Dynamic marketing mix modelling and digital attribution[1]

P.M Cain

Introduction

Digital media attribution aims to identify the combination of online marketing activities and touchpoints contributing to online sales conversion. Given the availability of unique user-identifiers, analysis conventionally traces the actions of single individuals. Traditional media attribution, on the other hand, evaluates the offline sales impact of offline marketing investments. Measurement is typically carried out at an aggregated level, using marketing mix analysis applied to groups of consumers, either at store, chain or market level.

With the advent of multi-channel marketing, comes the need to measure the sales impact of inherently micro-focused digital media alongside more macro-oriented traditional advertising. Consequently, any analytical approach that aims to incorporate both elements inevitably involves a degree of data aggregation or disaggregation depending on whether we adopt a macro or a micro route.

Disaggregated approaches are problematic for two reasons. Firstly, the growing use of multiple online platforms has led to an increasing inability to obtain consistent user-identifiers. Secondly, in the absence of single-source panel information, the impact of traditional market-level media is notoriously difficult to measure at disaggregated levels. Aggregated approaches can help resolve both problems by consolidating across consumers, facilitating the estimation of market-level factors and capturing the sum of individual actions across platforms.

This article presents an aggregate modelling framework for traditional and digital marketing attribution. The model structure is based on a theory of consumer purchase behaviour that naturally combines off and online marketing touchpoints, with response parameters estimated using appropriate dynamic econometric techniques. Outputs provide many managerial benefits, ranging from accurate ROI and media planning inputs through to simulation and demand forecasting.

The consumer purchase journey and the marketing mix model

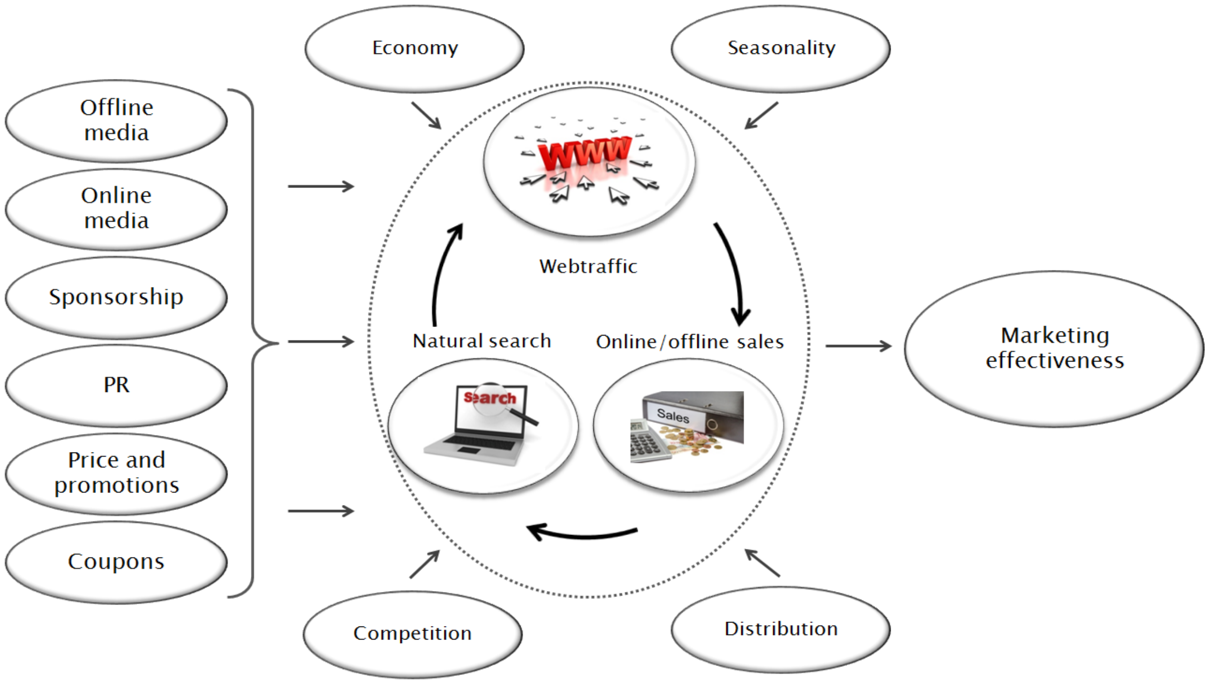
Marketing mix modelling (MMM) is a process that uses economic theory and econometrics to quantify the effectiveness of marketing activities on business performance. Model outputs are used to evaluate Return on Investment (ROI), allocate marketing resources for business growth and forecast the likely evolution of consumer demand.

