both qualitative and quantitative methods across banking, telecom, retail and the hospitality, restaurant and catering sectors. She has 20 years of managerial experience, having previously co-owned a research and development and manufacturing company. Olena holds an MBA and a master's degree in chemistry.

54, Velyka Vasylkivska str., Kyiv, 02000, Ukraine
E-mail: olena.volkova@research.ua

Abstract

The implementation of the EU General Data Protection Regulation (GDPR) poses significant challenges to the measurement of advertisement performance, necessitating a shift away from traditional tracking methods such as web beacons, cookies and browser fingerprinting. With the discontinuation of third-party cookies and increased privacy standards, it has become harder for marketers to optimise advertising spend and attribution models. This paper explores the GDPR's impact on the ad tech industry and anticipates challenges in adapting to a cookieless and more regulated data environment. In the wake of iPhone Operating System (iOS) updates and app tracking transparency, direct-to-consumer (D2C) companies have witnessed shifts in advertising spending across different channels. The study investigates three main measurement models: marketing-mix modelling (MMM), multi-touch attribution (MTA) and incrementality for D2C marketers, highlighting incrementality as the most effective method for analysing advertisement impact and optimising spending. The key contributions include a proposed triangulation framework that combines data from MMM, MTA and incrementality to support a data-driven approach, offering insights for tactical and strategic decision-making. To validate the proposed framework, a mixed-methods approach involving qualitative and quantitative surveys is designed. Targeting experienced advertising professionals, the survey evaluates the implementation of MMM and incrementality, assessing the various decision-making attributes of measurement models, such as ease-of-use, accuracy, validation, robustness, predictiveness etc. Results align with existing literature and the proposed framework, demonstrating the efficiency of each technique. The paper recommends adoption of the incrementality randomised control trial method and provides a roadmap for further research in this evolving landscape.

Keywords

marketing attribution, incrementality, MTA, marketing-mix modelling

DOI: 10.69554/WSAT6359

INTRODUCTION

Following the implementation of the EU General Data Protection Regulation (GDPR), the media measurement industry has faced significant challenges due to heightened privacy standards and restrictions on data sharing.[1,2] The digital marketing landscape has long relied on web beacons, pixel tracking, browser fingerprinting and cookies to track user activity and enable marketing attribution models. However, this paradigm is shifting due to changes such as the cessation of support for third-party cookies by some browsers.[3,4] With the iOS 14.5 release, apps now require user consent for tracking, further diminishing the availability of user-level data and impacting advertisement targeting precision and performance tracking, known as app tracking transparency.[5] Researchers and practitioners are analysing the GDPR's impact on the ad tech industry[6–8] and predicting challenges in adapting to a cookieless and more regulated data environment. With the decline in third-party cookie data, access to first-party data, previously held by major companies, has become increasingly important, but there are still limitations in tracking the entire customer journey at a granular level, as is necessary for attribution models.[9]

First-party trackers, which represent direct user interactions with websites or applications, are gaining prominence amid growing concerns over customer privacy.[10] Rather than relying solely on platform analytics, many medium-sized brands are asserting control by collecting first-party data through in-house pixel-tracking on all owned media points.[11] However, challenges persist in tracking the complete customer journey, as evidenced by reduced data collection post GDPR, hindering online businesses' ability to optimise marketing spending through attribution models.[12]

This paper explores alternative approaches to attribution, given the limitations of traditional models, to aid performance marketers in optimising advertising spending across multiple channels. Additionally, it underscores the importance of existing models in meeting decision-making criteria through a survey of industry practitioners.

Direct-to-consumer (D2C) companies are those that bypass intermediaries like retailers or wholesalers, opting to sell their products directly to consumers through channels such as online platforms, social media or company-owned stores. They rely on digital marketing, e-commerce platforms and innovative branding to establish direct connections with customers. According to a report by eMarketer,[13] D2C e-commerce sales in the USA are projected to experience double-digit growth, reaching nearly US$213bn by 2024, constituting 16.6 per cent of all e-commerce sales (see Figure 1).

Following iOS updates, D2C customers have increased spending on Google and offline channels while reducing spending on Facebook, significantly impacting both the e-commerce and D2C industries. In 2021, D2C e-commerce sales were valued at US$129.31bn, with a growth rate of 15.9 per cent, and this trend is expected to continue as more e-commerce brands adopt the D2C model. With the evolving media landscape, performance marketers who heavily relied on conversion attribution must now adapt their strategies to drive profitable and incremental growth. The central question facing D2C marketers is how best to achieve this goal in the face of changing market dynamics.

This study examines three main measurement models for D2C marketers in the post-GDPR world, namely, marketing-mix modelling (MMM), multi-touch attribution (MTA) and incrementality, providing key insights to guide the measurement of advertising performance. Notably, it finds incrementality to be the most effective way to analyse advertisement impact and optimise advertising spending.

The key contribution of the paper can be summarised as follows:

- A triangulation framework is proposed to overcome the inherent challenges posed by existing methods. This technique combines a benchmark dataset (last-touch attribution, LTA), covers aggregated data sets (MMM), and suggests

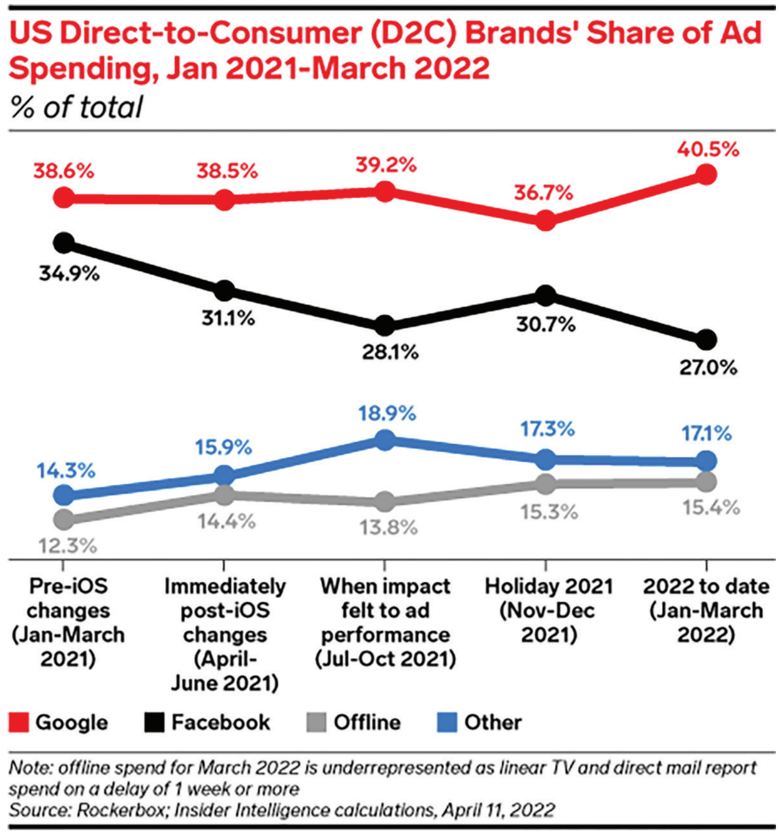

Figure 1 D2C ad spending shifting pre and post iOS
Source: eMarketer (2020) 'How D2C Retail Brands Are Evolving', 7th October, available at https://www.emarketer.com/content/how-d2c-retail-brands-evolving-5-charts (accessed 20th September, 2024)

validation using incrementality randomised control trials (RCTs). It aids ad marketers in both tactical and strategic decision-making.

- To validate the argument, a mixed approach of qualitative and quantitative survey questionnaires is designed. Targeting highly experienced advertising professionals, the survey evaluates ease of use, challenges and the advantages and disadvantages of different models. The results align with existing literature and the proposed framework, demonstrating the efficiency of each technique.
- Extensive analysis of collected data recommends the utilisation of the incrementality RCT method. Challenges and opportunities drawn from this analysis provide a roadmap for future research to build on this study.

The next section of this paper provides the background and discusses the related works, before moving on to a commentary on how media measurement has evolved post GDPR. The paper then describes the research methodology and the results.

BACKGROUND AND RELATED WORKS

Table 1 presents a summary of key studies in the field of online advertising. Rigorous causal inferences can be obtained through randomised experiments, which are often

Table 1: Key studies in the field of online advertising.

