Paper 10721-2016

Can you STOP the guesswork in your marketing budget Allocation??? Marketing Mixed Modeling using SAS® can help!!!

Delali Agbenyegah, Alliance Data Systems, Columbus, Ohio

ABSTRACT

Even though marketing is inevitable in every business, each and every year the marketing budget is limited and prudent fund allocations are required to optimize marketing investment. In many businesses, the marketing fund is allocated based on the marketing manager's experience, departmental budget allocation rules and sometimes 'gut feelings' of business leaders. Those traditional ways of budget allocation yield sub optimal results and in many cases lead to money wasting on certain irrelevant marketing efforts. Market Mixed models can be used to understand the effects of marketing activities and identify the key marketing efforts that drive the most sales among a group of competing marketing activities. The results can be used in marketing budget allocation and take out the guess work that typically goes into the budget allocation. In this paper, we illustrate how to develop and implement Market Mixed Modeling using SAS® procedures from a practical perspective. Real life challenges of market mixed model development and execution are discussed and several recommendations are provided to overcome some of those challenges whenever possible.

INTRODUCTION

Each and every year, firms spend billions of dollars in marketing. According to Chief Marketing Officer Council World Wide, global marketing spend amount to $540 billion dollars in 2014, with US marketing spend totaling $179.8 billion. Within the US, 28.2% of the marketing spend was on investments in digital marketing and 10.6% on mobile. Digital marketing spend increased about 17.7% in 2014 and is expected to increase by another 15.5% in 2015. With the increase in the popularity of digital engagement and social media, marketing budget allocation will only become more complicated. The questions here are simple. How do firms allocate their budget today? How do they evaluate the effectiveness of those investments and how do they plan to invest their limited marketing budget in the future to ensure the highest return on those investments whiles building brand image and reputation? Though these questions sound simple, the answer may not be that easy to come up with.

Many marketing managers use their experience, learnings from other businesses as well as their 'gut feelings' to allocate marketing budget. These approaches do not guarantee an optimal budget allocation and in many cases lead to money wasting in some media channels. However with the increasing successes achieved in using predictive analytics in advising businesses, there is some opportunity for Data Scientists to once again establish reputation by diligently applying predictive analytics to guide their firm's marketing budget allocation by using Marketing Mix Modeling (MMM).MMM is the use of statistical and analytical methods to quantify the effectiveness of various marketing activities (marketing mix) and systematically optimizing marketing budgets and allocating resources into profitable marketing efforts. Whiles this is not an easy task due to the numerous challenges that will be outlined later in this paper, the author believes that taking 'baby-steps' in applying MMM to budget allocation will not only outperform the current traditional approach which is mainly an educated guess work in many cases, but will also establish value for business leaders to make moves that will solve many of the challenges in marketing effectiveness measurements. In the next sessions, we will share our five step approach to a successful MMM implementation and provide illustrations with a fictitious client data where appropriate. We will then discuss some of the practical challenges that always come along the way, with the most challenging problem being data availability and marketing effect measurement due to confounding and simultaneous effects of marketing activities.

A FIVE-STEP GUIDE TO MARKETING MIX MODELING

1. BUSINESS APPLICATION AND DATA UNDERSTANDING

Understanding a firm's business environment is the first step of every model building process. It is necessary to understand the marketing strategies the corporation has been employing over the years, what guides their marketing budget allocation and how the effectiveness of their marketing activities is measured. Under this first step, the Analyst has to ask the right questions to understand the overall business plan as far as marketing and building brand image is concerned. It is also important to ask about how marketing activities are recorded and tracked. Are there campaign history databases, budget allocation records, specific product launch activity records, discount promotion records, in-store events records as well as any marketing related activity records. In many corporations, the Analysts may face challenges with regards to how marketing activity data is captured and stored.

2. DATA PREPERATION AND STRUCTURING

If there is any data that is scattered around between Marketing, Finance and Sales department, then it is marketing activity data. In many situations based on the author's experience, marketing activity spend and performance is not well recorded and structured and hence it is crucial for the Statistician to understand and come up with reasonable ways to collect and prepare data in a form that can be modelled. In many cases, some marketing data such as spend on TV ad, radio ad, Internet ad and spend on billboards may be recorded in different periodicity such as weekly, monthly or even quarterly. Other related measures such as Television and Radio Gross Rating Points (GRP) may also come in regional levels. This present some challenges to the Analyst and calls for a careful structuring of all available marketing investment and performance data to a form that can be easily modelled. I would add that the success of the model depends a lot on how accurate all the important data elements are organized and structured. Below is an example of some data elements that are useful for MMM.