In the modern digital economy, offline and online activity are inherently linked in the consumer journey to product purchase. For example, traditional marketing investments typically stimulate a purchase cycle that starts with natural online search, continues through to website research and finally onto online and offline purchase. This process is illustrated in

[1] Admap 2014

Figure 1, where digital and traditional marketing investments impact each step of the journey with controls for all external and macroeconomic drivers.

Figure 1: offline-online ecosystem



Simplified approaches to aggregate digital attribution typically include online marketing directly in conventional single equation models of product or brand sales. However, Figure 1 indicates that this is inadequate for two reasons. Firstly, it ignores the role of traditional media in driving online activity such as PPC, natural search and display, underestimating the true effectiveness of offline marketing. Secondly, and as a direct consequence, ignoring such offline and online interactions falls prey to the last-click attribution problem thus overestimating the impact of online marketing. To address these issues, the design and estimation of the marketing mix model needs to be re-structured accordingly.

Model design

Correct attribution of all traditional and digital marketing investments requires a model framework that reflects the multiple touchpoints and steps on the road to purchase. Figure 1 illustrates an example of a three-tier model structure, focusing predominantly on paid and owned (website) media. This can be extended to capture the role of word-of-mouth or earned media via incorporating additional steps such as Twitter feeds and Facebook posts. Model design then focuses on the flow and drivers of the hypothesised journey, together with the nature of consumer response to marketing activities.

Natural online search is specified as a function of offline marketing investments, online display and other relevant factors. Web traffic visits, typically split by source in order to capture all digital media interactions, are driven in terms of online and offline marketing investments together with natural search. Finally, sales are written as a function of product price, offline drivers and web traffic. In this way, it is clear that each step of the journey is a behavioural outcome to be explained, with search and web traffic 'handed down' as a driver into the next level.

Consumer response to marketing investments is determined by the functional relationship between each outcome and the explanatory variables and lies at the heart of business planning decisions such as optimal marketing budget allocation. Advertising response is generally concave reflecting decreasing demand elasticity with increasing media weight. This is a natural consequence of diminishing returns following repeated exposure to a given audience. Price response, on the other hand, is typically convex reflecting increasing returns, with demand rising exponentially to infinity as price falls to zero.

Model estimation

Estimation concerns quantification of the model parameters: that is, measurement of the driver impacts and response curves set out as part of the model design. This is where econometrics enters the picture: a statistical regression based procedure that applies the theoretical model structure to in-market data. Correct estimation needs to account for two key features.

Firstly, model design has established that web traffic and online search act as both dependent (explained) and independent (explanatory) variables in the system. This poses a problem for classical econometric analysis, where valid causal inference requires that each explanatory variable is exogenous or fixed outside the model. Variables that are determined inside the system are defined as endogenous. If used as contemporaneous explanatory variables, this confounds correlation and causation leading to biased response parameters.

One solution is to treat Figure 1 as a structural equation path model with a recursive structure. Under these circumstances, endogeneity is not an issue and Figure 1 can be directly interpreted as a causal chain with each step estimated separately by single equation regression techniques. However, it is highly likely that feedbacks exist between outcomes, particularly as we aggregate over consumers and across time. Consequently, alternative estimation approaches are often required such as network analysis or Instrumental Variable techniques involving joint estimation of sales, web traffic and natural search as a complete system. Only then may endogenous drivers be given a clear causal interpretation.

Secondly, economic theory only specifies the appropriate current period relationships between dependent and independent variables. This tells us nothing about how consumers adjust their behaviour over time in response to changing marketing investments and evolving brand tastes. However, marketing mix models are essentially time series models where dynamic behaviour plays an integral role in explaining the full impact of marketing strategy.

Dynamic specification is conventionally handled via lags of the advertising variables in the form of Adstocks. The idea is intended to capture how consumers adjust their purchase behaviour over time in response to media investments. However, this only captures short to medium term effects and ignores longer-term adjustment in baseline behaviour. Baseline adjustment is crucial for quantification of the long-run impact of marketing investments and is captured with the dynamic time series approach to the mix model ('Limitations of conventional marketing mix modelling', Cain, P.M., Admap, April 2008).