Experiment type	*Description*	*Reference(s)*	*Pros*	*Cons*
Randomised controlled trials	Utilises randomisation to assign participants or regions to experimental and control groups, enabling rigorous causal inferences.	Barajas *et al.* (2016), Blake *et al.* (2015), Johnson *et al.* (2017), Lewis and Reiley (2008), Sahni (2016)	• Establishes causal relationships • Allows for precise control over variables	• Can be costly and time-consuming to conduct • Ethical concerns regarding randomisation and control group assignment
Geo experiments	Conducted using geographic regions to assess the additional impact of advertising; often involves geo-based linear regression models or Bayesian structural time series analysis.	Vaver and Koehler (2011), Brodersen *et al.* (2015)	• Enables localised impact assessment • Utilises existing geographic data • Can be applied across various media	• Requires high number of regions for statistical power • May overlook regional differences not captured by geographic data
Ghost ads	Involves the creation of 'ghost' ads to measure the effectiveness of advertising efforts; commonly used in online display ad experiments.	Johnson *et al.* (2017)	• Provides insight into ad performance • Allows for controlled experimentation	• Can be resource-intensive to implement and manage • May encounter ethical concerns regarding deceptive practices
A/B testing	Compares two versions of a webpage or app against each other to determine which performs better; widely used in digital marketing for testing ad creatives and landing pages.	Barajas *et al.* (2016)	• Provides direct comparison between variations • Easy to implement	• Results may be influenced by external factors • Limited to comparison of only two variations
Time-based regression (TBR)	Employs regression analysis over time to assess the impact of advertising; suitable for experiments with a minimal number of geographic regions.	Kerman *et al.* (2017)	• Flexible and applicable for experiments with few regions • Captures temporal trends	• May not capture localised effects as effectively as Geo experiments • Relies on accurate time-series data
Matched-market tests	Involves careful selection and pairing of geographic regions instead of random assignment; cost-effective but may lack the accuracy of randomised trials.	Barajas *et al.* (2020), Gordon *et al.* (2019)	• Cost-effective and commercially viable • Allows for comparisons between similar regions	• May introduce bias in region selection and pairing • Results may not be as robust as those from randomised trials

conducted through geo experiments[14–16] 'ghost ads'[17] and A/B testing.[18] Notably, various studies have contributed significantly to the field of ad measurement through incrementality field experiments utilising RCTs.[19–24]

Geo-based linear regression models, as proposed by Vaver and Koehler,[25] can calculate the incremental return on advertising spend (ROAS) by assessing the additional impact of advertising within specific geographic regions. Geo-based linear regression models for incremental ROAS analysis enable businesses to assess the localised impact of advertising efforts, allowing for more targeted and efficient allocation of resources across different geographical areas. These experiments require a high number of geos to gain statistical power, which is not possible in practice as the geos are matched and paired in current geo-based market tests.[26] The causal Bayesian structural time series-based analysis methodology can be applied with a limited number of geos. The paper by Brodersen *et al.* proposes the inference of causal impact on the basis of a diffusion-regression state-space model that predicts the counterfactual market response in a synthetic control that would have occurred had no intervention taken place.

Time-based regression (TBR), developed by Kerman *et al.*,[27] is widely adopted in the industry. TBR is flexible as it is suitable for analysing experiments with a very low number of geos, such as those arising from experiments in smaller national markets and matched market studies.

In another popular matched-market test adopted in the industry,[28,29] the geos are carefully selected and paired rather than matched and randomised to the test and control groups. These tests are popular as they are inexpensive and commercially viable as they use small subsets of matched city (geos), although they may not provide the same level of accuracy as randomised trials with respect to hidden biases. Johnson[30] highlights recent developments and builds on earlier research primarily addressing the challenges, best practices and new developments in the context of studies of online display advertising. Johnson and Lewis[31] term this notion 'cost per incremental action' (CPIA), which practitioners are beginning to adopt. The authors argue that CPIA as a pay model solves many incentive problems in the advertising industry, in addition to optimising for the advertiser's desired metric.

Gordon *et al.*[32] provide a guide to measure the effectiveness of display advertising, with a focus on challenges such as low statistical power, treatment decisions guided by algorithms and marketplaces, identity fragmentation and incrementality optimisation. To improve inference from geo-based experiments, two broad solutions have been proposed: (1) pretesting outcome data to identify baseline differences using difference-in-differences [33] or synthetic approaches; and (2) applying block randomisation by grouping similar regions using past outcomes or regional characteristics (eg match market tests pair regions together).[34] Kerman *et al.*[35] and Vaver and Koehler[36] discuss design considerations for geo-based experiments regarding how to narrowly define regions, test and pretest period lengths and advertising spending differences across groups.

Geo experiments can be applied to measure a variety of user behaviour and can be used with any advertising medium that allows for geo-targeted advertising. In the USA, many advertising platforms use the 210 designated market areas (DMAs) as geo-targeting units.[37] Researchers have used Google's DMAs for various studies into the effectiveness of advertising; for example, Blake *et al.*[38] investigated paid search effectiveness. Gupta and Chokshi[39] introduce incremental lift as a metric to measure the impact of a marketing strategy. They use the viewability lift method for digital marketing strategy planning and campaign optimisations leading to improved campaign efficiency.

THE EVOLUTION OF MEDIA MEASUREMENT POST GDPR

Since the promulgation of the GDPR and similar legislation pertaining to data privacy, the marketing attribution and media measurement sector has been significantly impacted in a number of ways. These include the following:

- *Fragmented data sources:* The GDPR restricts the sharing of data between advertising exchanges and platforms. This fragmentation creates siloed data, making it difficult to obtain a comprehensive view of consumer behaviour across different channels. There is no single source of truth data for identifying campaign effectiveness at the channel or tactic level.
- *Reduced granularity:* Restrictions on tracking and collecting user-level data, especially without explicit consent, have reduced the granularity of available data. Marketers and advertisers now have less access to detailed user information, affecting their ability to track and target specific audiences effectively. Once the tracking of third-party cookies is retired, for example, it will be much harder for advertisers to target repeat visitors, as the identities of these users will no longer be unique, hence the distinction between what constitutes 'unique visitor' and what does not will become challenging, leading to a decreased understanding of user uniqueness. The traditional method of reporting impressions will continue, but at a more aggregated level, lacking the detailed insights marketers prefer. Moreover, advertisers will face limitations in attributing user actions, such as clicks and purchases on other sites, resulting in potential challenges in measuring and attributing return on investment.[40]
- *Attribution challenges*: In isolation, first-party data provide limited cross-channel visibility, increasing the risk of potential inaccuracies, as siloed information and data-quality issues hinder a comprehensive understanding of the customer journey. In a market with multiple content intermediaries, regulations designed to protect user privacy impose limitations on data collection that further impact the granularity of available information. At the same time, third-party data are associated with challenges related to reliability, accuracy and compliance with data-privacy regulations. Inaccurate or outdated external data can lead to flawed attribution models, and difficulties in cross-device tracking may introduce gaps in understanding user behaviour. Integrating third-party data seamlessly with first-party data also poses challenges, affecting the overall accuracy and effectiveness of attribution models. Addressing these challenges requires a focus on data quality, privacy compliance, integration and adapting to the evolving landscape of user behaviour and regulations.[41] More importantly, how can existing models support marketers to measure advertising effectiveness at the campaign level and take decisions to optimise advertising spend.
- *Impact on analytics and insights*: The ability to derive meaningful analytics and insights from fragmented and limited data is hampered. Marketers struggle to gain a holistic understanding of campaign performance and audience behaviour. For instance, the MMM provided by platforms can provide insights into overall channel performance, but performance marketers who require more granular data for tactical decision-making at the campaign level face challenges. This lack of detailed visibility makes it harder to

optimise individual campaigns and make data-driven adjustments in real time.

- *Compliance and transparency*: Advertising exchanges, advertising revenue models and measurement platforms have had to adapt to GDPR regulations by ensuring compliance and transparency in data collection and usage. The effort of delivering such compliance can create barriers to the seamless sharing of data. Numerous advertising buyers are hesitant to take risks with data-intensive audience-targeting methodologies. Instead, the prospect of employing contextual targeting has recently gained momentum due to its perceived enhanced safety and attractiveness. In this approach, advertisements are tailored to individuals based on the contextual content of the webpage they are currently viewing.[42] Advertising exchanges may face limitations in enriching their targeting capabilities, leading to a potential decrease in campaign effectiveness. Advertising revenue model has a significant impact on cross-site tracking without third-party cookie tracking, which is fundamental to cost-per-click and cost-per-action models. Advertisers using these models need to adapt by exploring alternative attribution approaches, relying on first-party data and embracing contextual targeting strategies, for example, retargeting revenue. While cost per thousand impressions is less directly affected, advertisers may still need to adjust their targeting and measurement approaches in a cookieless environment.[43]
- *Innovation and technology:* GDPR has pushed the industry to innovate in privacy-centric measurement technologies. Companies are exploring new methods like federated learning or differential privacy to derive insights while adhering to privacy regulations. Cookieless analytics are an alternative to web analytics explored by marketers such as Piwik PRO and Matomo Analytics.
- *Shift in strategies:* Marketers have had to shift strategies toward more contextual and less personalised advertising to comply with privacy regulations. This shift impacts targeting and personalised advertisement delivery, especially for e-commerce and D2C business.

Resurgence of MMM post GDPR/iOS

Marketing-mix modelling has taken several forms since the 1960s, evolving as and when advances in tracking technology have made new types of consumer data available.[44,45] Advertisers employ MMM to measure the effectiveness of their advertising spend.[46] Several researchers have examined the challenges and opportunities[47,48] with marketing-mix models. The following challenges need to be re-examined in light of recent regulatory developments:

- *Limited range of data available and data aggregation issues*: Marketing-mix models rely on historical data to identify patterns and relationships between marketing activities and business outcomes. With a reduced dataset, there is a risk of decreased model performance and less precise estimates of marketing effectiveness. Post GDPR, there has been a decline in the availability of common forms of data, whether response data, media metrics, marketing metrics or control factors such as seasonality, which calls for deeper examination of recent challenges and opportunities.
- *Selection bias of correlated input variable data*: GDPR emphasises the need for user consent and transparency in data processing. This can result in selection bias, where the available data may not represent the entire target audience due to consent

issues or opt-outs. In a multichannel advertising environment, advertisers tend to allocate their spending across channels in a correlated way. Additionally, platform-built marketing-mix models do not have access to competitor datasets.
- *Model selection process*: The model selection process has become critical post GDPR as advertisers need to choose methods that prioritise data privacy while still providing meaningful insights. The challenge lies in finding models that strike the right balance between accuracy and privacy compliance. Choosing inappropriate models may lead to unintended privacy breaches or inaccurate results depending on the MMM results of the respective platform, such as Lightweight (Google) or Robyn (Meta).

