Attribution and the Marketing Mix Model

Peter Cain
Marketscience, UK

Nitesh Sahay
Marketscience, UK

Classifications, Key Words:

- Endogeneity
- MMM
- MTA
- Consumer Journey
- Unobserved Component Models
- VAR
- VECM
- Cointegration

Abstract

Inconsistent user-identifiers and the walled garden policies of dominant social media players, together with the (imminent) abolition of third-party cookies has led to renewed interest in the marketing mix model as an attribution tool. However, to be useful in a post-MTA world any 'next generation' MMM framework needs to deliver on three fundamental business issues. Firstly, to serve as a true attribution solution, MMM needs to focus on causal estimation methods. Too often we see reliance on consumer journey solutions to address the problems of last-touch-attribution. However, these ignore the critical issues of selection bias endemic in much online media – leading to endogeneity bias and misallocation of the marketing mix. The growing popularity of automated machine learning approaches to the mix model only serve to exacerbate this problem, where the focus is on prediction not causation.

Secondly, MMM needs to quantify the long-term (base-building) effects of marketing and so inform brand-building strategy. Standard approaches are simply not set up to measure these effects, with fixed baselines and a focus on short to medium-term lag structures or Adstocks. Alternative time series structures are required that can quantify both short and long-term (base) variation – coupled with dynamic network models that can explain the causes of base variation and the economics of brand-building.

Finally, next-generation MMM needs to fill the gap left in a cookie-less world to deliver granular and swift insights on marketing ROI and optimal budget allocation. Suitably identified high-dimension mix models – across consumer cohorts by day or hour – can fit the bill. This can provide many of the claimed benefits of MTA such as granular online media effectiveness ranking by publisher and placement, together with the ability to quantify the impact of pricing, offline media, economic factors and longer-term brand-building.

1. Introduction

Marketing attribution attempts to quantify the incremental impact of each element of the marketing mix on consumer demand, typically in the form of a purchase conversion. With the advent of multi-channel marketing and the growing proliferation of off and online channels, accurate attribution represents a significant challenge. In response, modern analytical approaches have evolved into three broad strands.

Firstly, we have Multi-touch attribution (MTA). This focuses on the contribution of online touchpoints to a binary online conversion outcome such as 'buy or not buy', across a network of observed consumer journeys. Measurement is carried out at individual (cookie) level using parametric discrete choice modelling approaches such as Logit regression, or non-parametric machine learning algorithms such as random forests or neural networks. Outputs allow the marketer to assign credit to each element of the digital mix, addressing tactical questions such as how, where and when to spend the allocated budget and which publishers.

Secondly, we have nested (system) approaches to the marketing mix model (MMM). Here, consumer journey theories attempt to provide a complete explanation of the final purchase decision with pricing, paid, owned and earned media working together to drive demand. Measurement is typically carried out at an aggregated level, using least squares econometric methods applied to groups of consumers at store, chain, regional or market level for example. Outputs are used to quantify ROI, advise on optimal budget allocation across off and online channels and produce sales forecasts.

Finally, there are solutions that attempt to unify MTA and MMM into one framework. Treated separately, each approach is often seen as a competing attribution solution with conflicting outputs and recommendations: MTA is too narrow a representation of consumer demand, with no control for pricing, offline media and economic factors, whereas MMM is too broad to address the granular aspects of online marketing with little capacity for 'in-flight' measurement and 'real-time' optimisation. This dichotomy has led to attempts to unify the two approaches (Nail, 2015, MMA, 2021), where MMM deals with the wider macro view of consumer demand across off and online touchpoints, leaving MTA to focus on the narrower micro online view.

Notwithstanding the often-contentious issue of how best to combine the two, the very nature of customer-level MTA analysis is increasingly problematic. Firstly, the impact of traditional market-level media is notoriously difficult to measure at an individual consumer level. Secondly, the growing use of multiple online platforms has led to an increasing inability to obtain consistent user-identifiers. Thirdly, the walled-garden policies of dominant social media players, such as Facebook, have now made such identifiers unavailable altogether. Finally, and perhaps most crucially, GDPR and the ending of support for identifiers stored in third-party cookies will further impede the ability to connect on-site transactions with third-party ad placements.

In light of these issues, the focus is shifting in favour of the marketing mix model, where the aggregated nature of the data inputs can facilitate the estimation of market-level factors and capture the (de-identified) sum of individual actions across platforms. However, to constitute a valid attribution framework, any next-generation MMM solution needs to address three fundamental criteria: causal inference, short and long-term measurement, and granular 'real-time' insights. In this article, we explore these topics in detail, paving the way for more credible and actionable MMM approaches.

2. MMM and causal attribution

Marketing attribution and budget allocation rely on accurate causal attribution to each element of the marketing mix. Modern MMM attempts to inform this process via path-to-purchase theories of demand, where paid, owned, and earned media work together to drive sales. For example, a common hypothesis contends that marketing investments stimulate a journey that starts with natural online search, continues

[1] It is sometimes argued that experimental designs can handle walled garden data gaps. Even if this were possible, the problems of multiple cross-platform usage still remain.

[2] Although, note the potential for Clean Rooms (Forbes, 2021), Self-Attributing Networks (Apple) and Attribution API (Google), where some (narrowly-focused) attribution data would be available.

