

Marketing Mix Modeling (MMM) - Concepts and Model Interpretation

Sandeep Pandey
Global Head of Analytics
Wavemaker (a WPP Co.)

Snigdha Gupta
Head of Analytics & Data Science
Wavemaker (a WPP Co.)

Shubham Chhajer
Lead Data Scientist
Wavemaker (a WPP Co.)

Abstract-- Marketing mix modeling has existed for decades now. Everyone has been using it, some tapped its potential with enormous success while others are yet to see its true potential. Rapidly changing marketing environment, consumer dynamics and multi touch points have made it even more complex to get it right for any industry and product. The biggest challenge in the process of any marketing mix optimization is measuring real-time cross effects and cross channel impact on business. The intent of this paper is to introduce marketing managers, consultants, analysts, strategists, and researchers to an analytics application to optimize the allocation of a firm's marketing budget in such a way that provides the maximum likelihood of generating higher ROI. MMM uses advanced econometrics and marketing science to objectively measure the relative efficacy and effectiveness of an entire set of marketing and advertising investments, competitive steal, or initiatives to produce sales and growth in both short and long term. In this paper we discuss the methodologies used to perform such analysis, how to overcome major challenges, and the benefits that can be derived from the analysis. We also discuss opportunities for improvement in media mix models that can produce agile granular measurement for better strategy.

Keywords: Marketing mix, MMM, econometrics, higher ROI, budget allocation, agile measurement, strategy, marketing returns, marketing response

I. INTRODUCTION TO MMM

CMO's today are under immense pressure to provide quantifiable evidence of how their marketing expenditure is helping the organization achieve its "Business Goal". To add to the complexity, they not only have to manage ever increasing non-traditional channels (Digital) but also smarter consumers who are exposed to multi touch points resulting in faster fatigue. Channels work in synergy and interplay is unique for each market, industry, and company. In the current scenario of ever-increasing advertising and marketing budgets, advertisers and businesses have a need to understand the effectiveness and ROI of their media spend in driving business KPIs and optimize budget allocation to higher returning channels.

Marketers around the world are questioning their marketing tactic's effectiveness every single day. Business problems they are seeking answers to: 1) Quantify the impact of a specific marketing plan or strategy on business sales. 2) How does their current marketing tactics impact future sales? If the marketing tactics simply aren't working, you need to review the wider strategy and re-invest the budget for higher returns.

For the longest time, marketing mix modeling (MMM) has been an important technique for advertisers wanting answers to these two questions. MMM helps to understand the impact

of marketing tactics to then optimize the strategy and ensure that a business isn't wasting marketing dollars.

The problem? Even tech, social media and Madtech (Martech+Adtech) giants like Facebook has admitted that the old process of basing decisions on many years of data is now outdated.^[1] Initially, the idea was to learn about marketing tactics using stable data. In today's environment, one of the biggest problems is that stable data doesn't exist. Today, we have a complex web of interconnected digital channels that are always evolving and changing.

The marketing mix refers to analysis of variables that a marketing manager can control to influence a brand's KPI like sales or market share.^[2] Traditionally, these variables are summarized as the 4Ps of marketing: product, price, promotion, and place (i.e., distribution). Product refers to aspects such as the firm's portfolio of products, the newness of these products, their differentiation from competition, or their superiority to rivals' products in terms of quality. Promotion refers to advertising, detailing, or informative sales promotions such as features and displays. Price refers to the product's list price or any incentive sales promotion such as quantity discounts, temporary price cuts, or deals. Place refers to delivery of the product measured by variables such as distribution, availability, and shelf space.^[3] The perpetual question that business teams and stakeholders face is, what level or cross-combination of these variables maximize business KPIs like sales, market share, or growth? The answer to this question, in turn, depends on the following question: How does the business KPI respond to past levels of or spending on these variables?

II. MODELING PRINCIPLE

Over five decades, researchers, economists, marketing scientist have focused intently on solving this quest of identifying and measure the efficacy, effectiveness and sensitivity of every dollar spent or the impact of operating levels on the business KPIs and metrics of interest. To do so, they have developed a variety of econometric and statistical models of market-variable response to the media marketing mix. Most of these models have focused on market response to advertising and pricing.^[4] The reason may be that expenses on these factors seem the most discretionary, so business teams and stakeholder are most concerned about how they manage these factors. It focuses on modeling response to these factors, though most of the principles apply as well to other variables ranging from distributions to even the smallest of the investments in the marketing mix.

The underlying response and sensitivity modeling approach is principally aligned to the hypothesis that past data on

consumer and market response to the marketing mix contain valuable information. This data also enables us to predict how consumers might respond in the future and therefore how best to plan marketing variables.^[5] Thus, one would want to capture as much information as one can from the past to make valid inferences and develop accurate forecasts and strategies for the future.

Assume that we fit a regression model in which the dependent variable is a brand's sales, and the independent variable is advertising or price.

$$Y_t = \alpha + \beta A_t + \varepsilon_t \quad (1)$$

Here, Y represents the dependent variable (e.g., sales), A represents advertising, α and β are coefficients or parameters that the researcher wants to estimate, and the subscript t represents various time periods. The ε_t are errors in the estimation of Y_t that we assume to independently and identically follow a normal distribution (IID normal).^[2] Equation (1) can be estimated by regression. Then the coefficient β of the model captures the effect of advertising on sales. In effect, this coefficient summarizes much that we can learn from the past. It provides a foundation to design strategies for the future. Clearly, the validity, relevance, and usefulness of the parameters depend on how well the models capture past reality. This paper explains how we can implement multiple modeling techniques in the context of the marketing mix.^[2]

This has been addressed in a sequential manner, the first step is to understand the variety of patterns by which markets respond to the change in marketing and advertising variables. These patterns of response are also called the effect of media and advertising. We then present the set of most important econometric modeling techniques and discuss how these classic models capture or fail to capture each of these effects.