In general, MMM requires 3–4 parameters for each channel to evaluate the diminishing returns and optimisation of media budget, and a minimum of 7–10 data points per parameter for a stable linear regression, which is challenging to find in the industry setting.[49] In spite of all these challenges, the resurgence of MMM in the attribution industry is a response to the evolving marketing ecosystem, where the need for a comprehensive, cross-channel understanding of marketing effectiveness has become crucial for businesses aiming to optimise their marketing spend and strategies. Increasing privacy regulations and limitations in user-level data tracking have made it challenging for MTA and LTA models to function effectively. With its focus on aggregated data analysis, MMM is less reliant on user-level data, making it a viable option in the changing privacy landscape. MMM offers a holistic approach, considering both online and offline channels, to provide a comprehensive view of marketing impact. MMM allows for the integration of traditional advertising media, such as television, radio and print, with digital channels like online advertising and social media. This integration is crucial as brands seek to understand the synergies and interactions between different media types in influencing consumer behaviour. Additionally, the availability of extensive datasets and advancements in analytics tools have facilitated more sophisticated MMM techniques, improvements in statistical methodologies, machine-learning techniques and modelling approaches have enhanced the accuracy and applicability of MMM across various industries, making it a valuable tool in contemporary marketing strategies.

Multi-touch attribution or conversion attribution (bottom-up approaches)

Google's first heuristic model for attribution — the last-touch attribution model — was introduced in 2005, opening up a new way of calculating ROAS for online marketing channels. It is widely agreed, however, that the platform LTA report is at best flawed, either overestimating or underreporting the tactics/channels directly responsible for sales.[50,51] In a survey conducted for this study (the methodology of which will be discussed in due course), 69.2 per cent of the 51 respondents indicated that they do not believe LTA adequately captures the impact of marketing efforts, while 26 per cent respondents said they only partially agreed that LTA adequately captures the impact (Figure 2). Only 4.6 per cent of respondents agreed that LTA adequately captures the impact of marketing efforts in practice. Nevertheless, and despite the increasing scarcity of user-level data, LTA is still extensively utilised as a benchmark tool, because without it there would be no platform reports for online advertisers to improve/optimise media spending.

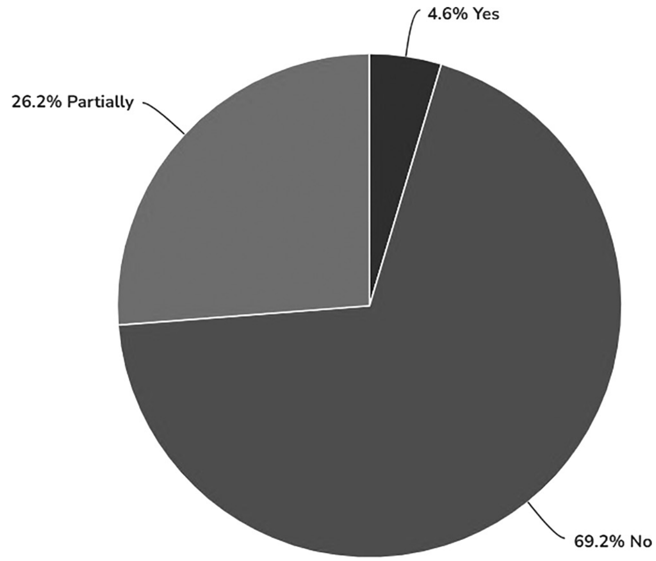

Figure 2 Response to the question, 'does last-touch attribution adequately capture the impact of your marketing efforts?'

Self-attributed conversions refer to the actions or conversions that platforms attribute to themselves based on their own tracking mechanisms. This includes mobile measurement partner (MMP) attribution, which refers to the process of attributing mobile app installs, user actions or conversions to specific marketing efforts or campaigns. MMPs are third-party platforms or services that help mobile app developers or advertisers track and attribute the source of app installs or in-app actions to various marketing channels or campaigns. For instance, if a user clicks on an advertisement on Facebook, then later sees the same product in a Google advertisement and makes a purchase, both Facebook and Google might claim credit for the conversion. Self-attributed conversions are the conversions that platforms claim based on their own tracking data, often using different attribution models (like LTA or MTA) to determine which touch point gets the credit for the conversion. These metrics can be valuable for the platform to showcase the effectiveness of its advertising services. However, discrepancies can arise when different platforms claim credit for the same conversion, leading to challenges in accurately measuring the true impact of each marketing channel.

The most popular MTA methods can be summarised as follows:[52]

- *Last-click and first-click:* The overall effect on the conversion is attributed to the last activity (source) or the first activity on the path.
- *Last non-direct click:* The overall effect on the conversion is attributed to the recent activity on a path that was not a direct access to a website.
- *Logistic regression:* The effect on the conversion is studied on the basis of logistic regression based, in turn, on the decomposition of all conversion paths and the binary assignment of the presence or absence of the channel on the path.
- *Linear:* The impact on the conversion is assigned proportionally to each activity on the path.
- *Position-based:* The effect on the conversion is assigned depending on the position of the activity on the path. For

example, Google Analytics assigns a default of 40 per cent of the impact to the first and last source, and the remaining 20 per cent is divided proportionally between other activities.

- *Customised weights:* The effect on the conversion is assigned arbitrarily and subjectively to each source (most frequently based on a previous, more advanced analysis).
- *Markov chain:* The effect of sources on the conversion is determined on the basis of an analysis of the incremental impact of the entire source in the population. Based on all conversion paths, chains are created with the probability of user migration between individual sources assigned. During the analysis, individual sources are removed from the calculation area and probability flows are examined in chains without an excluded source. The resulting difference is an incremental impact that illustrates the real impact of a given source on the final conversion.
- *Shapley value:* The game theory approach and the Shapley value method are a measure of a channel's average marginal contribution to each channel set (coalition, which is a unique path to the purchase scheme). The marginal contribution of a particular channel is an average difference between conversion results of channel sets (coalition) with and without a particular channel.

Incrementality experiments post GDPR

Running randomised experiments (or incrementality experiments) is becoming the standard approach for measuring the marginal effectiveness of online campaigns.[53,54] These controlled experiments have gained popularity post GDPR because they work as a validation tool for attribution methods when examining lifted conversions for platform attribution, MMP reports or platform reporting (Google Analytics).

Incrementality methods make it possible to measure the true impact of marketing efforts, specifically understanding whether a particular marketing channel is genuinely responsible for driving additional sales or conversions. By isolating the difference in behaviour between exposed and unexposed groups, these methods avoid the attribution errors seen in traditional models, providing clear insights into the real incremental lift generated by marketing activities without the need for user-level data. This enables businesses to optimise marketing budgets more effectively, allocating resources to channels that have causal impact than merely coinciding with correlation conversions.

Incrementality controlled techniques offer a more accurate understanding of the real impact of marketing activities; however, implementing them can be complex. Although the internet has reduced the cost of experimentation, contributing to the growth of field experiments in marketing literature, it is still not feasible to run experiments to test every single tactic/channel in a multi-channel environment. To obtain the best insight, Simester[55] recommends: (a) a combination of consumer survey and field experiment; and (b) complementing a single experiment with a large number of experimental treatments. In the next section, we propose a triangulation approach, which is a more effective way to meet marketers in an industry setting.

The combination of marketing-mix modelling and incrementality controlled experiments with attribution tracking provides the necessary entropy to make effective advertisement investment decisions. Implementing controlled incrementality results into MMM and MTA requires a systematic triangulation approach, combining data analysis, model adjustment, validation and ongoing optimisation to derive meaningful and actionable insights for improved marketing decision-making.

In the context of incrementality, triangulation refers to the process of using multiple methods or approaches to measure the incremental impact of a marketing campaign accurately. When determining the effectiveness or incremental lift generated by a campaign, triangulation involves employing different measurement techniques or methodologies to cross-validate and corroborate the findings. The results reported by platform LTA models are cross-validated with MMM and incrementality testing (lift + RCT) results. Repetitive tests/experiments are run to calibrate the 'multiplier' (common denominator) to identify true advertisement effectiveness, primarily for ROAS.[56]

Triangulation is an effective way to tackle the challenges relating to the diminishing availability of user-level disaggregated third-party data and attribution, as there are interdependencies in the main three models. Within the industry setting, other key reasons for the growing interest in triangulation include the following:

- *Measuring causal impact*: MMM attributes conversions or outcomes to different marketing channels based on observed correlations. Incrementality testing, on the other hand, helps measure the actual causal impact of specific marketing activities by isolating the effects of those activities through controlled experiments. Advertising's causal impact is referred to as 'incrementality'.[57] The actions that would not have occurred if the advertisements had not been displayed are referred to as the incremental actions caused by advertising. Rigorous causal inferences can be obtained through randomised experiments, which are often implemented in the form of geo experiments.[58] Before GDPR and iOS 14, common challenges associated with RCTs included opportunity costs and the high expense of conducting experiments across multiple channels. Post GDPR, however, these experiments have become vital due to the reduced availability of data and the inaccuracy of platform models. Advertisers now rely on incrementality results to make tactical decisions about their marketing spend.
- *Validation of correlation vs causation*: MMM relies on historical data and correlations between marketing activities and outcomes. Incrementality online controlled experiments validate whether these observed correlations are indeed due to causal relationships or if other factors might influence the outcomes.[59]
- *Accounting for external factors*: MMM might not always capture the full effect of external factors (eg seasonality, market trends) on outcomes. Incrementality tests can control these external factors by creating controlled environments, providing a clearer understanding of the true impact of marketing actions.
- *Platform-specific validation*: Different marketing platforms may have unique attribution models and methodologies. Incrementality testing allows for platform-specific validation of MMM results, ensuring that the insights derived from each platform align with the observed incremental effects. By running incrementality testing to validate MMM and LTA results from various platforms, businesses can ensure a more accurate understanding of the causal impact of their marketing efforts.[60] Researchers have begun using field experiments as a means of validating marketing models. Field experiments

provide a perfect environment for validation, as various policies can be implemented in the treatment and control settings and their results compared. This offers a 'model-free' basis for validation in addition to a comprehensive test of all the assumptions in the model. This approach enables more informed decision-making, leading to the more effective allocation of marketing resources and improved campaign strategies. Various RCT incrementality tests are now available to improve model accuracy and assist advertisers with their decision-making. Additionally, decision-making in geo incrementality experiments is vital for optimising marketing strategies, improving resource allocation and ensuring that businesses operate efficiently in diverse geographic markets. The ability to make informed decisions based on experiment results is key to achieving success in localised marketing efforts.[61]

- *Optimising future investments:* Validating MMM results through incrementality testing helps in making more informed decisions about future marketing investments. It guides marketers in optimising budgets and strategies by relying on validated, causally linked insights rather than purely correlational data.[62]
- *Adapting to changes:* As consumer behaviour, market dynamics and advertising platforms evolve, incrementality testing ensures that MMM remains relevant and adaptable. It helps in updating models to reflect changing consumer responses to marketing actions.

Competitive advantages of triangulation

The triangulation approach pursued in the present study aims not only to approximate the causal value of marketing impact but also to mitigate the weaknesses associated with individual methodologies post GDPR. Marketing-mix modelling, for instance, is adept at providing insights at the channel, source or tactical level, yet it falls short in addressing campaign-level nuances. For example, MMM excels in cross-channel performance evaluation but may fall short in terms of speed and granularity of results.[63] MMM is best suited for overall budget distribution rather than pinpointing the exact amount of incremental return driven by a specific channel or format of advertisement, which is necessary for establishing its profitability in effectively optimising campaigns. To bridge this gap, incrementality testing emerges as a valuable tool, offering a comprehensive understanding of a specific channel's true impact. Incrementality controlled tests (without reliable attribution models (MTA/LTA)) serve as a crucial validation method for MMM results obtained from various platforms. However, due to resource constraints, it is feasible only to conduct intermittent rather than continuous measurements, thereby necessitating the strategic selection of channels for analysis. Additionally, the triangulation approach supports effective decision-making at a tactical level to optimise advertising spending or to scale testing. Incrementality testing serves as an indicator for analysing the various factors involved in calculating return on investment. For example, when evaluating display advertising, it is critical to examine associated metrics, such as cost per thousand impressions (CPM), conversion rate, targeted audience conversion probability, etc.[64] The integration of these methodologies is imperative to achieve a holistic assessment, but the standard metrics may not remain same post GDPR due to ongoing changes. For instance, incrementality testing, despite its granular insights, is not always feasible for all channels due to cost and time constraints. Consequently, periodic readings are undertaken, serving as calibration points to refine and validate the marketing-mix model. This synthesis ultimately provides a nuanced estimation of the value attributed to platforms such as Facebook, although it remains insufficient when it comes to delineating the exact campaigns that contributed to this value.

In addressing the specific challenge of evaluating campaign effectiveness at the campaign level, the utility of MTA or a combination of LTA with platform attribution is advocated. Imagine a scenario where a vendor report claims a 40 per cent conversion rate for a particular metric; however, when more sophisticated attribution methods, such as MMM or LIFT/geo testing, are applied, they reveal a 30 per cent conversion rate.

The difference between the two results is calculated as the vendor report's 40 per cent minus the advanced attribution result's 30 per cent, resulting in a difference of 10 per cent. To express this difference as a percentage, it is divided by 2, yielding a multiplier of 0.5 per cent.

In simpler terms, this means that the advanced attribution methods suggest a 10 per cent lower conversion rate compared with what the vendor initially reported. The multiplier of 0.5 per cent serves as a way to adjust or account for this difference when interpreting or using the data from the vendor report in a more accurate manner.

For a more granular understanding aligned with campaign objectives, reliance on MTA or a combined platform attribution approach is recommended. This can determine how much of the advertising budget can be reduced/held at same ratio or how to scale lower-funnel campaign channels now that there is an idea of the relative ROAS in reference to MMM channel-level output and true ad effectiveness from incrementality tests, as shown in Figure 3. As you scale spend upwards, however, there is likely to be diminishing returns/decreasing return on investment.

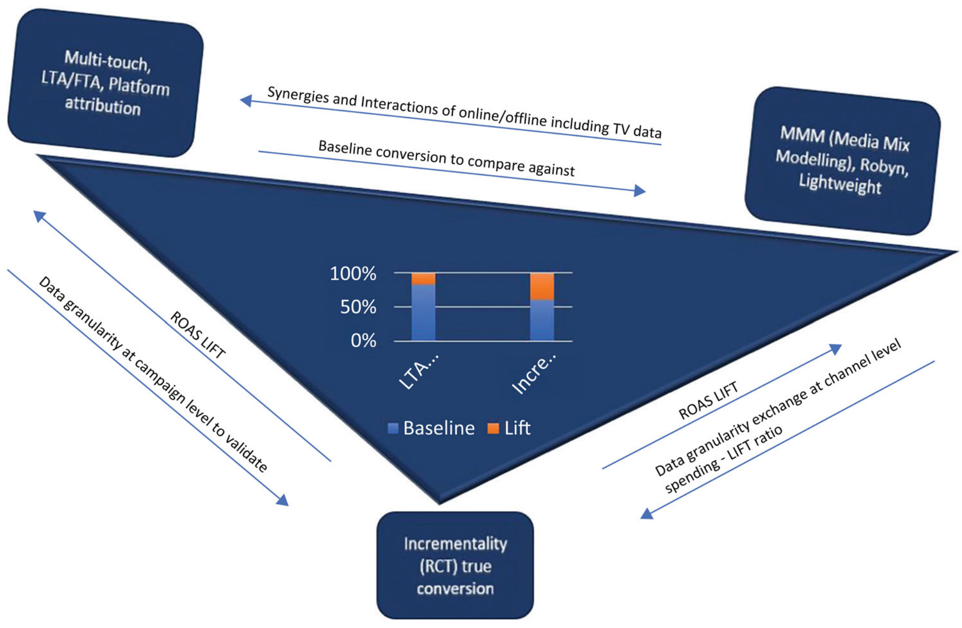

Figure 3 The triangulation approach

The redistributive potential of such methodologies, guided by relative shares determined through platform attribution, underscores the importance of a comprehensive assessment that incorporates the diverse strengths of each modelling technique. It is imperative to note that the outcomes of platform attribution should not be disregarded, as they contribute valuable insights to the overall evaluation. This redistributive conversion[65] serves as the multiplier to correct any discrepancies reported by website analytics attribution or self-attributed conversions, assuming that periodic incrementality controlled tests are conducted to check the true conversion rate for each supplier.

Challenges in implementing marketing decision-making models: Criteria and evolutionary trends

Although understanding regarding the development of models for marketing decisions has grown significantly, there remain concerns regarding the empirical foundations on which this knowledge is based. Brodie and Danaher [66] call for attention be given to the issues associated with the validation and empirical testing model, which are key decision-making criteria.[67] The proposed survey methodology represents an effort to obtain an unbiased analysis of the decision-making challenges faced by practitioners in implementing marketing models with a focus on experimentation/testing methodology.

As the use of models grows more widespread in many areas of marketing decision-making, the model-building criteria pertaining to model structure, ease of use and implementation strategy will become more widely acknowledged. These criteria form the foundation for constructing advertisement decision-making models.[68,69] They aim to guide marketers in making informed decisions, optimising advertising strategies and maximising the impact of their campaigns while remaining compliant and ethical.

Media-mix modelling functions as a pivotal decision-making model for marketers, aiding in the strategic allocation of advertising budgets across diverse media channels. By leveraging statistical analysis, MMM enables marketers to assess historical data, attribute the impact of each channel to business outcomes and optimise their advertising strategies. In essence, MMM serves as a data-driven tool that enhances the efficiency and effectiveness of marketers' decision-making processes in the dynamic landscape of media advertising. As this paper will discuss in due course, the granularity of the output of marketing-mix models for D2C customers is not satisfactory for optimisation at the campaign level.

Leading platforms such as Google Ads and Meta do offer some capabilities that align with MMM principles; however, they do not provide a full-fledged MMM solution in the traditional sense. This new technology is supposed to allow marketing-mix models to be built in a point-and-click environment, with less specialised MMM statistical knowledge required. These solutions offer automated modelling techniques that can be run on an ongoing basis and typically contain optimisation planning and simulation modules for forward-facing scenario planning. However, it is unclear how easy it is to implement these models and make budget decisions, highlighting the pressing need for a survey.