Table 1: Example Data Elements for Marketing Mix Modeling

Example Data Elements for Marketing Mix Modeling	
Variable	**Source**
Sales	Client Data base
TV Ad spend	Finance
Promo redemption transactions	Client Data base
TV GRP	Media Agency
DM Spend	Marketing team
Email Spend	Marketing team
DM circulation	Marketing team
Email Circulation	Marketing team
DM promo window(no of days)	Marketing team
Email Promo window(no of days)	Marketing team
In store promo **spend**	Marketing team

Number of in store promo days	Marketing team
Internet Spend	Marketing team
Consumer price index	https://research.stlouisfed.org/fred2
Unemployment Rate	https://research.stlouisfed.org/fred2
Gas Prices	https://research.stlouisfed.org/fred2

3. EXPLORATORY DATA ANALYSIS

Once data has been collected and structured, the next step is to run basic exploratory analysis on the data to check trends and observe relationships. For marketing and sales performance data, it is always useful to explore bivariate relations and correlations between sales and the different marketing activities such as TV ad spend, TV GRP, Direct mail spend, in-store promotions and any other variables available. Our friendly SAS procedures such as UNIVARIATE, CORR, MEANS, GPLOT, FREQ and the like would be useful in exploring MMM dataset. Since data is mostly in time series, such us monthly sales, monthly spend on advertisement, monthly spend on promotions, etc., it is also crucial to check stationarity of the data using various unit root tests. This will advise the choice of modeling methodology discussed in our next step. SAS ETS procedures such as AUTOREG and ARIMA have options that easily run unit root tests on time series data.

4. MODEL FITTING USING SAS

4.1 ADSTOCK ESTIMATION

Fitting the model is the heart of the MMM process. Before jumping to the model estimation, let me introduce an important concept called 'Adstock estimation'. Exposure to TV ads builds awareness in consumer markets, resulting in sales. Every new exposure to advertisement builds awareness whiles past exposures may have decay effect when new exposures are not added. This effect is termed Adstock and it is crucial to estimate this and use it as an explanatory variable in your MMM if data permits. The most basic and common Adstock estimation is done by using the following relationship.

$$
\begin{aligned}
&ADStock_t(\alpha) = GRP_t + \alpha \times ADStock_{t-1} \\
&0 < \alpha < 1
\end{aligned}
$$

where α is estimated using past measured effects of advertisement or half-life time of advertisement. GRP can be obtained from the media agency showing the advert on behalf of the corporation. Since this paper is focused mainly on the MMM fitting using SAS and the practical challenges it that comes with it, refer to (Joy, 2006) for further understanding of AdStock transformation. Once the AdStock is estimated, it can now be used as an explanatory variable in the MMM.

4.2 MARKETING MIX MODEL FITTING USING REGRESSION

The hope of most marketing activities is to affect sales, though some may aim at building brand awareness. The focus in this paper is to illustrate using model to explain the relationship between different marketing activities and sales over time.

Generally, the model is of the form:

$$Sales_t = f(x1, x2, x3, ...xn) + \varepsilon_t$$

where x_i refers to different marketing spend and effect measurements as well as their transformations, laggings and interactions at time t.

Depending on the level of data available and the results of the data exploration, you may fit a Linear Regression Model, Autoregressive Model or Linear Mixed Model using SAS. We illustrate using a fictitious specialty retailer's data below.

The data set (*MMM_SET*) has weekly sales as our response and weekly marketing activities and some external macroeconomic variables in Table 1 as our independent variables. The model fitting is more of an art work and requires the creation of many transformed variables as well as interaction variables.