III. PATTERNS OF ADVERTISING RESPONSE

The patterns of response to advertising and marketing can be clubbed into seven segments. These include current effect, carryover effect, shape effect, competitive effect, dynamic effect, content effect, and media effect. The first four of these effects are common across all marketing and media variables. The last three are specific to media and advertising

investments. In this paper we have tried to explicitly explain each of these effects in detail.

A. Current Effect

The current effect of advertising is the change in sales caused by an exposure (or pulse or burst) of advertising occurring at the same time period as the exposure. Consider Figure 1.A it plots time on the x-axis, sales on the y-axis, and the normal or baseline sales as the dashed line.^[2] The current effect of advertising is the spike in sales from the baseline given an exposure of advertising. Years of research indicate that this effect is small as compared to that of other marketing parameters and is quite fragile. For example, the current effect of price is 20 times larger than the effect of advertising.^[6,7] Also, the effect of advertising/ media burst is so small that it can easily be drowned out by the noise signals in the data. Thus, one of the most important tasks of the researcher or a marketing scientist is to specify the model very carefully to avoid exaggerating or failing to observe an effect that is known to be fragile.^[8]

B. Carryover Effect

The carryover effect of advertising is that portion of its effect that occurs in time periods following the pulse of advertising. Figure 1.B and 1.C shows long and short carryover effects respectively. The carryover effect may occur for several reasons, such as delayed exposure to the ad, delayed consumer response, delayed purchase due to consumers' backup inventory, delayed purchase due to shortage of retail inventory, and purchases from consumers who have heard from those who first saw the ad (word of mouth). The carryover effect may vary in intensity, from as large as or larger than the current effect itself.

Typically, the carryover effect is observed for short duration, as shown in 1.C, rather than of long duration, as shown in Figure 1.B. The long duration carryover effect, that researchers often find is due to the use of data with long intervals that are temporally aggregate.^[9] For this reason, researchers should use data that are as granular, stable and disaggregate over time as they can find.

The total effect from an exposure of advertising is the sum of the current effect and all the carryover effect due to it.

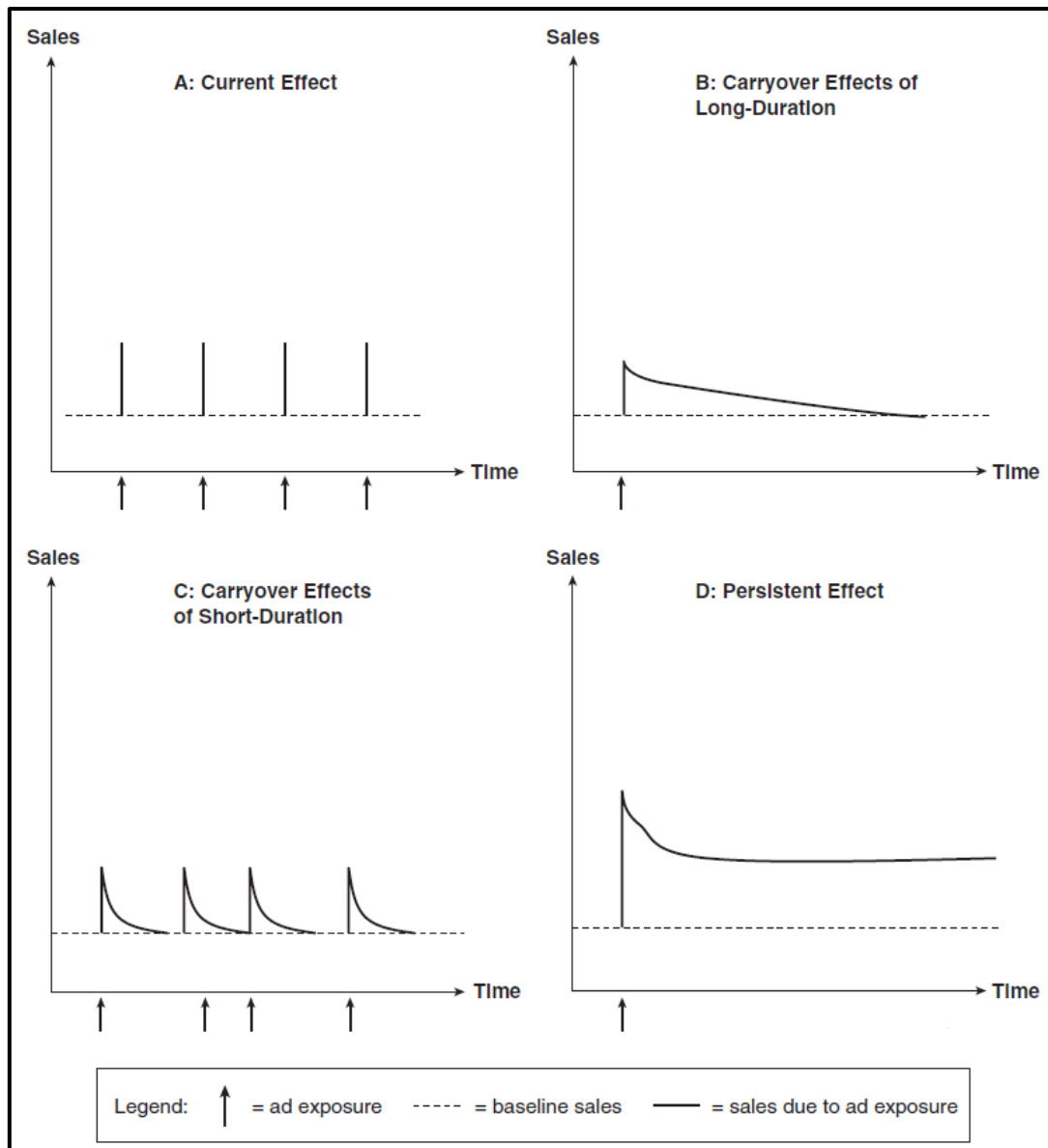


Figure 1: Temporal Effects of Media and Advertising (Illustrative)