As discussed, the advanced legacy MTA is a bottom-up approach that, until recently, enjoyed phenomenal success in the industry. The reason for the widespread adoption of such models was because of their ability to attribute conversions to specific media sources in a cross-channel environment, thereby helping marketers to make better informed decisions to optimise ROAS. Advanced MTA models that operate without incrementality RCT testing fail to meet essential decision-making criteria for implementation, such as ease of use, model accuracy and interpretability, for performance marketers to make informed decisions regarding the optimisation of their advertising.[70] This is one of the main reasons for MMM and MTA models to evolve to an advanced attribution 2.0 incrementality-based framework.[71,72]

Both heuristic and advanced attribution (MTA) methods have been helpful in day-to-day optimisation and providing key input for AI-powered advertising campaigns, but the combination of privacy updates and the managed decline of third-party cookies has made it more difficult to rely on them alone for decisions regarding budget allocation. Additionally, marketers need a holistic view of all activities that may have influence in the attribution model and to refer to the framework proposed by Hosahally and Zaremba.[73] Especially in times of economic uncertainty, ROAS for all channels must be evaluated at frequent intervals, bearing in mind that platform-level LTA results have a tendency to over-estimate or under-report compared with incrementality test results. This inconsistency is not limited to paid media alone; there is also a need to re-examine the earned media and competition media contribution[74] as a changing economic climate induces significant changes in consumer behaviour.

To ensure industry adoption, advertisement decision-making models, like MMM, MTA, LTA and controlled incrementality, must meet the standard decision-making criteria. The benchmark for such criteria will vary based on the specific needs, goals and characteristics of that industry. That said, there are several common criteria that organisations typically consider when deciding to implement a model.

Per Little,[75] a decision calculus is defined as a model-based set of procedures for processing data and judgments to assist managers with their decision-making. The main criteria for a decision-making model are as follows:

- *Ease of use:* The marketing model needs to be easy to use, easy to understand and easy to implement.
- *Easy to control:* Marketers should be able to make the model behave the way they want it to. MMM should guide marketers on where and how to allocate resources effectively.
- *Accuracy:* The model must provide accurate results. The model should accurately reflect the impact of advertising efforts on desired outcomes, such as sales, brand recognition, conversion, ROAS or customer engagement. It needs to provide reliable insights into the effectiveness of different advertising channels and strategies. As stated earlier, open source MMM tools only partially meet the accuracy criteria as they fail to provide insight for tactical decision-making as the data on which they operate are platform-restricted and lack granularity.
- *Relevance:* It is crucial for the model to align with the specific goals and objectives of the advertising campaign. The model should focus on metrics and measurements that relate directly to the intended outcomes, such as increasing conversions, driving traffic or boosting brand awareness.
- *Validation:* The model must support the validation of results. Many platform-provided marketing-mix models lack transparency in terms of the algorithms used, data-processing methods and the

underlying assumptions. Inbuilt models rely heavily on the data available within the platform even though such data can be limited in scope. Advertisers might not have access to comprehensive cross-channel data, including offline or non-digital touch points, making it impossible to validate model data or model performance.

- *Robustness:* the model should have great potential to be robust and can handle high volume of data and complexity.
- *Actionability:* Marketers need actionable insights. The decision-making model should provide information that can be translated into practical steps and strategies.
- *Transparency:* Marketers must be able to understand and trust the results. Putting one's faith in the back-end data-processing of black-box models is a risky practice.
- *Predictiveness:* The model should have predictive capabilities, allowing marketers to anticipate the potential outcomes of different advertising strategies or changes in the campaign. Predictive models enable better decision-making by forecasting how alterations may impact results.
- *Ethical and legal compliance:* With increasing regulations around data privacy and advertising practices, decision-making models must comply with ethical standards and legal requirements regarding consumer privacy, data usage and advertising practices.

METHODOLOGY

The survey questionnaire incorporated both quantitative and qualitative elements in what is known as mixed-methods research. Such surveys collect numerical data to quantify trends and patterns, in addition to gathering qualitative data to explore the underlying reasons or to provide context. The combination of quantitative and qualitative data helps provide a more comprehensive understanding.

Quantitative surveys often use closed-ended questions with predefined response options, such as multiple-choice questions or Likert scales. For example:

- Do you believe last-touch attribution adequately captures the impact of your marketing efforts?
- (a) yes; (b) no; (c) partially.
- What specific challenges or concerns do you anticipate when implementing RCT incrementality tests?
- (a) lack of expertise in designing the experiments; (b) data privacy issues; (c) resource constraints; (d) budget constraints; (e) other (please specify).
- Rate your satisfaction with the product on a scale of 1 to 5, with 5 being the highest.

To obtain qualitative data, questions may be augmented with a supplementary question, such as:

- Additionally, please provide comments explaining your rating (qualitative).

Mixed methods (qualitative and quantitative) research, specifically conducting surveys with expert marketers, is the ideal approach for identifying challenges in the implementation of marketing models, particularly in the context of decision-making. The utilisation of a survey allows for a comprehensive exploration of perspectives and experiences from a diverse range of experts, providing valuable insights into the challenges faced in the field. By targeting a broad audience of marketing measurement experts, the research gains a holistic view, encompassing varied backgrounds and contexts within the industry. Surveys also facilitate the collection of nuanced and detailed responses, allowing participants to elaborate on specific challenges they have encountered when implementing decision-making models. This approach is effective in capturing both common and unique issues, enabling a thorough understanding of the landscape. Moreover, surveys offer a structured means of data collection, ensuring consistency in the information gathered and allowing for the quantitative analysis of responses, thereby contributing to a robust and evidence-based exploration of the challenges associated with the implementation of marketing models.

Survey design and implementation

A web-based survey was undertaken using computer-assisted web interviewing; an internet surveying technique in which the interviewee follows a script provided on a website.

By targeting expert marketers, the research aims to capture the intricate details and real-world experiences related to implementation of decision-making models. The post-GDPR setting introduces additional complexities and considerations, making it crucial to understand how these regulations impact decision-making processes.

Participant selection criteria

Interviewees were recruited through the tailored screening of individuals with pertinent backgrounds in marketing measurement and were recruited from both brands and measurement agencies, including top incrementality measurement agencies and D2C brands marketers such as Measured, Ruler Analytics, Nielsen, Search Discovery, Lifesight, Adfactors PR, Visual IQ Technology Services, Canvas worldwide, Tinuiti, and Incrmntal.

Eligibility criteria were determined based on the respondents' familiarity with attribution and incrementality methodologies and their professional engagement in the online advertising industry. Only 50 respondents met the specified eligibility criteria and were included in the subsequent analysis.

Over 50 per cent of respondents had over ten years of work experience in the marketing field and were using incrementality RCTs to validate results obtained from either LTA or MMM.

Interview questions were carefully designed to identify the decision-making criteria with specific marketing-mix models (Robyn, Lightweight, and in-house) and the usage of RCT incrementality models.

The questionnaires were developed using a program for creating web interviews. This software allows the flow of the questionnaire to be adjusted based on the answers provided, as well as information already known about the participant.

Data analysis

Survey responses were subjected to thorough analysis to extract meaningful insights and trends. This analysis included the calculation of percentages and mean values, in addition to identifying the distribution of numerical ratings for each criterion in order to detect common themes and patterns. The analysis sought to provide a quantitative understanding of how respondents perceive different aspects of the models.

Limitations of survey methodology

While diligent efforts were undertaken to capture diverse perspectives in this study, it is crucial to acknowledge the limitations associated with survey-based research. First, it is possible that the initial recruitment process introduced a degree of selection bias. Given also the relatively small size of the sample, it cannot be assumed to be representative. Specifically, if certain demographics are over or under-represented, this compromises the representativeness of the population, introducing the risk of sample bias. The dependence on self-reported information also introduces the risk of memory lapses, inaccuracies or intentional misrepresentation. For example, respondents may provide biased or socially desirable answers, thus distorting the accuracy of the data.

In addition, questionnaires, although efficient, have limitations. For example, structured response options are more likely to miss qualitative nuances, thus limiting the exploration of complex issues. Ambiguity and misinterpretation in question wording can lead to inconsistent responses. Lack of context, inability to probe for deeper insights and limited flexibility in adapting to emerging insights or changing circumstances are additional constraints. Non-response bias and the potential overemphasis on quantitative data also pose challenges to the study's generalisability.

Lastly, the qualitative nature of the study restricts the generalisability of findings beyond the sampled population.

Ethical considerations

All participants provided informed consent before participating in the survey, and their confidentiality and anonymity were strictly maintained throughout the research process

RESULTS AND DISCUSSION

The main marketing-mix models used by the marketers are Lightweight by Google, Robyn by Meta and in-house models. We designed this survey to highlight the decision-making challenges faced by marketers when implementing these models given the greater reliance on MMM to measure the performance of online advertisements post GDPR.

The survey uses a five-point scale to measure respondent sentiment, where 1 is the lowest possible rating, and 5 is the greatest. As reported in Tables 2–4, respondents rated decision-making criteria for three different models, namely, Lightweight, Robyn and in-house models.

Lightweight is a Bayesian MMM library that allows users to train marketing-mix models and obtain channel attribution information. Lightweight is designed to make it easy for marketers to build Bayesian marketing-mix models by providing the

Table 2: Decision-making criteria ratings with percentages for MMM Lightweight model used by marketers.