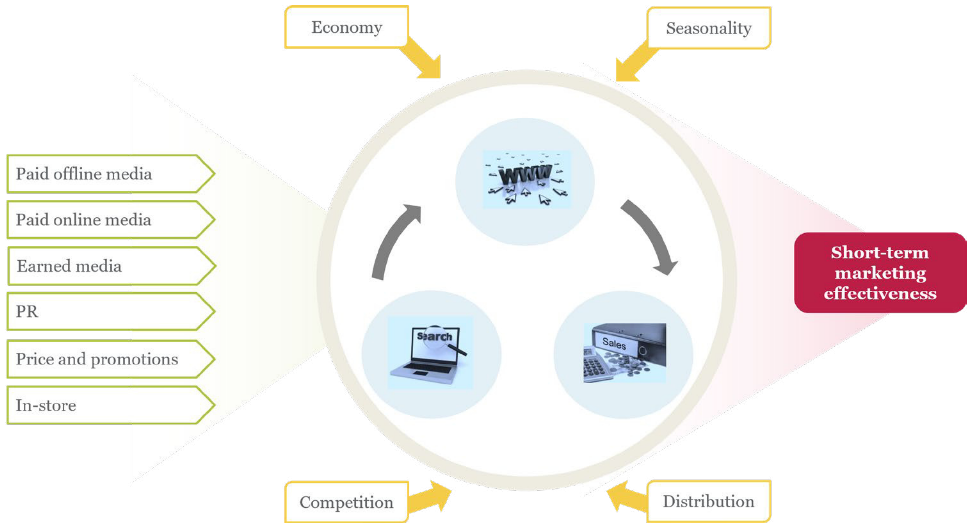

Figure 1. Network model of short-term sales

through to website research, and finally onto online and offline product purchase. This can be depicted as a Directed Acyclic Graph (DAG), illustrated in **Figure 1** and set out in equations **(1)-(3)**.

$$LnS_{it} = \dot{\alpha}_{it} + \dot{T}_{it} + \dot{\sigma}_{it} + \sum_{j=1}^{n}\sum_{k=1}^{n}\dot{\beta}_{ijk}\ln X_{ijkt} + \sum_{k=1}^{K}\dot{\delta}_{ik}D_{kt} + \dot{\gamma}_{is}lnW_{st} + \dot{\varepsilon}_{it} \tag{1}$$

$$LnW_{st} = \ddot{\alpha}_{st} + \ddot{T}_{st} + \ddot{\sigma}_{st} + \sum_{j=1}^{n}\sum_{k=1}^{n}\ddot{\beta}_{jks}\ln X_{jkt} + \sum_{k=1}^{K}\ddot{\delta}_{ks}D_{kt} + \ddot{\theta}_{s}lnNS_{t} + \ddot{\varepsilon}_{st} \tag{2}$$

$$LnNS_{t} = \dddot{\alpha}_{t} + \dddot{T}_{t} + \dddot{\sigma}_{t} + \sum_{j=1}^{n}\sum_{k=1}^{n}\dddot{\beta}_{jk}\ln X_{kt} + \sum_{k=1}^{K}\dddot{\delta}_{k}D_{kt} + \dddot{\varepsilon}_{t} \tag{3}$$

S_{it} = Off and online product sales by cross-section i (e.g., household, store, region)

W_{st} = Unique web traffic visits by source s (e.g., paid search, SEO, direct to site, display, social)

NS_t = Natural keyword search (branded and generic)

α_t = Model intercepts

T_t = Linear trends or drift

σ_t = Seasonal terms

X_{jkt} = k market level/cross-sectional own and competitor marketing variables (j=1-n) covering pricing and off and online media

D_{kt} = k interventions and external controls

ε_t = Equation error terms

β, δ, γ and θ denote estimated parameters

Equation (1) analyses the sales behaviour of groups of consumers over time and cross-section i, in terms of a range of marketing and economic covariates. **Equations (2**) and **(3)** describe the relationships between, and the drivers of, web traffic sources and natural search.[3] This type of 'nested' structure is designed predominantly to address the problems of *last-touch attribution*, helping to reattribute a part of media, such as paid search, back to sources earlier in the chain, such as TV advertising. In this way, it can be seen as an aggregate form of MTA, where appropriate credit is allocated to each touchpoint leading to improved budget allocation. However, much like MTA, this approach ignores the fundamental problem of *selection bias* leading to serious consequences for the estimated relationship between sales and web traffic sources in **Equation (1).**[4]

[3] All equations as expressed in natural logarithms, capturing non-linear response and synergies between the driver variables. Other types of transforms such as Generalised Additive Models are also used.

[4] Some practitioners bypass **Equation (2)** and use online impressions directly in **Equation (1)**. This is based on the notion that consumers do not have to click through to the site for impressions to impact demand. While this is certainly possible, the selection bias problem remains.

2.1. The endogeneity problem

Selection bias arises when (part of) the difference in the 'treatment' outcome (sales) is caused by a factor that predicts the likelihood of selection into treatment (paid search) rather than due to the treatment itself. That is, consumers with a greater propensity to buy predict the level of search traffic, which in turn predicts the sales outcome. In this way, a large proportion of site visits are simply an artefact of the sales process. This creates an *endogeneity* or *identification* problem, leading to biased estimates of the traffic-sales impact of **Equation (1)** and all marketing effects that work through it. Consequently, even if offline TV advertising does lead to more paid search activity, it does not necessarily mean it drives *incremental* sales in this way.