C. Shape Effect

The shape of the effect refers to the change in sales in response to increasing advertising investments or intensity of advertising in the same period. The intensity of advertising could be in the form of exposures per unit time and is also called frequency or weight.^[2] Figure 2 describes varying shapes of advertising response. Note, that x-axis is the intensity or units of advertising exposures/ investments, while y-axis denotes the response/ change in sales. With reference to Figure 1, Figure 2 charts, the height of the bar in Figure 1.A may increase, as we increase the exposure of advertising.

Figure 2 shows three typical response curve shapes: linear, concave (increasing at a decreasing rate), and S-shape. Of these three shapes, the S-shape seems the most probable. The linear shape, logically can be highly impossible because it implies that sales will increase indefinitely up to infinity as advertising increases. The concave shape (diminishing

returns) in that area addresses the implausibility of the linear shape by getting saturated at very high operating levels. However, the S-shape seems the most probable because it suggests that a minimum threshold of advertising is required else it might not be effective at all because it gets phased out in the noise due to miniscule presence. While at some very high level, it might not increase sales because the market is saturated, or consumers suffer from fatigue due to over-exposure with repetitive advertising.

The elasticity or as we call responsiveness of sales to advertising is the rate of change in sales as we change media and advertising intensities. It is captured by the slope of the curve in Figure 2 or the coefficient (β in Equation (1)) in the model estimates of the curve and its respective equation. Just as we expect the advertising sales curve to follow a certain shape, we also expect this responsiveness of sales to advertising to show certain characteristics. First, the estimated response should be in the form of an elasticity. The elasticity of sales to advertising (also called advertising

elasticity) is the percentage change in sales for a 1 percentage point change in advertising. As defined, an elasticity is units-free and does not depend on the measures of advertising or of sales. It is a pure measure of advertising and media channels' responsiveness whose value can be compared across brands, products, companies, markets, and over time. Second, the elasticity should not plainly follow a linear shape i.e., neither always increase with the level and intensity of advertising nor be always constant but should show an inverted bell-shaped or inverted normal distribution pattern in the level of advertising.^[2] We would expect responsiveness to be low at low levels of advertising because it would be phased out by the noise in the market at such small operational levels, and low at very high levels of advertising because of diminishing returns and saturation. Thus, we can comprehensively expect the maximum responsiveness of sales at moderate operating levels of advertising. It turns out that when advertising has a sigmoid S-shaped response with sales, the advertising sensitivity/elasticity would have the inverted non-linear bell-shaped response to advertising levels. Therefore, the model that can capture this S-shaped relationship will also capture advertising elasticity in its theoretically most stable appealing form.

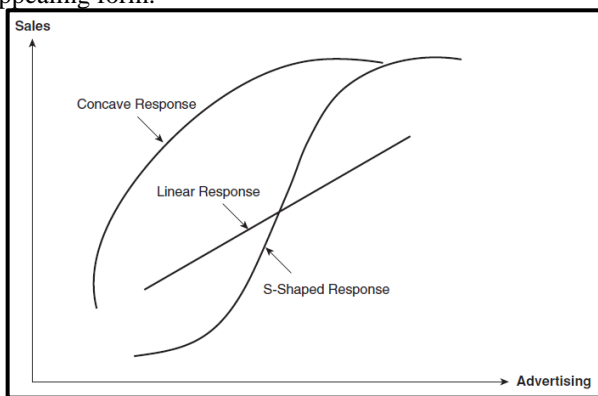


Figure 2: Linear and Non-linear Responses to Media and Advertising investments(illustrative)

D. Competitive Effects

It is assumed that Advertising normally takes place in free markets. The Game theory application in developing competitive strategy classically proves that in fair market conditions whenever one brand advertises a successful innovation or successfully uses a new advertising form/campaigns, other brands quickly imitate it. Competitive advertising tends to increase the threat and steal market share thus reduce the effectiveness of target brand's advertising. The competitive effect of a target brand's advertising is its responsiveness to that of the competitive brands in the category and market. As most advertising takes place in the presence of competition, trying to understand advertising of a target brand in isolation may be erroneous and leads to highly biased incorrect estimates of responsiveness.

In addition to just the steal effect of competitive advertising, a target brand's advertising might differ due to its position in the category and market or its familiarity with consumers. For example, established or larger brands may generally get more traction than new or smaller brands from the same

level of advertising because of the higher brand loyalty and awareness in the market. This effect is called differential advertising responsiveness due to brand position or brand familiarity.^[2]

A very uncommon phenomenon that is part of competitive effect is the category halo where the competition and category advertising increase the sales of the target brand. This can be observed in nascent categories and segments which are just entering the market and thus consumers are not much aware of the brand/ category itself, thereby all levels of advertising across the category by brands acts as a primer and catalyst in pushing the category collectively.

E. Dynamic Effects

Dynamic effects are those effects of advertising that varies with time. This includes carryover effects discussed earlier along with wearin, wearout, and hysteresis discussed here. To understand wearin and wearout, we need to return to Figure 2. As we have seen for the concave and the sigmoid S-shaped advertising response, sales increase until they reach some peak as advertising intensity increases. This advertising response can be captured in a static context—say, the first week or the average week of a campaign. However, in reality, this response pattern changes as the campaign progresses.