Criteria	*1*	*2*	*3*	*4*	*5*
Ease of use	14.29%	14.29%	0.00%	28.57%	42.86%
Easy to control	0.00%	0.00%	42.86%	28.57%	28.57%
Relevance	0.00%	0.00%	28.57%	57.14%	14.29%
Accuracy	0.00%	0.00%	28.57%	57.14%	14.29%
Validation	0.00%	14.29%	57.14%	14.29%	14.29%
Robustness	0.00%	14.29%	14.29%	28.57%	42.86%
Actionability	0.00%	14.29%	28.57%	14.29%	42.86%
Transparency	0.00%	28.57%	0.00%	42.86%	28.57%
Predictiveness	0.00%	28.57%	0.00%	28.57%	42.86%
Ethical and legal compliance	0.00%	0.00%	14.29%	14.29%	71.43%

Table 3: Decision-making criteria ratings with percentages for MMM Robyn model used by marketers.

Criteria	*1*	*2*	*3*	*4*	*5*
Ease of use	0.00%	20.00%	40.00%	13.33%	26.67%
Easy to control	6.67%	6.67%	46.67%	26.67%	13.33%
Relevance	0.00%	0.00%	13.33%	40.00%	46.67%
Accuracy	0.00%	0.00%	26.67%	26.67%	46.67%
Validation	0.00%	0.00%	26.67%	33.33%	40.00%
Robustness	0.00%	0.00%	26.67%	26.67%	46.67%
Actionability	0.00%	0.00%	20.00%	20.00%	60.00%
Transparency	0.00%	6.67%	20.00%	53.33%	20.00%
Predictiveness	0.00%	0.00%	26.67%	33.33%	40.00%
Ethical and legal compliance	0.00%	0.00%	20.00%	20.00%	60.00%

Table 4: Decision-making criteria ratings with percentages for MMM In-house models used by marketers.

Criteria	*1*	*2*	*3*	*4*	*5*
Ease of use	0.00%	26.67%	33.33%	0.00%	40.00%
Easy to control	6.67%	20.00%	26.67%	13.33%	33.33%
Relevance	0.00%	0.00%	6.67%	40.00%	53.33%
Accuracy	0.00%	0.00%	33.33%	40.00%	26.67%
Validation	0.00%	6.67%	20.00%	33.33%	40.00%
Robustness	0.00%	0.00%	0.00%	40.00%	60.00%
Actionability	0.00%	0.00%	0.00%	20.00%	80.00%
Transparency	0.00%	0.00%	13.33%	20.00%	66.67%
Predictiveness	0.00%	6.67%	26.67%	13.33%	53.33%
Ethical and legal compliance	6.67%	0.00%	0.00%	46.67%	46.67%

functionality to appropriately scale data, evaluate models, optimise budget allocations and plot common graphs used in the field.

As shown in Table 2, for the Lightweight model, respondents expressed a positive view, with 'ease of use' and 'actionability' receiving the highest percentages of 42.86 per cent and 42.86 per cent, respectively. 'Validation' and 'robustness' also garnered favourable responses, while 'relevance' received a relatively lower percentage of 14.29 per cent.

Per Table 3, respondents showed positive sentiment with respect to the Robyn model, with 'easy to control', 'robustness' and 'predictiveness' obtaining high percentages of 46.67 per cent, 46.67 per cent and 40 per cent, respectively. However, 'ease of use' and 'relevance' had lower percentages, suggesting potential areas for improvement.

As shown in Table 3, marketers displayed positive sentiments towards the 'actionability', 'transparency' and 'predictiveness' of in-house models, with the highest percentages of 80 per cent, 66.67 per cent and 53.33 per cent, respectively. However, 'ease of use' and 'relevance' had relatively lower percentages, indicating potential focus areas.

Overall, these tables provide insights into marketers' perceptions of decision-making criteria for different marketing models, highlighting both strengths and potential areas for improvement. Post GDPR, the challenges associated with MMM include the fact that the data feeds of open source tools (such as Robyn or Lightweight) are limited to the respective platform's data. Building a comprehensive MMM model requires data to be collected from other media channels, including offline ones. Additionally, when marketers need ongoing analysis or insights at a tactical level, MMM fails to deliver. Due to a lack of granularity, platform-supported MMMs such as Meta's Robyn and Google's Lightweight do not provide accurate high-level insights into campaign-level performance and decision-making.[76,77]

Comparative analysis

Comparative analysis of the average ratings for each criterion across the 'in-house', Lightweight and Robyn models yields the following insights:

- *Ease of use:* Lightweight has the highest average rating (3.71), suggesting it is perceived as the easiest to use among the three. The Robyn and in-house models have similar ratings, with 3.47 and 3.53, respectively.
- *Easy to control:* Here, Lightweight again scores slightly higher (3.86) than the other two models, indicating it may offer better control to marketers. Robyn has the lowest average rating (3.33) in this category.
- *Relevance:* The in-house model scores the highest (4.47), indicating it aligns well with specific needs. Robyn also scores high (4.33), while Lightweight has a lower average (3.86).
- *Accuracy:* Robyn scores the highest in accuracy (4.20), closely followed by the in-house model (3.93). Lightweight has an average rating of 3.86.
- *Validation:* The in-house model is perceived to be the best for supporting validation (4.07), with Robyn closely behind (4.13). Lightweight has the lowest average rating (3.29) in this criterion.
- *Robustness, actionability, transparency, predictiveness and ethical and legal compliance:* In these areas, the ratings vary, but generally, the in-house and Robyn models tend to score higher, indicating they are perceived as more robust, actionable, transparent, predictive and compliant compared with Lightweight. Notably, the in-house model scores exceptionally high in actionability (4.80) and robustness (4.60).
- *Overall trends:* The in-house model appears to be highly regarded for its robustness, actionability and relevance, scoring the highest in these categories. Lightweight, while scoring slightly lower in some areas, is perceived to be the easiest to use and control. Robyn offers a good balance across all criteria, scoring particularly well in accuracy and relevance.

These insights suggest that the choice between these models may depend on the specific priorities of users or the context in which the model is to be used. For instance, if ease of use and control are paramount, the Lightweight model might be preferred. However, for applications where accuracy, relevance and compliance are critical, the 'in-house' or Robyn models may be more suitable.

Recommendations

The predominant obstacles cited were a lack of expertise in experiment design (56 per cent), resource constraints (52.6 per cent), budget limitations (38.6 per cent) and concerns related to data privacy (22 per cent), as shown in Figure 5. A subsequent enquiry regarding data availability indicated that respondents were notably concerned about GDPR-related issues (47.7 per cent). Additional challenges included statistical complexity (33.8 per cent), as shown in Figure 4, time constraints (38.5 per cent) and budget limitations (44.6 per cent). Future research could focus on these areas with a view to building commercially viable and effective decision-making models.

When assessing the effectiveness of marketing models, respondents highlighted key decision-making criteria. Notably, accuracy emerged as crucial, with 46.2 per cent assigning it the highest rating of 5. Robustness, the model's ability to navigate market fluctuations, was also emphasised, garnering ratings of 4 or 5 from 71.1 per cent of respondents. The ease of

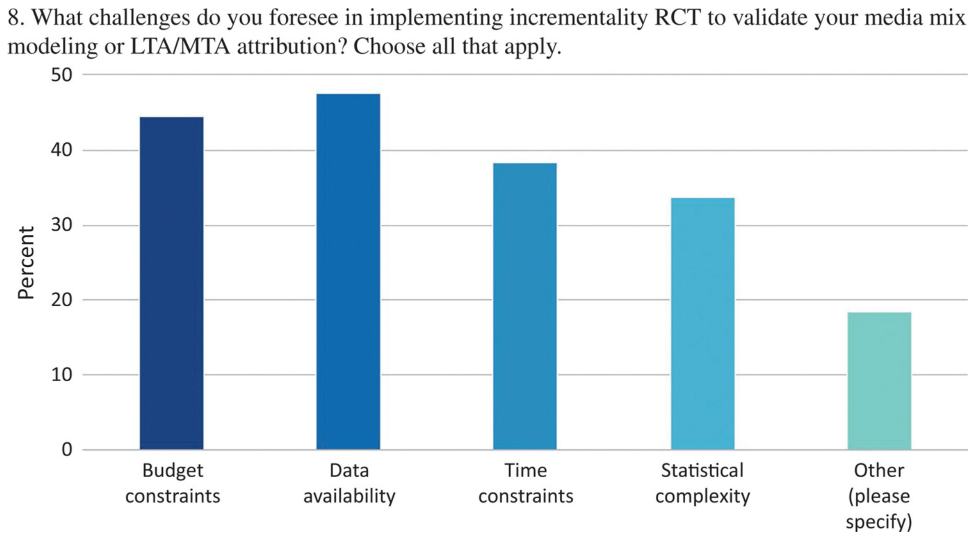

Figure 4 Key challenges faced by marketers when implementing incrementality randomised controlled trials