The textbook solution is Instrumental Variable (IV) estimation, where the causal effect of an independent variable (paid search) on an outcome variable (sales) is estimated using an instrumental variable z which affects sales only through its impact on search (the *exclusion* principle). If successful, z provides the necessary *exogenous* variation in search, such that the outputs are more akin to those of experimental trials.[5] However, valid instruments are notoriously difficult to find. Consequently, many alternative solutions have been proposed ranging from the difference in differences, regression discontinuity designs, and Heckman correction, through to Latent IV (Ebbes *et al*., 2009), Gaussian Copulas (Park & Gupta, 2012), DAG analysis (Chen *et al*., 2018, Pearl, 2000) and incorporating experimental results as Bayesian priors (Ugena *et al*., 2021).[6] Whichever route is taken, the message is clear: for meaningful attribution, the chosen identification scheme needs to be clearly specified as part of any MMM engagement.[7]

3. Short and long-term marketing effectiveness

For a complete view of marketing ROI and optimal allocation, marketing mix models need to reflect both short and long-term marketing effects. Short-term effects explain mean-reverting or transitory sales variation. Long-term effects explain persistent changes in underlying base sales, reflecting permanent additions to the loyal customer base. Measuring the true long-run impact of marketing investments, therefore, requires a focus on the base sales component of the mix model.

3.1. The standard approach

Standard mix models use ordinary or generalised least squares regression techniques, with fixed or deterministic baselines, and focus solely on short to medium-term sales effects with stationary Adstock transforms. Consequently, all such models fail to reflect any persistent changes in core brand preferences by construction. A popular remedy simply adds attitudinal brand metrics to the short-term model together with sub-models in terms of advertising variables. The indirect effects of advertising on sales are then interpreted as long-term effects.[8] However, this approach is flawed in several respects.

[5] On face value, the nested mix model structure of **Equations (1)-(3)** *appears* to fit the bill. Provided (at least one of) the variables driving web traffic satisfies the exclusion principle, the web traffic fitted values could be substituted into the sales equation to give a two-stage least squares estimate. However, X_t and D_t generally affect both web traffic and sales and cannot serve as valid instruments.

[6] Experimental priors in MMM rely on valid A/B testing or 'lift' studies for all endogenous variables. However, these are rarely available as part of the routine data collection process.

[7] Note that 'causal AI' methods seek to automatically identify DAGs such as **Figure 1**. However, since human context is always required, such techniques 'do not yet work as stand-alone methods for causal learning' (Peters *et al.*, 2017). Furthermore, there is no one unique chain. Endogeneity bias stems from ignoring the simultaneous likelihood of all other plausible DAGs, leading to the correlation between search and the error term in **Equation (1)**. We need to control for all paths to help identify causal effects.

[8] Alternative 'long-term' approaches simply extend the short-term structure, either by adding Adstocks with very high retention rates or multiplying the short-term effects by an *ad hoc* scaling factor.

1. Ignores the fundamentals of time series econometrics

If (observable) brand-building effects exist, sales should exhibit evolutionary behaviour.[9] If not, then the impact of brand metrics on sales can only be a short-term relationship by definition. If, on the other hand, sales are evolving, we cannot just run simple regressions of sales on marketing and brand metrics. Firstly, if marketing and brand metrics are stationary, then the mix equation is unbalanced, and we must first-difference sales. Alternatively, if brand metrics are also evolving, then there is potential for spurious regression problems. Consequently, brand equity metrics must also be first-differenced, and valid cointegrating relationships between sales and brand metrics need to be incorporated.

2. Mindset metrics are regressed directly on short-term sales

A plausible theory of brand-building needs to link the long-term brand preferences embodied in mindset metrics directly to the long-term purchase demand revealed through base sales. This follows naturally from the fact that base sales and attitudinal data both represent brand health (*inter alia*, Kamakura & Russell, 1993, Hanssens *et al.*, 2014). As such, they are essentially two sides of the same coin. Therefore, the use of actual sales is inconsistent and obscures long-term movements risking contamination with short-term transactional effects.

3. Does not reflect the brand-building process

Simply adding brand metrics as additional regressor(s) precludes feedback between (base) sales, earned media, and other long-term drivers. Feedback effects mimic word-of-mouth as consumers talk about brand experiences leading to new trialists and growth of the loyal customer base, which wears in over time. Only by identifying these *endogenous* relationships in a suitable network structure can we estimate the true incremental long-term impact of marketing on base sales via brand perceptions.

3.2. An alternative approach

To resolve these issues, brand metrics need to be linked directly to variation in base sales in a long-term network model of brand-building. To achieve this, the marketing mix model needs to be re-cast in a form that allows measurement of both short-term sales and long-term base variation. One candidate is the Unobserved Component Model (Harvey, 1989), illustrated in Cain (2005, 2008), where sales behaviour is decomposed into a trend, seasonal, regression effects and measurement error. This re-writes the marketing mix model **Equations (1)-(3)** as:

$$LnS_{it} = \dot{\mu}_{it} + \dot{\sigma}_{it} + \sum_{j=1}^{n}\sum_{k=1}^{n}\dot{\beta}_{ijk}\ln X_{ijkt} + \sum_{k=1}^{K}\dot{\delta}_{ik}D_{kt} + \dot{\gamma}_{is}lnW_{st} + \dot{\varepsilon}_{t} \quad 1(a)$$

$$LnW_{st} = \ddot{\mu}_{st} + \ddot{\sigma}_{st} + \sum_{j=1}^{n}\sum_{k=1}^{n}\ddot{\beta}_{jks}\ln X_{jkt} + \sum_{k=1}^{K}\ddot{\delta}_{ks}D_{kt} + \ddot{\theta}_{s}lnNS_{t} + \ddot{\varepsilon}_{st} \quad 2(a)$$

$$LnNS_{t} = \dddot{\mu}_{t} + \dddot{\sigma}_{t} + \sum_{j=1}^{n}\sum_{k=1}^{n}\dddot{\beta}_{jk}\ln X_{kt} + \sum_{k=1}^{K}\dddot{\delta}_{k}D_{kt} + \dddot{\varepsilon}_{t} \quad 3(a)$$

$$\mu_t = \mu_{t-1} + \lambda_{t-1} + \eta_t \quad 4(a)$$

$$\lambda_t = \lambda_{t-1} + \xi_t \quad 4(b)$$

$$\sigma_t = -\sum_{j=1}^{p-1}\sigma_{t-j} + \kappa_t \quad 4(c)$$

The intercepts α in each equation are replaced with a time-varying (stochastic) trend μ_t comprising two components. **Equation 4(a)** allows the underlying level of each time series to follow a random walk with a growth factor λ_t analogous to the conventional trend term T. **Equation 4(b)** allows λ_t to also follow a random walk. Depending on the estimated values of the