Wearin is the increase in the response of sales to advertising, from one week to the next of a campaign, even though advertising occurs at the same level each week. Figure 3 shows time on the x-axis and sales on the y-axis. It assumes an advertising campaign of n weeks, with one exposure per week at approximately the same time each week. Observe small spikes in sales with each exposure. However, the spikes keep increasing during the first 3 weeks of the campaign, even though the advertising level is the same. This is the phenomenon of wearin. Wearin effects are typically observed at the start of a campaign. Researchers may attribute this effect to repetitive campaign communication in subsequent periods enabling more people to see the ad, talk about it, think about it, and respond to it than would have done with only one high value burst of the campaign. Wearout is the decline in sales response to advertising from week to week of a campaign, even though advertising occurs at the same level each week. Wearout typically occurs at the end of a campaign because of consumer fatigue. Figure 3 shows wearout in the last 3 weeks of the campaign.

Hysteresis is the permanent effect of an advertising exposure that persists even after the pulse is withdrawn or the campaign is stopped (see Figure 1.D). Typically, this effect does not occur more than once. It occurs because an ad established a connect through a dramatic impact and through a previously unknown fact, linkage, or relationship. Hysteresis is an unusual effect of advertising that is quite rare.

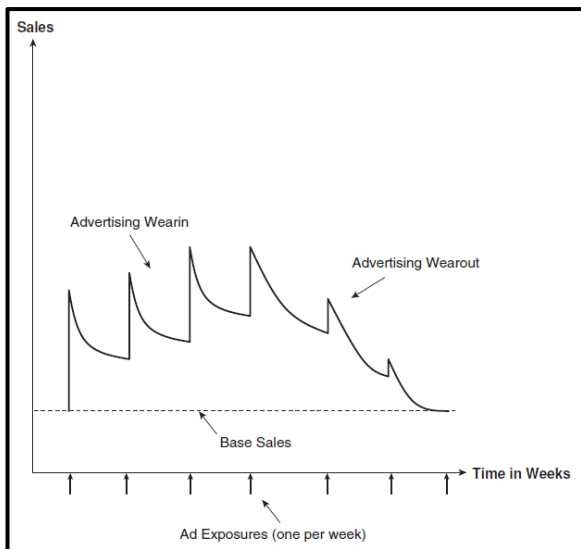


Figure 3: Wearin and Wearout Effects of Advertising (illustrative)

F. Content Effects

Content effects are the variation in response to advertising due to variation in the content or creative cues of the ad. This is the most important source of variation in advertising responsiveness and the focus of the creative talent in every agency. This topic is essentially studied in the field of consumer behavior using laboratory or theatrical experiments. However, experimental findings cannot be easily and immediately translated into management practice because they have not been replicated in the field or in real markets. Typically, modelers have captured the response of consumers or markets to advertising measured in the aggregate (in dollars, gross/total ratings points, or exposures) without regard to advertising content. So, the challenge for analysts is to include measures of the content of advertising when modeling advertising response in real market. Many researchers are now working in the domain of cognitive consumer neuroscience for assessing this impact qualitatively and quantitatively.

G. Media Effects

Promoting the sales of a product, via price discounts, advertisements and/or special displays, often has a huge impact on its sales. Successful execution of a sales promotion is possible if and only if the increase in sales volume is accounted for in all phases of the supply chain. With proper supply chain planning, driven by promotion forecasts, it is possible to satisfy the increased demand of the promoted product without introducing spoilage or surplus stock. However, the sales promotion of one product may in addition have significant secondary effects (halo or cannibalization) on the sales of other products not in promotion – a fact that is often forgotten or left with little attention.

Media effects are the difference in advertising response through various media channels (traditional and digital) of target brand as well as the other product lines much popular (Halo Effect/ Cannibalization), such as TV/ Newspaper/ Search/ Display etc. It also includes the efficacy by different

attributes, like channels/ genre/ schedule etc. for TV/ section/ story for display ads.

IV. MODELING ADVERTISING RESPONSE

In this section we discuss four different models of media mix/ advertising response, which address one or more of the above effects. The models below are in the order of increasing computational complexity. We also discuss the advantages and limitations of each model, which can help readers in comprehensively understanding the value and the progression from simple marketing mix models to more complex ones. Through ensemble models, a researcher or a marketing scientist may be able to develop a model that can capture many of the effects discussed above. However, that task is achieved at the cost of huge computational complexity. In an ideal theoretical scenario, the marketing mix model should be rich enough to capture all the seven effects discussed above. No one has proposed a model that has done so, though a few have come close by combining these models.

A. Basic Linear Model

Linear models describe a continuous response variable as a function of one or more predictor variables. The basic linear model captures only a few of the advertising effects, most common one is the current effect. The model takes the following form:

$$Y_t = \alpha + \beta_1 A_t + \beta_2 P_t + \beta_3 R_t + \beta_4 Q_t + \varepsilon_t \quad (2)$$

Here, Y is the dependent variable (e.g., sales), while the other capital letters represent variables of the marketing mix, such as advertising (A), price (P), sales promotion (R), or quality (Q). α and β_k are the coefficients that a researcher wants to estimate. α here represents some form of the base of the dependent variable. β_k represents the effect of the k^{th} independent variable on the dependent variable. The subscript t represents various time periods. Below we also discuss the problem of the appropriate time interval, but for now, we can think of time to be in weeks or days. The ε_t are errors in the estimation of Y_t that is assumed to be independently, identically normally distributed (IID normal).^[2] This assumption means that there is no pattern to the errors, hence, it constitutes only random noise (also called white noise). This simple model assumes we have multiple enough observations/ data (over time) for sales, advertising, and the other marketing variables and therefore can best be estimated by regression, a simple and powerful statistical tool.