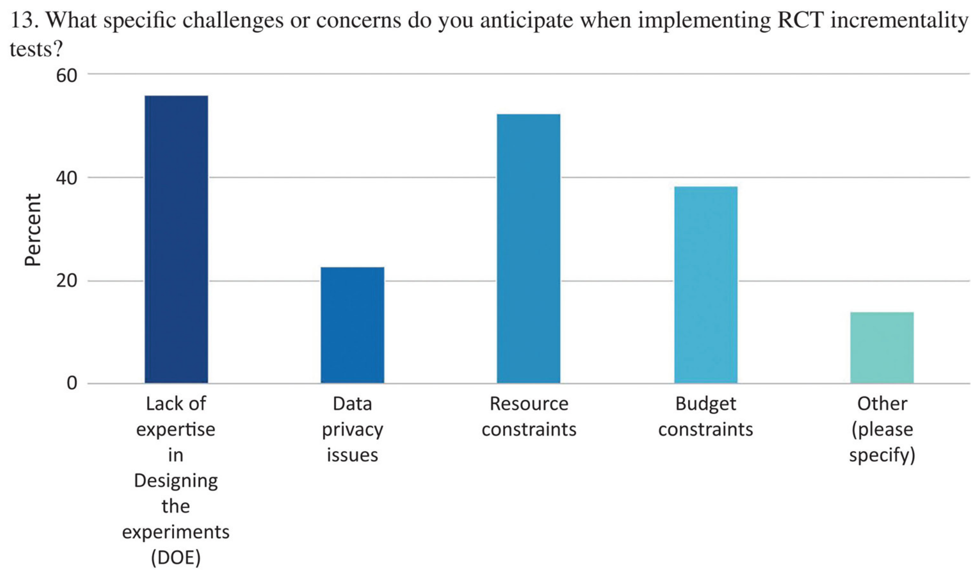

Figure 5 What specific challenges or concerns do you anticipate when implementing randomised controlled trials?

interpretation and implementation followed suit, with 60.6 per cent recognising its significance by rating it 4 or 5. Cost-effectiveness, while relevant, did not stand out prominently, with responses distributed across various ratings, showing a slight preference towards 3 and 4. Predictiveness, indicating the model's ability to foresee outcomes, was considered important by 69.2 per cent of respondents, who assigned ratings of 4 or 5. Relevance, aligning the model with specific goals and objectives, received the highest emphasis, with 78.8 per cent rating it as 4 or 5. The analysis underscores a strong emphasis on accuracy, robustness, predictiveness and relevance. Ethical and legal compliance, though important, was slightly less highlighted, and cost-effectiveness, while acknowledged, did not emerge as a top priority.

The insights from the survey suggest that marketers prioritise incrementality models due to their prowess in accuracy, robustness, predictiveness and relevance as shown in Table 5. Over 80 per cent of respondents rated the model's accuracy as 3 or 4, while only 40 per cent consider it cost-effective. The incorporation of RCTs in incrementality testing contributes to the validity, usability and applicability of insights, bolstering trustworthiness and practicality. Future research is deemed necessary for refining incrementality-based approaches. When considering MMM in conjunction with incrementality testing, customer feedback surveys, advertising testing and data enrichment, a comprehensive, data-driven measurement approach emerges. Additionally, in the absence of user-level data, advertisers may explore alternative attribution methods, such as relying on first-party data and contextual targeting. The evolving landscape, marked by the diminishing significance of cookies and iOS changes, positions the combined MMM and incrementality approach as a potential gold standard for marketers, warranting ongoing research and exploration in the field.

It can also be argued that RCT and MMM collectively can meet the other decision-making criteria. RCT incrementality contributes to the validity, usability and applicability of the insights derived from the testing. Advertisers rely on incrementality results to make effective decisions regarding the optimisation of their advertising spend, and adherence to these criteria enhances the trustworthiness and practicality of the findings.[78]

Practitioners and researchers have accepted incrementality as the gold standard for the measurement of advertisement performance as, unlike attribution methods, they depend heavily on disaggregated user-level data.[79,80] However, controlled experiments require robust experimentation methodologies, control groups and a deep understanding of data

Table 5: Decision-making criteria ratings with percentages for RCT incrementality model used by marketers.

Criteria	*1*	*2*	*3*	*4*	*5*
Accuracy of the model	2 (3.8%)	1 (1.9%)	6 (11.5%)	19 (36.5%)	24 (46.2%)
Robustness	0 (0.0%)	2 (3.8%)	13 (25.0%)	27 (51.9%)	10 (19.2%)
Ease of use, interpretation and implement	1 (1.9%)	2 (3.8%)	18 (34.6%)	20 (38.5%)	11 (21.2%)
Cost-effectiveness	1 (1.9%)	8 (15.4%)	20 (38.5%)	12 (23.1%)	11 (21.2%)
Predictiveness	1 (1.9%)	5 (9.6%)	10 (19.2%)	23 (44.2%)	13 (25.0%)
Relevance	0 (0.0%)	4 (7.7%)	7 (13.5%)	18 (34.6%)	23 (44.2%)
Ethical and legal compliance	4 (7.7%)	6 (11.5%)	15 (28.8%)	13 (25.0%)	14 (26.9%)

analysis. Moreover, the platform runs the tests in the background without sharing the design or methodology details with marketers in the case of ghost ads, for example. Therefore, further investigation into the incrementality-based approach is required from practitioners and academics alike. For a more data-driven approach to measurement, MMM may be coupled with other techniques, such as incrementality testing, customer feedback surveys, advertising testing, data partnerships and the enrichment of first-party data. Kohavi *et al.*[81] recommend that online controlled experiments be conducted to inform organisational decisions at all levels from strategy to tactics. Incrementality experiments provide a suitable tool for both tactical and strategic decision-making.[82] Incrementality experimental results help avoid correlation/causality issues by using incrementality outputs to train MMM models. Incrementality provides fresh data on which to make rapid tactical decisions, complementing the long-term strategic and planning strengths of MMMs. Moving forward, the combination of MMM and incrementality will likely be the gold standard in a post-iOS, post-cookie world.

References

(1) Rammos, T. and Harttrumpf, M. (2022) 'Cookies, TTDSG and ePrivacy Regulation: What Can We Expect in 2022 in Terms of Data Protection in the Online Sector?', available at https://www.taylorwessing.com/en/interface/2022/the-year-ahead-ii/cookies-ttdsg-and-eprivacy-regulation-what-to-expect-in-2022-data-protection-and-online-sector (accessed 20th September, 2024).
(2) Goldberg, S., Johnson, G. and Shriver, S. (2019) 'Regulating Privacy Online: The Early Impact of the GDPR on European Web Traffic', *American Economic Journal: Economic Policy*, Vol. 16, No. 1, pp. 325–358.
(3) Çınar, N. and Ateş, S. (2022) 'Data Privacy in Digital Advertising: Towards a Post-Third-Party Cookie Era', in: 'Privacy', Routledge, New York, NY, pp. 55–77.
(4) Goel, V. (2021) 'An Updated Timeline for Privacy Sandbox Milestones', June, available at https://blog.google/products/chrome/updated-timeline-privacy-sandbox-milestones (accessed 20th September, 2024).
(5) O'Flaherty, K. (June 2021) 'IOS 14.5: How This Outstanding New Feature Will Change Your iPhone Forever', available at https://www.forbes.com/sites/kateoflahertyuk/2021/04/24/ios-145-how-this-outstanding-new-feature-will-change-your-iphone-forever/?sh=775d710911bf (accessed 20th September, 2024).
(6) Cinar and Ateş, ref. 3 above.
(7) Goldberg *et al.*, ref. 2 above.
(8) Rammos and Harttrumpf, ref. 1 above.
(9) O'Brien, C., Thiagarajan, A., Das, S., Barreto, R., Verma, C., Hsu, T., Neufield, J. and Hunt, J.J. (2022) 'Challenges and Approaches to Privacy Preserving Post-click Conversion Prediction', arXiv preprint, available at https://arxiv.org/pdf/2201.12666 (accessed 25th September, 2024).
(10) Jerath, K. (2022) 'Mobile Advertising and the Impact of Apple's App Tracking Transparency', available at https://www.apple.com/privacy/docs/Mobile_Advertising_and_the_Impact_of_Apples_App_Tracking_Transparency_Policy_April_2022.pdf (accessed 20th September, 2024).
(11) Cinar and Ateş, ref. 3 above.
(12) Goldberg *et al.*, ref. 2 above.
(13) Feger, A. (2022) 'How D2C Retail Brands Are Evolving in 5 Charts', available at https://www.insiderintelligence.com/content/how-d2c-retail-brands-evolving-5-charts (accessed 20th September, 2024).
(14) Kohavi, R., Tang, D. and Xu, Y. (2020) 'Trustworthy Online Controlled Experiments: A Practical Guide to A/B Testing', Cambridge University Press, Cambridge.
(15) Lewis, R.A, Rao, J.M. and Reiley, D.H. (2011) 'Here, There, and Everywhere: Correlated Online Behaviors Can Lead to Overestimates of the Effects of Advertising', in: Proceedings of the 20th International Conference on World Wide Web, Hyderabad, 28th March–1st April, pp. 157–166.
(16) Vaver, J. and Koehler, J. (2011) 'Measuring Ad Effectiveness Using Geo Experiments', Google Research, available at https://research.google/pubs/measuring-ad-effectiveness-using-geo-experiments/ (accessed 25th September, 2024).
(17) Johnson, G.A., Lewis, R.A. and Nubbemeyer, E.I. (2017) 'Ghost Ads: Improving the Economics of Measuring Online Ad Effectiveness', *Journal of Marketing Research*, Vol. 54, No. 6, pp. 867–884.
(18) Barajas, J., Zidar, T. and Bay, M. (2020) 'Advertising Incrementality Measurement Using Controlled Geo-Experiments: The Universal App Campaign Case Study', ACM, Washington, DC.