[9] The absence of evolution does not imply the absence of brand-building *per se*: merely that it is *unobservable*. Evolution could be offset by customer churn rendering observed sales stationary.

covariance parameters η_t and ξ_t, the system can accommodate both stationary and non-stationary product demand allowing the data to decide between them. **Equation 4(c)** specifies seasonal effects, which are constrained to sum to zero over any one year. If κ_t is zero, then seasonality is deterministic.

Equations 1(a)-4(c) provide a direct separation of sales behaviour into short and long-term components. The estimated regression parameters capture short-term (transitory) marketing effects, informing short-term ROI and budget allocation decisions. Long-term effects can then be analysed through a network model of the permanent sales component μ_t in terms of consumer brand perceptions and external long-term controls (Cain, 2010, 2022). A representative example is illustrated in **Figure 2**, where marketing investments stimulate brand awareness, drive brand consideration, and increase social media interest leading to underlying base sales growth. If we can show that marketing significantly impacts the permanent (baseline) component, then we can state that marketing campaigns have persistent long-term effects, as existing purchase incidence increases and/or new buyers are converted into permanent loyal consumers. These effects are then combined with short-term effects to provide total ROI and budget allocation recommendations.

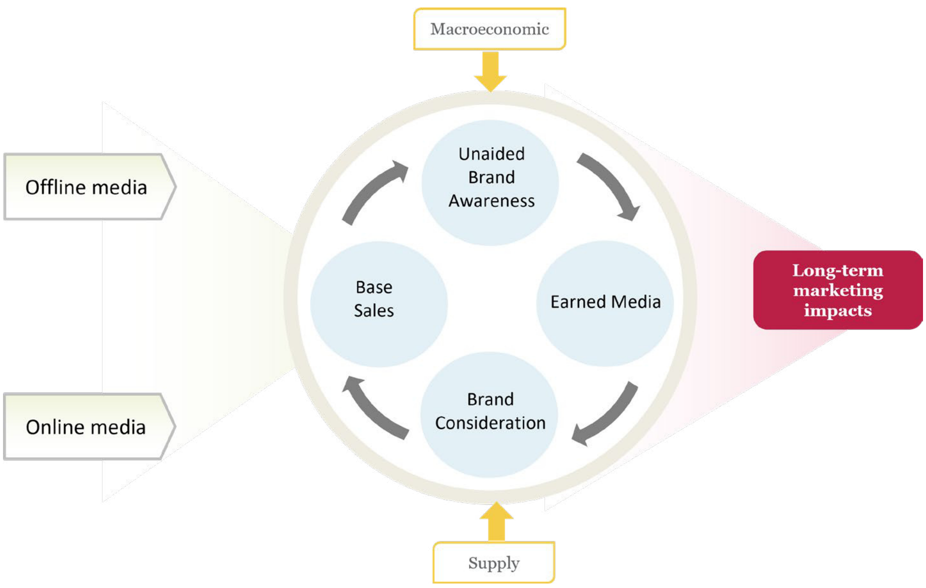

Figure 2. Network model of long-term sales

Estimation of the long-term network model requires a suitable systems approach to capture the long-term relationships between the nodes of Figure 2 and the persistent brand-building role of media. Popular systems approaches are Path Models or Structural Equation Models. However, these frameworks are typically static and ignore the dynamic relationships between the network variables. As such, they are unsuitable for long-term trend and cointegration analysis and cannot measure feedback between the nodes and the dynamics of how brand-effects *wear in* over time. To overcome these issues, a dynamic systems approach such as a Vector Autoregression (VAR) is required (*inter alia*, Hendry, 1995), written as a cointegrated Vector Error Correction Model (VECM):

$$\Delta lny_t = \Psi_1 \Delta lny_{t-1} .. \Psi_{l-1} \Delta lny_{t-l+1} + \alpha\beta' lny_{t-1} + \Omega_k(L) lnx_{kt} + \Upsilon_k D_{kt} + \varepsilon_t \quad (5)$$

where y_t denotes a vector of n endogenous variables capturing base sales and path-to-purchase or brand-building 'steps', x_t denotes a set of k marketing variables with lags L and D_k denotes a set of dummy variable events. The $\alpha\beta' lny_{t-1}$ term represents the error correction component, comprising r cointegrating (equilibrium) relationships β between the nodes and associated error-correction parameters α. With n endogenous variables, there may be up to n-1 such relationships with a minimum of one common trend driving the non-stationary (brand-building) properties of the system.

Equation (5) is then estimated using the Johansen technique (1988) and identified using either a Cholesky decomposition, restrictions based on economic theory or instrumental variable techniques (Juselius, 2006). Once identified, impulse response analysis traces out the dynamic long-term base sales impact of changes in brand metrics and earned media. The long-term impact of marketing activity x_t then cumulates indirectly and permanently into the level of base sales.