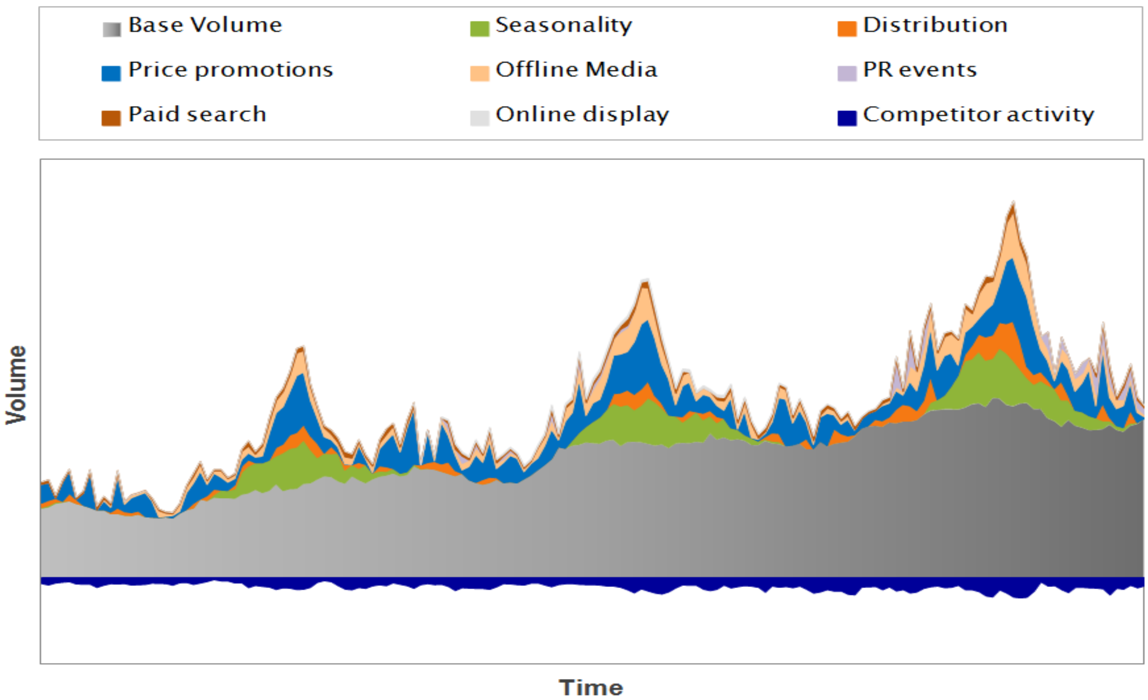
Managerial benefits

The proposed framework provides a dynamic approach to aggregated traditional and digital marketing attribution. The outputs inform the business planning process in four key areas.

1. Traditional and digital ROI

Combining offline and online marketing investments into one dynamic system provides a complete aggregated view of product sales attribution as illustrated in Figure 2. Incremental volumes allow calculation of short-term traditional and digital marketing ROI. Evolutionary base volume represents the dynamic trend component of the sales time series, capturing the long-term effects of marketing on brand sales. This is of particular importance in a digital setting, where the viral nature of online sharing implies that social media has potential long-term trend-setting effects.