SAMPLE CODE 1: Assuming data is stationary and there is no autocorrelation, we fit a linear regression model using our final selected variables.

```
PROC REG DATA= MMM_SET;
MODEL Weekly_sales = TV_ad_2  Gas_prices promo_spend_days  DM_EM
Internet D_SAS D_bf D_holiday D_val /VIF TOL COLLIN  SELECTION=stepwise;
ODS  OUTPUT CollinDiag=colldiag ParameterEstimates=ParaEst;
RUN;
```

SAMPLE CODE 2: If autocorrelation is present, we correct it by fitting an autoregressive model

```
PROC AUTOREG DATA=MMM_SET;
MODEL Weekly_sales = TV_ad_2  Gas_prices promo_spend_days  DM_EM
Internet D_SAS D_bf D_holiday D_val /NLAG= 8 PARTIAL DW=1;
OUTPUT OUT=autoreg P= autoregpred UCL=ucl LCL=lcl;
QUIT;
```

SAMPLE CODE 3: Suppose we have regional data of marketing activity spend as well as their corresponding sales at the regional level, we fit a linear mixed model with a random class structure as the region. The RANDOM option specifies the variables whose effect on sales is assumed to be random among the different regions.

```
PROC MIXED DATA=MMM_SET METHOD=ml;
  CLASS  REGION;
  MODEL Weekly_sales = TV_ad_2  Gas_prices promo_spend_days  DM_EM
  Internet D_SAS D_bf D_holiday D_val / S;
  RANDOM INTERCEPT   Gas_prices promo_spend_days  DM_EM
  D_bf D_holiday D_val/SUB= REGION;
RUN;
```

The resulting output from the above SAS sample codes are similar to all the regression procedure outputs and hence are not included here.

5. MODEL IMPLEMENTATION AND INTERPRETATION

Unlike many other predictive models that utilize scoring and ranks to select subjects or classify individuals, marketing mix models are more like econometric models that need to be interpreted properly especially in terms of sensitivity and contribution of different marketing activities to overall sales.

After the model fitting, diagnostics and validation, the final step of our five step approach is to interpret the model to key stakeholders and explain the contribution of relevant marketing activities to the corporation's incremental sales. Other relevant business calculations such as the ROI or flow-thru of each marketing activity are also computed using the model results.

Finally, simulations can be run with different levels of marketing investments to compare the effectiveness of those investments and recommend to most effective marketing mix base on overall business goals and marketing strategy.

Completing these five steps discussed above come with a lot of challenges at each step and contributes to the unpopularity of MMM and the continuous use of 'guesswork' and 'gut-feelings' in marketing budget allocation. In the next session, we discuss some of those challenges and recommendations to deal with the challenges are provided wherever possible.

6. CHALLENGES OF BUILDING AND IMPLEMENTING MARKETING MIX MODELS

6.1 DATA AVAILABILITY

Model can only be built on data. An excellent modeler can only do a god job with accurate data to discover relationships between variables. Quantifying and forecasting marketing effectiveness using statistical methods is a relatively recent practice and most corporations just do not have data about marketing activities that have affected their sales performance in the past. Even in cases where a company has some data on TV ads and Radio ads, they may not have information on newspapers and magazines as well as bill boards that are carrying the same marketing message. Without the right information about all the marketing activities that affect sales such as TV ads, newspapers, magazines, billboards, direct mails, emails, in-store activities, catalogues and any other marketing events, it is almost impossible to build any meaningful model. In the author's experience, information on some marketing activities has to be retrieved from marketing managers past emails. In circumstances where data is even retrieved, they are in different periodicities and some are recorded at levels that need a lot of re-structuring to be made model ready. Though the best way to tackle this challenge is to encourage corporations to start building good marketing databases, data scientists and analysts can help solve the problem by showcasing the value of MMM and the need to collect and keep data in a model friendly format. Even in many cases where MMM was applied in the midst of bad data, results showed improvements over gut-feeling budget allocation and this will only improve with good data availability.

6.2 DIFFICULTY MEASURING ADVERTISEMENT CONTENT AND EFFECT

Let's face it. Advertisement content is hard to measure. The emotional appeal and connectivity created by a marketing ad is hard to quantify. The real effect of marketing activity on sales is not easy to measure. Many practitioners use gross rating points (GRP), impressions and expenditure on ads as the measurement tool. However, TV ads with similar GRP may have different effects on sales based on the content and the creativity employed in the ad. Likewise, a high expenditure on an ad does not guarantee better performance of that ad than a less expensive ad.