3.3. Worked example

The complete short and long-term modelling approach is formally demonstrated in Cain (2022). Here we present a simple example to illustrate the principles involved. We first take daily data for sales, web traffic, and natural branded search for a seasonal brand, together with a range of off and online marketing factors, pricing, monthly unaided awareness data, and an index of monthly business economic activity. Monthly business activity and awareness data were then disaggregated to daily level and introduced directly into the sales equation (1). Standard OLS (fixed base) estimation gives an awareness coefficient of 0.045 but is insignificant with a t-ratio of 1.1. Furthermore, the base price coefficient is positive and a Durbin Watson (DW) statistic of 1.08 indicates significant model error autocorrelation. The implication is that neither awareness, price, nor economic growth manages to adequately capture long-term sales movements.[10]

We then applied the UCM framework of 1(a)-4(c) and aggregated the extracted baseline to the weekly frequency - illustrated in Figure 3 alongside unaided awareness, base price, and business economic activity (BEA). Standard ADF tests indicate that all are non-stationary

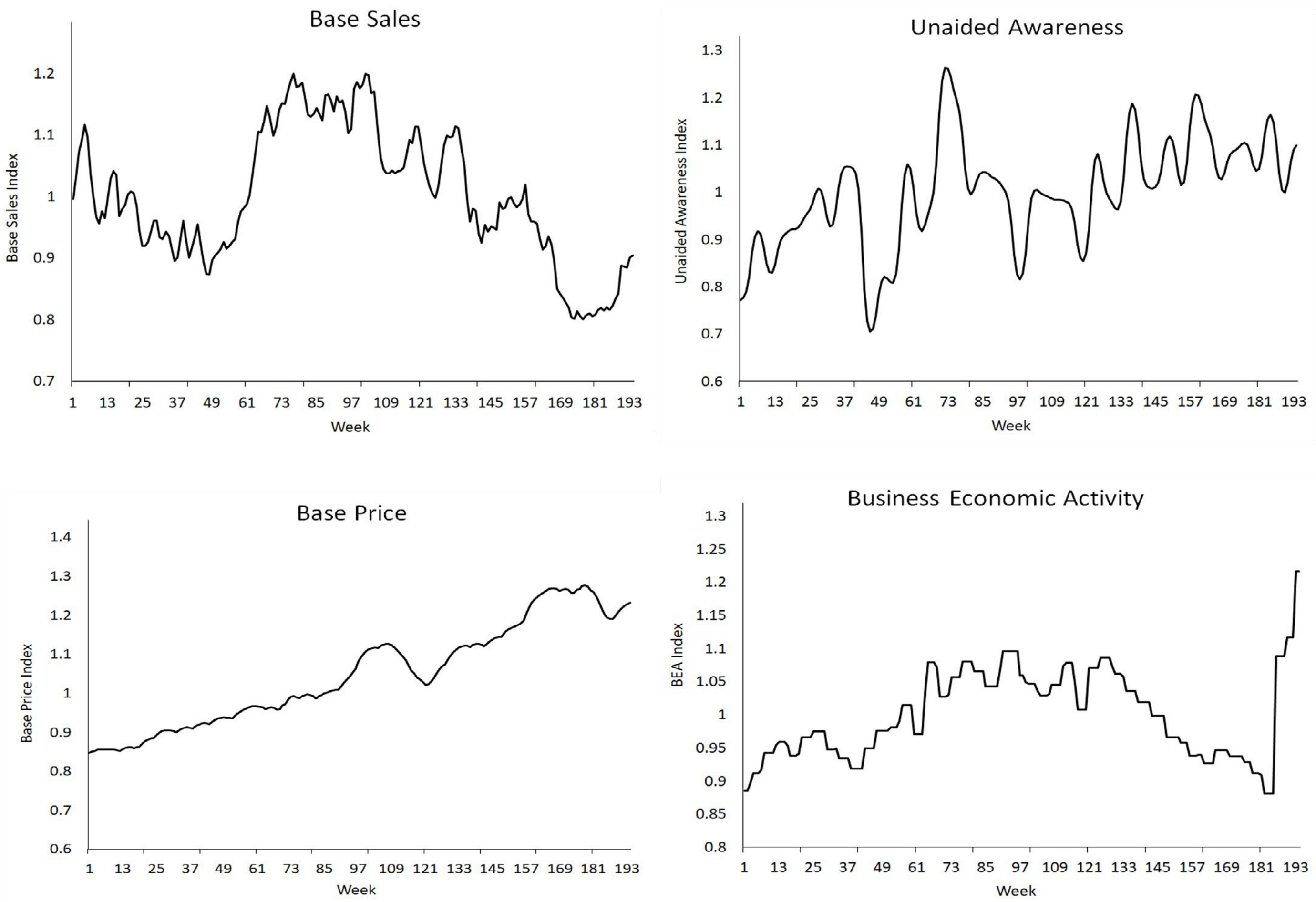

Figure 3. Extracted base sales, brand awareness and controls

[10] An AR(1) error structure improves autocorrelation with a DW stat of 2.01. However, the awareness coefficient is -0.03 and insignificant. Weekly frequency models made little difference to the results.

I(1) series. Equation (5) was then estimated with two lags of the endogenous variables to ensure well-behaved residuals. Marketing regressors x_{kt} comprise paid TV GRPs and social media commentary (earned media).

The cointegrating relationship between the variables is illustrated in the left-hand panel of **Table 1**. This captures the underlying equilibrium (attractor) relationship between base sales, price, unaided awareness, and business activity, with feedback reflected in the alpha (error correction) parameters. The full long-term (impulse response) coefficients are illustrated in the right-hand panel, where the first row shows the final (permanent) elasticity of a 1% impulse in unaided awareness on base sales of 0.13 with a significant t-ratio of 2.7. Note too, that the final long-term effects of base price and economic activity are also correctly signed and significant.