Figure 2: sales attribution

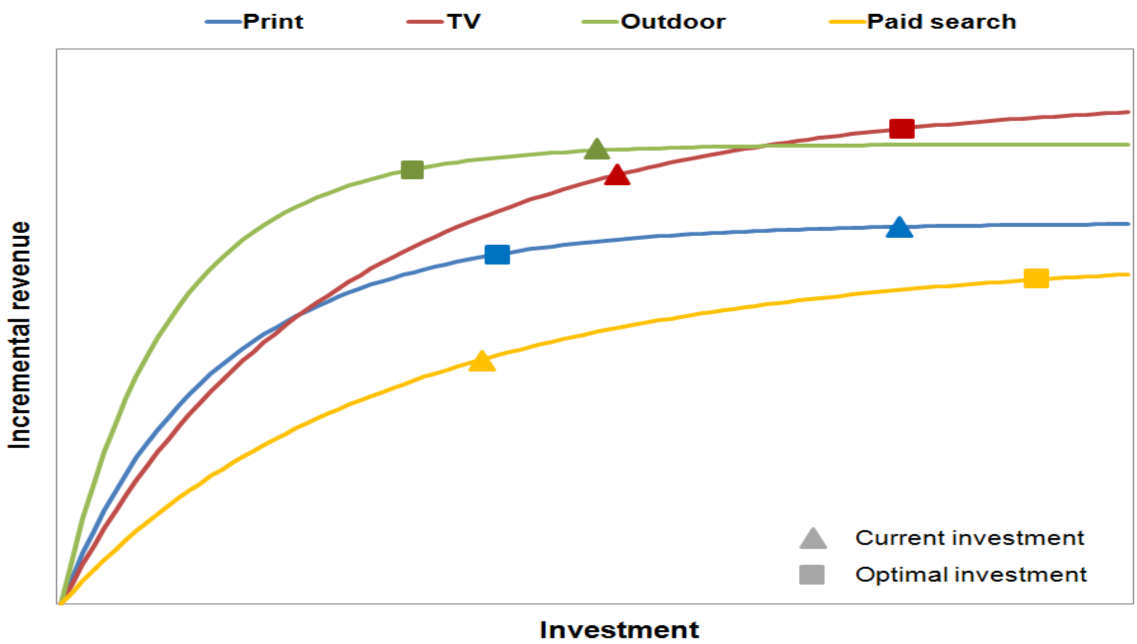


2. Marketing budget allocation

Marketing ROI provides a valuable snapshot of marketing performance. However, it is based on historical investments and is inherently backward looking. Optimal allocation of planned marketing budgets requires a view into the additional sales revenue from the next pound spent on marketing investments. This is given by estimated response functions, such as those illustrated in Figure 3 for above the line media.

Marketing budget is allocated across channels according to marginal rates of revenue response with a view to maximisation of a defined business objective. This typically involves either revenue maximisation under fixed marketing budgets, minimisation of total budget to achieve desired revenues or straight profit maximisation. Promotional response is generally incorporated to extend optimal allocation across both above and below the line components. For long-term brand building objectives, short-term response curves are then adjusted for any measured long-run (baseline) effects.

Figure 3: media response curves



In each optimisation scenario, specific channel spending constraints are generally required. For example, upper limits to paid search budgets are often set due to a limit in the number of keywords that can be purchased. On the other hand, lower limits are often imposed since many marketing channels require minimum threshold levels of investment, below which response is negligible. Allocation results then guide the optimal division of planned marketing spend between traditional and digital components and are typically used to inform the annual media planning process. Once annual channel budgets are set, planners must then determine how best to distribute investments throughout the year.

3. Media lay-down

Optimal allocation of a given channel budget over time is mainly governed by the nature of the estimated response curve. For example, the concave response forms of Figure 3 imply that a continuous drip policy is the optimal strategy: specifically, an initial investment followed by continuous weekly top-up to maintain awareness. The rate of media decay determines the weight of top-up, with adjustments made for variation in media costs, base sales and seasonal demand in order to maximise incremental volume.

In practice, however, such optimal strategies are often difficult to apply. Annual channel budgets are typically insufficient to spend continuously over the year and maintain weekly reach thresholds. Under these circumstances, campaigns are planned over fewer weeks, targeting particular areas of the year where seasonal base demand is higher and exploiting interactions between traditional and digital marketing to help guide holistic communication strategies. Once weekly channel lay-downs are set, media planners can then determine the best daily and day-part allocation for optimum reach and frequency.

4. Simulation and forecasting

Simulation concerns 'what-if' scenario testing, where the dynamic model structure is used to forecast the revenue and profit impacts of budget allocation recommendations and media plans. Results are typically used to evaluate planned media strategies compared to the optimal mix or simulate the lay-down structure and budgets required to achieve desired revenue and profit objectives.

Focusing solely on media, however, presupposes fixed levels of other key business drivers. A more inclusive approach incorporates planned marketing, pricing and distribution strategies together with assumed competitor actions and levels of macroeconomic activity. Simulation then evaluates the combined demand impact of varying levels of all the driver variables. This leads to accurate demand forecasting, playing a key role in supply chain management.

Conclusions

In order to maintain its relevance in the modern digital economy, the traditional marketing mix model needs to be re-structured to incorporate digital media. This article has presented a general dynamic approach that explicitly models the offline-online consumer purchase journey. The framework provides several commercial benefits.

Firstly, the model structure accurately captures interactions between traditional and digital media, together with digital media interactions across web-traffic sources. As a result, each element of the marketing mix receives the appropriate credit for final off and online sales conversion. Secondly, the full interactive approach leads to more accurate ROI measurement and improved marketing budget allocation. Thirdly, identified marketing interactions help media planners devise integrated on and offline communication strategies. Finally, the estimated model allows simulation of the effects of all relevant drivers on consumer demand, generating accurate forecasts and playing an important part in the supply chain management process.