B. Multiplicative Model

The multiplicative model takes its name from the fact that the independent variables of the MMM are multiplied together. It is a description of the effect of two or more predictor variables on an outcome variable that allows for interaction effects among the predictors. This contrasts with an additive model, which sums the individual effects of several predictors on an outcome. Thus,

$$Y_t = \text{Exp}(\alpha) \times A_t^{\beta_1} \times P_t^{\beta_2} \times R_t^{\beta_3} \times Q_t^{\beta_4} \times \varepsilon_t \quad (3)$$

While this model seems complex, a transformation can render it quite simple. In particular, the logarithmic transformation linearizes Equation (3) and renders it like Equation (2); thus,

$$\log(Y_t) = \alpha + \beta_1 \log(A_t) + \beta_2 \log(P_t) + \beta_3 \log(R_t) + \beta_4 \log(Q_t) + \varepsilon_t \quad (4)$$

The main difference between Equation (2) and Equation (4) is that the latter has all variables as the logarithmic transformation of their original state in the former. After this transformation, the error terms in Equation (4) are assumed to be IID normal.

The multiplicative model has many benefits:

- a) First, the multiplicative model implies that the business KPI/ dependent variable is affected by an interaction of the variables of the marketing mix. In other words, the independent variables have a synergistic effect on the dependent variable. In many advertising real life situations, the variables indeed interact to have such an impact. For example, higher advertising *combined* with a price drop may enhance sales more than the *sum* of higher advertising or the price drop occurring alone.
- b) Second, Equations (3) and (4) imply that response of sales to any of the independent variables can take on a variety of shapes as seen above depending on the value of the coefficient. In other words, the model is flexible enough that it can capture and simulate relationships that take a variety of shapes by estimating appropriate values of the response coefficient.
- c) Third, the coefficients not only estimate the effects of the independent variables on the dependent variable, but they are also elasticities.

The multiplicative model has two main limitations. First, it only estimates two, (the current and the shape effect) of the seven effects described above. Second, the multiplicative model is unable to capture a sigmoid S-shaped response of independent variables to sales.

C. Exponential Attraction and Multinomial Logit Model

Attraction models are based on the premise that market response is the result of the attractive power of a brand relative to that of other brands with which they compete. The attraction model implies that a brand's share of market sales is a function of its share of total marketing effort; thus,

$$M_i = S_i / \sum_j S_j = F_i / \sum_j F_j \quad (5)$$

Here, M_i is the market share of the i^{th} brand (measured from 0 to 1), S_i is the sales of brand i , \sum_j implies a summation of the values of the corresponding variable over all the j brands

in the market, and F_i is brand i 's marketing effort and is the effort expended on the marketing mix (advertising, price, promotion, quality, etc.). Equation (5) has been called Kotler's fundamental theorem of marketing. Also, the righthand-side term of Equation (5) has been called the attraction of brand i . Attraction models intrinsically capture the effects of competition.^[2]

A simple but inaccurate form of the attraction model is the use of the relative form of all variables in Equation (2). So for sales, the researcher would use market share. For advertising, he or she would use share of advertising expenditures or share of gross rating points (share of voice) and so on. While such a model would capture the effect of competition, it would suffer from other problems of the linear model, such as linearity in response. It is also incorrect because the RHS would be the sum of the individual shares of effort of each element and not exactly the share of marketing effort in total. A modified linear attraction model can be used to resolve the problem of linear response curves and the theoretical implausibility in stating the RHS of the model. The modification uses brand share of market with an exponential transformation in the marketing mix; thus,

$$M_i = \text{Exp}(V_i) / \sum_j \text{Exp} V_j \quad (6)$$

where M_i is the market share of the i^{th} brand (measured from 0 to 1), V_j is the marketing effort of the j^{th} brand in the market, \sum_j stands for summation over the j brands in the market, Exp stands for exponent, and V_i is the marketing effort of the i^{th} brand, expressed as the righthand side of Equation (2).^[2]

$$V_i = \alpha + \beta_1 A_i + \beta_2 P_i + \beta_3 R_i + \beta_4 Q_i + e_i \quad (7)$$

where e_i are error terms. By substituting the value of Equation (7) in Equation (6), we get

$$M_i = \text{Exp}(V_i) / \sum_j \text{Exp} V_j = \text{Exp}(\sum_k \beta_k X_{ik} + e_i) / \sum_j \text{Exp}(\sum_k \beta_k X_{jk} + e_j) \quad (8)$$

where X_k (0 to m) are the m independent variables or elements of the marketing mix, and $\alpha = \beta_0$ and $X_{i0} = 1$. The use of the ratio of exponents in Equations (6) and (8) ensures that market share is an S-shaped function of share of a brand's marketing effort.

However, Equation (8) also has two limitations. First, the models are very difficult to interpret because the RHS of Equation (8) is in exponential form. Second, the denominator of the RHS is a sum of the exponents of the marketing effort of each brand summed over each element of the marketing mix. Fortunately, both these problems can be solved by applying the log-centering transformation to Equation (8).^[10] After applying this transformation, Equation (8) reduces to

$$\text{Log}(M_i M^-) = \alpha^*_i + \sum_k \beta_k (X^*_{ik}) + e^*_i \quad (9)$$

where the terms with $*$ are the log-centered version of the

normal terms; thus, $\alpha^*_i = \alpha_i - \bar{\alpha}$, $X^*_{ik} = X_{ik} - \bar{X}_i$, $e^*_i = e_i - \bar{e}$, for $k = 1$ to m , and the terms with (\cdot) are the geometric means of the normal variables over the m brand in the market. The log-centric transformation of Equation (8) reduces it to a type of multinomial logit model in Equation (9). The advantage of this model is that it is relatively simpler, more easily interpreted, and more easily estimated than Equation (8). The right-hand side of Equation (9) is a linear sum of the transformed independent variables. The left-hand side of Equation (9) is a type of logistic transformation of market share and can be interpreted as the log odds of consumers preferring the target brand relative to the average brand in the market. The particular variant of the multinomial logit in Equation (9) is the aggregated form. That is, this form is estimated at the level of market data obtained in the form of market shares of the brand and its share of the marketing effort relative to the other brands in the market. An analogous form of the model can be estimated at the level of an individual consumer's choices.^[11] This other form of the model estimates how individual consumers choose among rival brands and is called the multinomial logit model of brand choice.^[12] The multinomial logit model (Equation (9)) has a number of attractive features that render it superior to any of the models discussed above.