Figure 4 then illustrates the corresponding pattern of dynamic adjustment of base sales to unaided awareness, where the full impact wears in over approximately 16 weeks.[11]

The corresponding VECM is given in **Table 2**, which shows how marketing investments impact the dynamic adjustment of each of the network variables. Here TV and earned media impact unaided awareness with elasticities of 0.002 and 0.022, respectively. Weighted by the long-term impact of unaided awareness on base sales gives final base elasticities of 0.0003 and 0.003. These are used to quantify long-term base contributions over the sample and extrapolated over a 3-5 year forecast horizon. Combined with the short-term effects from the UCM, this provides total ROI and budget allocation recommendations.

Regressor	CV	Alpha
Base sales	1	-0,151 (-3.8)
Base Price	-0,588 (5.1)	0
Awareness	**0.280 (2.1)**	0.082 (2.10)
BEA	1.81 (4.9)	0.060 (3.70)

Equation	$\sum \hat{\varepsilon}_{Base}$	$\sum \hat{\varepsilon}_{Price}$	$\sum \hat{\varepsilon}_{Aware}$	$\sum \hat{\varepsilon}_{BEA}$
Base sales	0.455 (6.40)	-0.688 (-3.6)	**0.130 (2.7)**	0.960 (4.0)
Base Price	-0,005 (-0.1)	1.37 (10.0)	0.006 (0.20)	-0.020 (-0.10)
Awareness	0.196 (1.80)	0.080 (0.30)	0.889 (12.5)	-0.721 (-1.21)
BEA	0.221 (5.70)	0.054 (0.50)	-0.063 (-1.1)	0.639 (4.9)

Table 1. Cointegrating economic structure and impulse response matrix

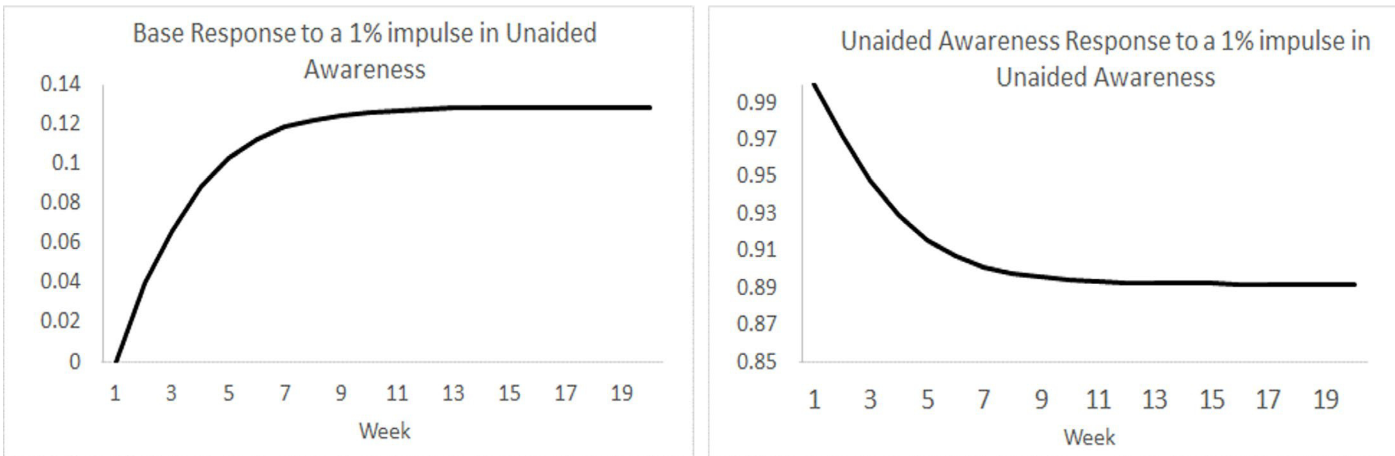

Figure 4. long-term base sales adjustment

[11] Note that these results imply that long-term effects are under-estimated using the traditional approach. However, the relationship(s) between brand metrics and sales can often be over-estimated if the long-term network dynamics are not accounted for. It depends on the data and model structures, requiring careful modelling on a case-by-case basis.

Equation	$\Delta Base_t$	$\Delta Price_t$	Aware	ΔBEA_t
Intercept	-3.14 (-3.76)	0.002 (1.31)	1.69 (2.09)	1.25 (3.68)
$\Delta Base_{t-1}$	-0.188 (2.63)	-	-0.130 (-1.86)	0.019 (0.94)
$\Delta Price_{t-1}$	-0.334 (-1.82)	0.28 (3.96)	-0.15 (-1.01)	-0.054 (0.98)
$\Delta Aware_{t-1}$	-	-	-0.014 (1.01)	-
ΔBEA_{t-1}	-0.365 (-2.06)	-	0.221 (1.26)	0.273 (3.78)
ECM_{t-1}	--0.151 (-3.76)	-	0.082 (2.10)	0.060 (3.68)
TV	-	-	0.002 (2.10)	-
ΔEM_t	-	-	0.022 (2.80)	-

Table 2. dynamic network adjustment

4. MMM and tactical planning

It is often argued that MMM is too slow and lacks the necessary granularity to handle the tactical and 'real-time' attribution problems that solutions such as MTA purport to solve. Given the detailed and rapid solutions marketers have now come to expect, any viable MMM framework needs to be able to rise to the challenge.