6.3 SIMULTANEOUS IMPLEMENTATION OF MARKETING ACTIVITIES

Many different marketing activities actually occur at the same time. Whiles TV and radio ads are going on, other direct mail coupons and emails might be going on. This makes it difficult to really measure the effect of all the different marketing efforts individually. It also causes multi collinearity in many of the data available for modeling. This problem may be compounded in many other cases by seasonal peaks in sales performance, especially in retail.

Some of these problems may be resolved by combining correlated variables to obtain interaction variables, creating seasonal dummy variables or using VARCLUS or PRINCOMP procedures to construct composite indexes.

6.4 DYNAMIC AND LAG EFFECTS OF MARKETING

Imagine you watch a TV ad tonight and you really like the product, you will not immediately purchase that product. Probably you are now watching that for the first time and this might be the seventh day of that advert. There is a carryover effect of marketing activities as well as decay effect and this is challenge that many practitioners face. One way to resolve this is through ad stock estimation which was discussed in step four of our five step process. Another way to handle this is to include lagged variables as well as their interactions into the model though this also poses another problem of collinearity. Determining an accurate number of lags to use is also not an easy task.

6.5 NON LINEARITY OF MARKETING RESPONSES

We all know that many of the things we approximate with linear models are not really linear and marketing effectiveness measurement is even more nonlinear. Thinking about various marketing activities and the role they play in bring more sales, it is hard to believe that the effect is really linear. Some practitioners use complicated nonlinear functions which at times make the model interpretation very difficult. Using an S-shaped curve to estimate media response is also popular and it’s based on the idea that marketing activities become more effective after reaching some threshold and then becomes ineffective after exceeding some other threshold (Hanssens et al 2001).Though this is true and can solve some of the nonlinearity problems if applied correctly, the existence of different media channels, their lags and interactions makes it very difficult for linear functions to survive.

6.6 INSTABILITY OF PARAMETER ESTIMATES

Due to the presence of multicolinearity, most of the parameter estimates of MMM are not stable. The data scientist has to be very cautious in building and validating the models as well as conducting multiple stability checks in order to be more confident in the model. Omitted variables and many measurement errors in marketing effectiveness even makes this problem worse. Structural changes in corporation’s activities such as creative design change, value proposition change, etc. also affects the parameter estimates. The analyst has to keep an eye on all these changes and make adjustments where possible to ensure a meaningful MMM.

6.7 UNOBSERVABLE EFFECTS

Marketing is affected a lot by many psychological factors which are probably hard to quantify and let alone including it into the model. Brand awareness and brand image play a big role interactively with marketing activities to affect sales. One way to resolve this problem may be to use survey data designed to measure these behavioral and psychological effects and then use them as latent variables in your model. In that case, the analyst may use structural equation models by invoking the CALIS procedure in SAS to obtain a meaningful model.

7. CONCLUSION

Marketing mix modeling involves the use of different statistical tools to measure the effectiveness of different marketing activities and using the insight from the statistical methods to recommend marketing mix that will result in a higher marketing effectiveness. Though a good MMM can significantly increase a corporations sales whiles even saving costs, it has a lot of challenges. Will data scientists and analysts dwell on the challenges and not harness the analytical power available to them? The author believes that the analytical community can establish more reputation in their corporations by combining the art and science of predictive modeling and data analysis in the midst of the challenges to influence their corporations budgeting decisions.

Also, establishing the framework and show casing the potential value of MMM with limited available data will create the awareness and the need to capture marketing data in a more effective way.

REFERENCES

- Bhattacharya, Patralekha (2012). *Marketing Mix Modeling: Techniques and Challenges*,Thinkalytics, SESUG Paper ST-152
- Cain,P.M(2010*),Marketing Mix Modeling and Return on Investment, Integrated Brand Marketing and Measuring Returns*, Palgrave, Macmillan
- Joy, Joseph (2006 *) Understanding Advertising AdStock Transformations*, MPRA Paper No. 7683
- Hanssens, Dominique M., Leonard J. Parsons, Randal L. Schultz. (2001) *Market Response Models*: Econometric and Time Series Analysis, 2nd ed. Boston, MA, Kluwer Academic Press

ACKNOWLEDGMENTS

The author will like to thank Yin Chen, Dong Yan and the entire Predictive Analytics team at Alliance Data Card Services for their support.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Delali Agbenyegah
Alliance Data Systems
3100 Easton Square Place
Columbus, OH, 43219
delali.agbenyegah@alliancedata.com
www.delaliagbenyegah.com