- a) First, the model includes the competitive terms, so that prediction of the model is actually sum and range constrained, same as the original data. That is, the predictions of the market share of any brand range between 0 and 1, and the sum of the predictions of all the brands in the market equals 1.
- b) Second, and more important, the working form of Equation (6) suggests a characteristic sigmoid S-shaped response between market share and any of the advertising variables (Figure 2). In the case of advertising, for example, this shape implies that sales response to advertising is low at levels of very low or very high advertising. This characteristic is particularly appealing based on advertising theory. The main reason is that miniscule levels of advertising may not be effective and have any impact because they get phased out in the noise. As discussed in shape effect very high levels of advertising also may not have any impact because of market/ channel saturation and diminishing returns. If the estimated lower or minimum threshold of the sigmoid S-shaped relationship does not occur at 0, this indicates that market share or sales maintains some minimum levels even when marketing effort is reduced to a zero operational point. We can interpret this minimal floor to be the base loyalty of the brand. Alternatively, we can interpret the level of marketing effort that coincides with the threshold (or first turning point) of the sigmoid S-shaped curve as the minimum threshold point necessary for consumers or the market/ category to even

notice a significant change in marketing-advertising effort.

- c) Third, because of the S-shaped curve of the multinomial logit model, the elasticity of market share to any of the independent variables shows a characteristic bell-shaped relationship with respect to marketing effort. This relationship implies that at very high levels of marketing effort, a 1% increase in marketing effort translates into ever smaller percentage increases in market share. Conversely, at very low levels of marketing effort, a 1% decrease in marketing effort translates into ever smaller percentage decreases in market share. Thus, market share is most responsive to marketing effort at some intermediate level of market share. This pattern is what we would expect intuitively of the relationships between market share and marketing effort. Despite its many benefits, the exponential attraction or multinomial model as defined above does not capture the latter four of the seven effects identified above.

D. Hierarchical Models

The remaining effects of advertising that we need to capture (content, media, dynamic effect- wearin/ wearout) involve changes in the responsiveness itself of advertising (i.e., the β coefficient) due to advertising content, media used, or time of a campaign. These effects can be captured in one of the two ways: dummy variable regression or a hierarchical model. Dummy variable regression is the use of various interaction terms to capture how advertising responsiveness varies by content, media, wearin, or wearout. We illustrate it in the context of a campaign with a few ads. Suppose the advertising campaign uses only a few, say two different types of ads. Assuming, we start with the basic regression model of Equation (3). Then we can capture the effects of these different ads by including suitable variables. One simple formulation is to include a dummy variable or flag for the second ad, plus an interaction effect of advertising times this dummy variable.^[2] Thus,

$$Y_t = \alpha + \beta_1 A_t + \delta A_t A_{2t} + \beta_2 P_t + \beta_3 R_t + \beta_4 Q_t + \epsilon_t \quad (10)$$

where A_{2t} is a dummy variable that takes on the value of 0 if the first ad is used at time t and the value of 1 if the second ad is used at time t . δ represents the effect of the interaction term $A_t A_{2t}$. In this case, the main coefficient of advertising, β_1 , captures the effect of the first ad, while the coefficients of β_1 plus that of the interaction term (δ) capture the response estimate of the second ad. While simple, these models quickly become quite computationally complex when we have simultaneous occurrence of multiple advertisements, channels, media, and time periods. This is the situation in real world. The problem can be solved using the hierarchical models. Hierarchical models are multistage models in which coefficients (of advertising) estimated in one stage become the dependent variable in the other stage. The

V. FUTURE OF MMM

second stage contains all the characteristics by which advertising is likely to vary in the first stage, such as ad content, medium, or campaign duration.^[2]

Two features are essential for hierarchical models:

- a) We should be able to simulate and obtain multiple estimates of the effects (or coefficient values) of advertising on some dependent variable such as sales or market share for the same brand across different contexts. It may include at least one of the following: the ad campaign, week of the campaign, market, or medium. Then we can use the estimates of the effects of advertising from the first stage as dependent variables in the second stage.
- b) As far as possible, we also need to minimize the excessive covariance among factors. Co-occurrence of similar time series data leads to the problem of multicollinearity among the created variables in the second-stage model. If these factors have sufficient cross-variance, coefficient estimates of the second-stage model should be reliable and usable with minimum error propagation. Depending on the sample size and richness of the data, hierarchical models can estimate the above discussed effects of advertising. That is, with such models and given suitable reliable data, the researcher can measure the most effective ad content, its duration, current returns, and relevant advertising mix for higher returns. The duration of the campaign could be estimated in terms of months or weeks or days. For example, if the effectiveness and impact of the ad first increases slowly in a gradual manner and then decreases suddenly, one could conclude that wearin is slow but wearout is rapid. On the other hand, if the impact of the ad sharply declines over time, then there is no wearin, and wearout sets in from the very start. Furthermore, if the data is sufficiently rich and detailed, the marketing scientists can also obtain synergistic interaction effects and much more granular information such as which channels are most suitable for particular ads or which content needs to be run over campaigns of long versus short duration.