4.1. Tactical decision making

Whereas MTA focuses on cookie-level data over very short time windows, MMM can provide similar learnings across both off and online through higher frequency time series data. The process is illustrated in **Figure 5**, where the dynamic aggregated framework set out in Sections 2 and 3 is first estimated at daily level, providing trend and seasonal factors and incremental contributions for off and online media investments. We then take the hourly data for sales and remove the (proportions of) trend and the contributions for all variables not available at an hourly level. We then model the remaining portion of hourly sales in terms of the detailed 'sub-tactic' elements of all off and online media variables – subject to the estimated 'upper-level' contributions. In tandem with offline media effectiveness, media synergies, and long-term brand-building of the main daily MMM model, this type of approach can then provide granular online media effectiveness by day-part, ranking by publisher, placement, and web page.[12]

4.2. Real-time attribution

Daily network UCM marketing mix models, as set out in Sections 2 and 3 and summarised in the left-hand panel of **Figure 5**, typically take approximately eight weeks to build – depending on the number of models and cross-sections. Updates in response to business needs, or potential structural/parameter changes typically take place every three to six months. Hourly level models – summarised in the right-hand-

[12] Since consumer cohorts are time-based rather than geography-based, this approach is not subject to the matching problems typically faced with cookie or household-level models, where a mapping between individual data and more aggregated (mass-market) offline data is required.

side panel of **Figure 5** – are constructed on the last three months of data used to build the main daily models, with contribution/parameter constraints set from the daily models to ensure consistency. The hourly models are then updated weekly to deliver the types of rapid in-campaign attribution illustrated in **Figure 6**.

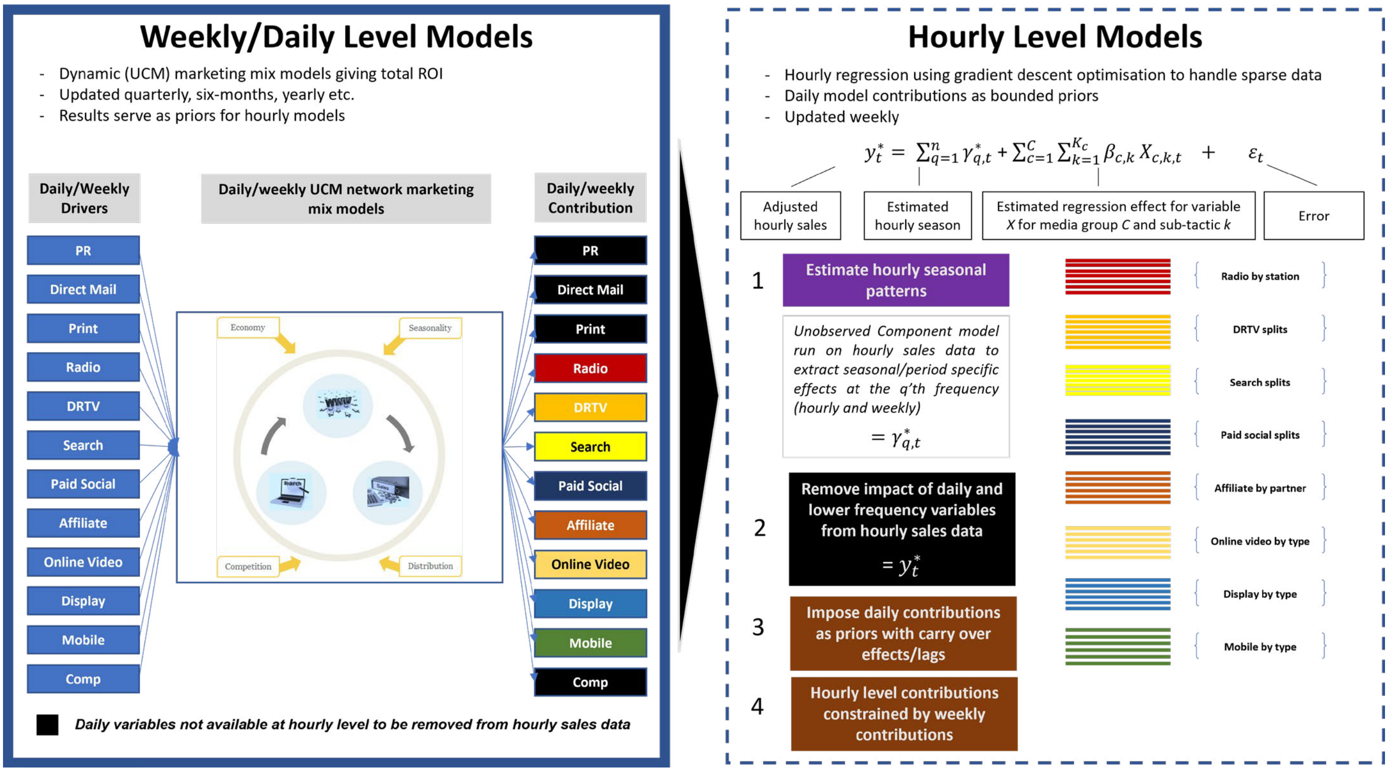

Figure 5. High frequency MMM

Figure 6. Campaign response attribution

5. Conclusions

With the potential demise of MTA, the focus is once more back on the marketing mix model as an attribution framework. However, to be useful and live up to the exacting standard that marketers have come to expect, any next-generation MMM approach needs to satisfy three fundamental criteria.

Firstly, to serve as a true attribution solution, MMM needs to focus on causal estimation methods. Too often, we see reliance on consumer journey solutions to address the problems of last-touch-attribution. However, much like micro-level MTA methods, these ignore the endemic selection bias in many online media – leading to endogeneity bias and misallocation of the marketing mix. The growing popularity of automated machine learning (ML) approaches to the mix model only serves to exacerbate this problem. To address this issue, all MMM work - whether based on regression, neural nets, or other ML methods - needs a transparent identification scheme to isolate true incrementality.