(19) Barajas, J., Akella, R., Holtan, M. and Flores, A. (2016) 'Experimental Designs and Estimation for Online Display Advertising Attribution in Marketplaces', *Marketing Science*, Vol. 35, No. 3. pp. 465–483.
(20) Blake, T., Nosko, C. and Tadelis, S. (2015) 'Consumer Heterogeneity and Paid Search Effectiveness: A Large-Scale Field Experiment', *Econometrica*, Vol. 83, No. 1, pp. 155–174.
(21) Johnson, G.A. (2023) 'Inferno: A Guide to Field Experiments in Online Display Advertising', *Journal of Economics & Management Strategy*, Vol. 32, No. 3, pp. 469–490.
(22) Kerman, J., Wang, P. and Vaver, J. (2017) 'Estimating Ad Effectiveness Using Geo Experiments in a Time-Based Regression Framework', Google Technical Report, available at https://research.google/pubs/estimating-ad-effectiveness-using-geo-experiments-in-a-time-based-regression-framework/ (accessed 25th September, 2024).
(23) Lewis, R.A. and Reiley, D. (2008) 'Does Retail Advertising Work? Measuring the Effects of Advertising on Sales via a Controlled Experiment on Yahoo!', available at https://ssrn.com/abstract=1865943 (accessed 25th September, 2024).
(24) Sahni, N.S. (2016) 'Advertising Spillovers: Evidence from Online Field Experiments and Implications for Returns on Advertising', *Journal of Marketing Research*, Vol. 53, No. 4, pp. 459–478.
(25) Vaver and Koehler, ref. 16 above.
(26) Brodersen, K.H., Gallusser, F., Koehler, J., Remy, N. and Scott, S.L. (2015) 'Inferring Causal Impact Using Bayesian Structural Time-Series Models'.
(27) Kerman *et al.*, ref. 22 above.
(28) Barajas *et al.*, ref. 18 above.
(29) Gordon, B.R., Zettelmeyer, F., Bhargava, N. and Chapsky, D. (2019) 'A Comparison of Approaches to Advertising Measurement: Evidence from Big Field Experiments at Facebook', *Marketing Science*, Vol. 38, No. 2. pp. 193–225.
(30) Johnson, ref. 21 above.
(31) Johnson, G. and Lewis, R.A. (2015) 'Cost per Incremental Action: Efficient Pricing of Advertising', Simon Business School Working Paper No. FR 15-29, available at https://ssrn.com/abstract=2668315 (accessed 20th September, 2024).
(32) Gordon, B.R., Jerath, K., Katona, Z., Narayanan, S., Shin, J. and Wilbur, K.C. (2021) 'Inefficiencies in Digital Advertising Markets', *Journal of Marketing*, Vol. 85, No. 1. pp. 7–25.
(33) Blake *et al.*, ref. 20 above.
(34) Brodersen *et al.*, ref. 26 above.
(35) Kerman *et al.*, ref. 22 above.
(36) Vaver and Koehler, ref. 16 above.
(37) Nielsen (2024) 'DMA Regions', available at https://www.nielsen.com/dma-regions (accessed 20th September, 2024)).
(38) Blake *et al.*, ref. 20 above.
(39) Gupta, S. and Chokshi, S. (2020) 'Digital Marketing Effectiveness Using Incrementality', in: Advances in Computing and Data Sciences: 4th International Conference, ICACDS 2020, Valletta, 24th–25th April, Revised Selected Papers 4, Springer, pp. 66–75.
(40) Stapleton, S. (2022) 'The Cookiepocalypse: Why First-party Data Are Going to Matter to Your Newsroom', available at https://reutersinstitute.politics.ox.ac.uk/cookiepocalypse- why-first-party-data-going-matter-your-newsroom (accessed 20th September, 2024).
(41) *Ibid*.
(42) Davies, J. (2018) '"Personalization Diminished": In the GDPR Era, Contextual Targeting Is Making a Comeback', available at https://digiday.com/media/personalization-diminished-gdpr-era-contextual-targeting-making-comeback/ (accessed 20th September, 2024).
(43) Stapleton, ref. 40 above.
(44) Borden, N.H. (1964) 'The Concept of the Marketing Mix', *Journal of Advertising Research*, Vol. 4, No. 2, pp. 2–7.
(45) Leeflang, P.S., Wittink, D.R., Wedel, M. and Naert, P.A. (2013) 'Building Models for Marketing Decisions', Springer Science & Business Media, New York, NY.
(46) Borden, ref. 44 above.
(47) Chan, D. and Perry, M. (2017) 'Challenges and Opportunities in Media Mix Modeling', Google Report No. 16, available at https://research.google/pubs/challenges-and-opportunities-in-media-mix-modeling/ (accessed 25th September, 2024).
(48) Wedel, M. and Kannan, P.K. (2016) 'Marketing Analytics for Data-Rich Environments', *Journal of Marketing*, Vol. 80, No. 6, pp. 97–121.
(49) Chan and Perry, ref. 47 above.
(50) Lee, G. (2010) 'Death of "Last Click Wins": Media Attribution and the Expanding Use of Media Data', *Journal of Direct, Data and Digital Marketing Practice*, Vol. 12, pp. 16–26.
(51) Gordon *et al.*, ref. 29 above.
(52) Zaremba, A. (2020) 'Conversion Attribution: What Is Missed by the Advertising Industry? The OPEC Model and Its Consequences for Media Mix Modeling', *Journal of Marketing and Consumer Behaviour in Emerging Markets*, Vol. 10, No. 1. pp. 4–23.
(53) Chittilappilly, A. (11th July, 2012) 'Using Experiment Design to Build Confidence in Your Attribution Model', MediaPost, available at https://www.mediapost.com/publications/article/178511/using-experiment-design-to-build-confidence-in-you.html (accessed 25th September, 2022).
(54) Johnson *et al.*, ref. 17 above.
(55) Simester, D. (2017) 'Field Experiments in Marketing', in: 'Handbook of Economic Field

Experiments', Vol. 1, Elsevier, Amsterdam, pp. 465–497.
(56) Runge J., Patter, H. and Skokan I. (2023) 'A New Gold Standard for Digital ad Measurement?', *Harvard Business Review,* March, available at https://hbr.org/2023/03/a-new-gold-standard-for-digital-ad-measurement (accessed 20th September, 2024).
(57) Johnson and Lewis, ref. 31 above.
(58) Vaver and Koehler, ref. 16 above.
(59) Kohavi *et al.*, ref. 14 above.
(60) Simester, ref. 55 above.
(61) Kohavi *et al.*, ref. 14 above.
(62) *Ibid*.
(63) Ohlinger, J. D. and Nedyalkov, N. (2023) 'Incrementality Testing: The Key to Unlocking Profitable Growth in a Changing Industry', available at https://www.thinkwithgoogle.com/intl/en–emea/marketing–strategies/data–and–measurement/incrementality-testing/ (accessed 20th September, 2024).
(64) Einhorn, T. (2022) 'How Do I Measure Incrementality on Display Advertising?', available at https://www.measured.com/faq/how-do-i-measure–incrementality-on-display-advertising/ (accessed 20th September, 2024).
(65) Testwuide, T. (2023) 'Worried about Ad Costs? Don't Waste Your Time (and Money)', available at https://www.measured.com/blog/worried-about-ad-costs-dont-waste-your- time-and-money/ (accessed 20th September, 2024).
(66) Brodie, R. J. and Danaher, P. J. (2000) 'Building Models for Marketing Decisions: Improving Empirical Procedures', *The International Journal of Research in Marketing*, Vol. 17, No. 2/3. pp. 135–139.
(67) Leeflang *et al.*, ref. 45 above.
(68) *Ibid*.
(69) Little, J.D. (1970) 'Models and Managers: The Concept of a Decision Calculus', *Management Science*, Vol. 16, No. 8, pp. 462–466.
(70) Hosahally, S. and Zaremba, A. (2023) 'A Decision-Making Characteristics Framework for Marketing Attribution in Practice: Improving Empirical Procedures', *Journal of Digital & Social Media Marketing*, Vol. 11, No. 1, pp. 89–100.
(71) Leeflang *et al.*, ref. 45 above.
(72) Little, ref. 69 above.
(73) Hosahally and Zaremba, ref. 70 above.
(74) Zaremba *et al.*, ref. 53 above.
(75) Little, ref. 69 above.
(76) Bharadwaj, M. (2022) 'Media Mix Modeling or Incrementality Measurement: Which Will Replace Multi-touch Attribution?', *Forbes*, July, available at https://www.forbes.com/sites/forbestechcouncil/2022/07/08/media-mix-modeling-or-incrementality-measurement-which-will-replace-multi-touch-attribution/ (accessed 20th September, 2024).
(77) Lim, S. (2023) 'A New Era of Ads Analytics? How Media Mix Modeling Is Reshaping Strategies in the Cookieless World: Analysis', available at https://www.campaignasia.com/article/a-new-era-of-ads-analytics-how-media-mix-modeling-is-reshaping-strategies-in-the/492927 (accessed 20th September, 2024).
(78) Kohavi *et al.*, ref. 14 above.
(79) Runge *et al.*, ref. 56 above.
(80) Barajas *et al.*, ref. 19 above.
(81) Kohavi *et al.*, ref. 14 above.
(82) Bharadwaj, ref. 76 above.