Note that to address all the seven effects of advertising identified above, the researcher would have to use a combination of multiple constructs of hierarchical model with a special check on error propagation across hierarchies, which will itself ensemble an exponential attraction or multinomial logit model along with a Koyck-type distributed lag enhancement. In other words, ensemble models described above would enable a researcher to address the most important phenomena associated with advertising. In reality, such fully integrated models that can capture all the effects of advertising are very complex and require substantial data. If researchers want to focus on only a few effects or their data is not rich enough, they might want to simplify the model they use to focus on only the most important effects.

With the growing investments in digital channels, ever growing consumer base of ecommerce and social platforms, disappearance of cookies and the emergence of tech and AI has disrupted the dynamics of marketing. The pace of change in today's transformed marketing world has increased significantly and requires advertisers and agency to be all prepared for this. Now is the time to take a hard look at the current measurement framework and investments to understand which areas will be affected most with the change in events in this disconnected ecosystem. The real-life marketing world comes with its own increasingly complex set of challenges, including a limited, siloed view of behaviors and limited lookback stable windows.

Due to this complex ecosystem marketers and advertisers face many difficulties in measuring effectiveness of the investments they make. Immense amount of research and innovations are being worked out in the areas related to effectiveness measurement and optimization of these investments. We have identified three major areas of working which is very important in the current scenario and to transform the entire measurement system. As mentioned, they are not the only challenges that exist but solving them would make a significant impact. These focus areas and innovations will elevate the techniques of measurement from an aggregated (less granular) to a disaggregated attribute-based (more granular) ecosystem with faster turnarounds.

A. Granular attribute-based deep dive: Causality, effectiveness, and efficacy measurement

In medicine, proving cause and effect is literally a question of life and death. Survival rates improve with the introduction of a new drug. But is this causation or merely correlation? In the world of marketing the stakes may not be as high, but the problem remains the same.^[13] How do we attribute Sales increase after an advertising campaign, to the specific marketing effort only, is there any other channel that enhanced the effect of one channel—or this is due to some other impact altogether? The process of estimating the true effect of an effort on a business KPI known as '*causal inference*' is the essence of measurement. And common effectiveness and efficacy measurement methods don't always get it right. Rather than abandoning these trusted methods altogether, marketers can adopt medicine's approach: a 'hierarchy of evidence' that favors methods higher up the hierarchy. Where it's not possible to employ the correct sane method of evaluation, marketers and researchers should be aware of the limitations of their research and the bias, ambiguity and uncertainty attached to the inferences and outcomes. Experts should communicate this uncertainty and propose valid contexts and assumptions, so that it doesn't hamper decision-making. Below are the three major focus areas of granular correct measurement:

- a) Assessing causality from observational data: When we can't run experiments, one must rely on methods that use observational data such as marketing mix modelling or digital attribution.

These measurement techniques are not always good at estimating causal effects therefore valid hypothesis and testing frameworks must be put in place to verify the underlying assumptions of operation. Algorithms now exist to determine when a particular observational method is a good estimate of causal effects and when it is not. Such algorithms can analyze causal cross variable diagrams which encode our understanding of how marketing efforts causes an outcome alongside other factors. These have seldom been applied to marketing effectiveness. What if we could develop tools allowing marketers to build causal diagrams, to analyse them automatically and to recommend optimal methods for estimating causal effects?

- b) Design of experiments: Randomized controlled experiments are by far the most accurate method of measuring causal effects available. But they can typically only test one or two things at a time, they can be difficult to administer, and require much more granular user attribute data. Continuous evaluation systems must be put in place that enable marketers to take on the studies and measure the effectiveness of different marketing plans and strategies to quantify for future deployments in enhanced MMM outputs. The feeds from this user level ecosystem to that of MMM measurements can validate and improve the efficacy and effectiveness of the MMM ecosystem.
- c) Communicating uncertainty: Regardless of the measurement, we need to acknowledge and understand the error margins present in estimates of marketing effectiveness. It is very important to put in place all the underlying assumptions and hypothesis (formulations and testing frameworks) to accurately integrate the outputs of measurements in day-to-day planning and operations.

B. Measure the long term, today

CXOs (CEOs, CMOs, etc.) must continually balance their business strategies and decisions that drive long-term growth against the short-term returns as demanded by investors or shareholders. Marketers walk the same tightrope: investing in a brand is essential in growing a sustainable business but may not drive quarterly sales at a good return on investment. According to some sources, marketers have become too focused on the short-term returns, and this damages effectiveness. A successful long-term strategy requires better measurement of long-term effects, without having to wait years for the results.

- a) Integrating long-term and short-term MMM results: Advanced MMM approaches can now estimate the long-term sales delivered by marketing and create a 'multiplier' comparing this to short-term sales. But this may require many years of data and is more complex than regular MMM, meaning the approach isn't often used and is multi-layered. Marketers may also employ multipliers out of context or without understanding

their error margins which can introduce huge bias and error in the outputs.

- b) Integrating customer lifetime value (LTV) and first-party data into MMM and its short-term outcomes: With good customer data, marketers can model the expected LTV of potential customers, so they can be targeted with customized messages or increased spend for higher engagement. But only a few advertisers have quality data and required systems to modify assumptions of these underlying models based on real consumer behavior which can then be connected with automated marketing systems.
- c) Using online behaviors as a leading indicator of long-term outcomes: With growing internet penetration and evolution of social media and e-commerce- online behaviors (such as search, social queries etc.) can provide frequent, granular data with very large sample sizes. Research shows that this data is related to brand health and may be a leading indicator of long-term outcomes. The effectiveness experts must innovate to take these input feeds into MMM to better assess the responses well in advance. If one could quantify an estimated financial value to the uplifts and gains in online behavior caused by marketing will change the course of marketing measurement.