Secondly, MMM needs to quantify the long-term (base-building) effects of marketing and so inform brand-building strategy. Standard approaches are simply not set up to measure these effects, with fixed baselines and a focus on short to medium-term lag structures or Adstocks. Alternative time series structures are required that can quantify both short and long-term (base) variation – coupled with dynamic network models that can explain the causes of base variation and the economics of brand-building.

Finally, next-generation MMM needs to deliver near 'real-time' granular insights on marketing ROI and optimal budget allocation. Suitably identified high-dimension mix models – by day and hour – can fit the bill. This can provide many of the claimed benefits of MTA, such as online media effectiveness ranking by publisher and placement, with the added benefit of controlling for the wider economic environment and quantifying the contribution of pricing and offline media.

6. Acknowledgements

The authors are grateful to David Dixon for helpful comments and to Arnab Pal and Kunal Rambhia for input into section 4 and model outputs of **Figure 6**.

References

1. Cain, P.M. (2005). "Modelling and forecasting brand share: A dynamic demand system approach". *International Journal of Research in Marketing*, 22(2), 203–220.

2. Cain, P.M. (2008). "Limitations of conventional market mix modelling". *Admap*, 48-51.

3. Cain, P.M. (2010). "Marketing Mix Modelling and Return on Investment". In: P. J. Kitchen (ed.) *Integrated Brand Marketing and Measuring returns*, Palgrave Macmillan, 94-130.

4. Cain, P.M. (2022). "Modelling short and long-term marketing effects in the consumer purchase journey". *International Journal of Research in Marketing*, 39(1), 96–116.

5. Chen, A., Chan, D., Koehler, J., Perry, M. L., Wang, Y., Sun, Y., & Jin, Y. (2018c). "Bias Correction For Paid Search In Media Mix Modeling". *ArXiv (Cornell University)*.

6. Ebbes, P., Wedel, M., & Böckenholt, U. (2009). "Frugal IV alternatives to identify the parameter for an endogenous regressor". *Journal of Applied Econometrics, 24*(3), 446–468.

7. Hanssens, D. M., Pauwels, K., Srinivasan, S., Vanhuele, M., & Yildirim, G. (2014). "Consumer Attitude Metrics for Guiding Marketing Mix Decisions". *Marketing Science, 33*(4), 534–550.

8. Harvey, A. C. (1989). "Forecasting, Structural Time Series Models and the Kalman Filter". *Cambridge University Press*.

9. Hendry, D. F. (1995). "Dynamic econometrics". *Oxford University Press*.

10. Johansen, S. (1988). "Statistical analysis of cointegration vectors". *Journal of Economic Dynamics and Control, 12*(2–3), 231–254.

11. Juselius, K. (2007). "The Cointegrated VAR Model: Methodology and Applications" (Advanced Texts in Econometrics). *Oxford University Press.*

12. Kamakura, W. A., & Russell, G. L. (1993). "Measuring brand value with scanner data". *International Journal of Research in Marketing, 10*(1), 9–22.

13. Mobile Marketing Association (MMA) (2021). "MTA is dead, long live MTA".

14. Nail, J. (2015). "Introducing Unified Marketing Impact Analytics". *Forrester*.

15. Park, S, and Sachin G. "Handling Endogenous Regressors by Joint Estimation Using Copulas." *Marketing Science, 31*, no. 4 (2012): 567–86.

16. Pearl, J. (2000), "Causality", New York: *Cambridge University Press*.

17. Peters, J., Janzing, D., & Scholkopf, B. (2017). "Elements of Causal Inference: Foundations and Learning Algorithms" (Adaptive Computation and Machine Learning series). *The MIT Press*.

18. Schoen, M. (2021). "Can Data Clean Rooms Be The Answer For Privacy-Safe Marketing?" *Forbes. (CLE) blending. Deloitte working paper*.

19. Ugena, C.R, Conde de Simón, M, Answer, M. and Fabrizio, D., (2021), "Exploring Marketing Mix Modeling (MMM) and Conversion Lift Experiment (CLE) blending", *Deloitte working paper.*

Author

Peter Cain is executive partner and co-founder at Marketscience, an award-winning independent consultancy specialising in advanced econometric modelling, statistical analysis and bespoke advice on marketing investments. Prior to this, Peter was the founder and CEO of Marketscience Consulting, established in 2012 with the goal of blending academic, commercially relevant analytics and strategic advice for business. He has more than 20 years of commercial and academic experience in economics and marketing science designing econometric business solutions for blue-chip companies and organizations across a wide range of industries. He writes extensively on economics and econometrics in marketing and published in top peer-reviewed journals. Before marketing research, Peter was in academia, specializing in monetary economics and econometrics. He holds BSc and MSc degrees in Economics from the University of Warwick, and PhD in Monetary Economics from the University of Nottingham.

peter.cain@marketscience.co

Nitesh Sahay is a Partner at Marketscience. Prior to this, Nitesh worked at different roles in analytics at GE Capital, Symphony Marketing Solutions, Datamonitor and Ninah (The Publicis Groupe). He has more than 20 years of commercial experience across multiple industry verticals and clients in Fortune 500 companies working on techniques such as Marketing Mix Models, Market Structure, Customer Segmentation and other predictive methods. He has written 2 papers on economics published as Working Paper series and another in an edited book. Before marketing research, Nitesh was in academia, specializing in macroeconomics and econometrics (computable general equilibrium models). He holds a BA in Economics from Delhi University and MA and MPhil degrees in Economics from the Jawaharlal Nehru University, New Delhi.

nitesh.sahay@marketscience.co