C. Unified methodology and more granularity

For decades, scientists have sought a way to marry the theory of the very small (quantum mechanics) with the theory of the very large (classical physics), to develop a so-called 'theory of everything'. A similar challenge is emerging in effectiveness measurement. Consumer-level models like digital attribution measure at the level of the very small. Aggregate-level models like marketing mix modelling measure the very large. And they can lead to very different results. The logical next step, therefore, is to find a way to bring together measurement methods to get a wholistic view of their effectiveness: a 'theory of everything' for marketing.

1. Blending one method with another: While some advertisers and measurement providers claim they are blending MMM with digital attribution, MMM with experiments, or digital attribution with experiments, they form a minority, and best practice remains unclear.
2. Blending multiple methods at once: Beyond traditional methods, effectiveness experts seek to blend data and insight from different sources. This can involve collating and presenting it all in a consumable way or blending it on an analytical level and presenting a unified result. Some promising approaches exist, but there are no industry standards.

What if effectiveness experts could agree on the ideal process for bringing together and presenting multiple sources of information? What if researchers could build transparent models to blend data of different granularities (user, cohort, geo,

aggregated, etc.) to get consistent and holistic measures of effectiveness?

3. AI amalgamation for scalable production grade MMM advancements and innovations: The primary objective of AI investments is to digitally and technologically transform the way measurements are done, from traditional manual single model-based analysis which is time consuming to generating millions of models providing agility, scalability and broad range of selection criteria to choose from in an automated timely manner. With the help of AI, technology, and cloud infrastructure available we can utilize highly advanced and computationally complex simulation based multi-layered econometric causal structural model approach (n-layered X m-factors) which is an advanced measurement approach for generating scenarios and calculating efficacy and effectiveness of investments in a complex connected ecosystem.^[14]

VI. CONCLUSION

For marketers, the advertising opportunities are continually changing as the consumer behavior is continuously evolving and digitally transforming. Back in the days, assessing the impact of advertising on TV and print media was simple because it was a constant environment with very few changes. However, measuring the true influence of marketing tactics in today's ever-changing environment is like a cat chasing a laser light around the room.

Does this mean an end for marketing mix modeling? No, this isn't the case at all. We're sure that you've seen marketers complaining about the end of MMM, but this is probably because they tried the original method and gave up when it wasn't effective. Instead, what you need to do is completely forget about your tactics from the past. MMM isn't dead; it has gone through a period of evolution, and you need to update your techniques regarding how you operate with MMM.

Planning the marketing mix is a central task in marketing management. Prudent planning requires that marketing managers consider how markets have responded to the marketing mix in the past. The underlying assumption is that the past predicts the future with absolute certainty but that it contains valuable lessons and patterns that if extracted correctly might enlighten the future and better shape and aid our strategic decision-making.

The econometrics of response modeling describes how a researcher should model response to the marketing mix to capture the most important effects validly. This paper provides an overview of the essential issues and principles in this area. It first describes the important advertising effects that occur today. It then discusses in detail the computational complexities, strengths and limitations of various models that capture those effects.

REFERENCES

- [1] Trapica Content Team. Facebook Modern Marketing Mix Models. <https://medium.com/trapica/facebook-modern-marketing-mix-models-2777ff99a6d0>
- [2] Tellis, G. (2006). Modeling marketing mix. In R. Grover, & M. Vriens *The handbook of marketing research: Uses, misuses, and future advances* (pp. 506-522). SAGE Publications, Inc., <https://www.doi.org/10.4135/9781412973380.n24>
- [3] McCarthy, J. (1996). *Basic marketing: A managerial approach* (12th ed.). Homewood, IL: Irwin.
- [4] Sethuraman, R., & Tellis, G. J. (1991). An analysis of the tradeoff between advertising and pricing. *Journal of Marketing Research*, 31, 160–174.
- [5] Tellis, G. J., & Zufryden, F. (1995). Cracking the retailer's decision problem: Which brand to discount, how much, when and why? *Marketing Science*, 14(3), 271–299.
- [6] Sethuraman, R., & Tellis, G. J. (1991). An analysis of the tradeoff between advertising and pricing. *Journal of Marketing Research*, 31, 160–174.
- [7] Tellis, G. J. (1986). Beyond the many faces of price: An integration of pricing strategies. *Journal of Marketing*, 50, 146–160.
- [8] Tellis, G. J., & Weiss, D. (1995). Does TV advertising really affect sales? *Journal of Advertising*, 24(3), 1–12.
- [9] Clarke, D. G. (1976). Econometric measurement of the duration of advertising effect on sales. *Journal of Marketing Research*, 13, 345–357.
- [10] Cooper, L. G., & Nakanishi, M. (1988). *Market share analysis*. Norwell, MA: Kluwer.
- [11] Tellis, G. J. (1988a). Advertising exposure, loyalty and brand purchase: A two-stage model of choice. *Journal of Marketing Research*, 15, 134–144.
- [12] Guadagni, P., & Little, J. D. C. (1983). A logit model of brand choice calibrated on scanner data. *Marketing Science*, 2, 203–238.
- [13] Think it with Google. Measuring effectiveness Three Grand Challenges: The state of the art and opportunities for innovation. <https://www.thinkwithgoogle.com/>
- [14] Pandey, Sandeep and Gupta, Snigdha and Chhajed, Shubham, ROI of AI: Effectiveness and Measurement (June 2, 2021). INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 10, Issue 05 (May 2021), Available at SSRN: <https://ssrn.com/abstract=3858398> or <http://dx.doi.org/10.2139/ssrn.3858398>