Appendix 1. Advertising Models

1) Product-level models

Basu and Batra (1988) proposed a multi-brand advertising budget allocation model called ADSPLIT to overcome the drawbacks of single-product maximizing models that ignore company's total budget constraints. In the optimisation process to maximise business profit, the total budget constraints are factored in. If the sum of the recommended optimal budgets from each product exceeds or does not reach the total budget constraints, budget reallocation process is followed by comparing the marginal return on advertising on each product. While this modeling framework allows for the product with the higher profit contribution to receive more weights in the allocation process, the ADSPLIT model has a few limitations. First, advertising carry-over and decay effects are not considered. Second, in its advertising response model, advertising is treated as a random variable. However, advertising may be determined strategically by a manager, which may result in biased parameter estimates. Therefore, the optimisation results based on these parameters are questionable.

2) Media-level models

As advertising via various media results in differentially effective, managers want to understand this differential impact via different media because they may invest more on the medium that yields higher return. Media-mix models can be used to disentangle the relative contribution from each medium in explaining observed sales variation. For example, Montgomery and Silk (1972) proposed a media-mix model with diminishing returns and advertising decay effects. As one of the earliest empirical study in adverting, the contribution of this study is to compare the relative per dollar contribution of different media (e.g., magazine, direct mail, and brochure) on market share. Naik and Raman (2003) developed the media mix model allowing an interaction effect between TV and print, advertising decay effects, and diminishing returns on advertising. This study provided a counterintuitive suggestion that marketers increase the media share for less effective medium. In the presence of synergy, optimal spending depends not only on its own effectiveness, but also on the spending level for the other activity. Doyle and Saunders (1990) proposed allocation rules for advertising decision process across products and media. Their response function, where sales is a dependent variable and advertising is an independent variable, allows diminishing returns and advertising lagged effects but media interaction is not considered. The optimal advertising amount for each product and medium is proportional to the sum of the response coefficients for the associated products and media, where this approach is similar to Basu and Batra (1988) in that the product/media with the higher profit contribution would get more weights in the allocation process. While this study tries to integrate two different dimensions of product and media, the analysis has ignored the issue concerning the timing of advertising.

3) Time-level models

Based on Koyck (1954)'s lag structure model, Nerlove and Arrow (1962) built a normative model by pointing out that advertising expenditure should be treated the same way as investment in durable goods. They assumed that there is a stock of goodwill that determines the current demand. This stock of goodwill summarises the advertising in the past and, like capital stock, depreciates over time. While the Nerlove-Arrow model (1962) assumes a concave response, the ADBUG model by Little (1970) is a conceptual model, which is flexible to allow both concave and S-shaped depending on the magnitude of advertising response parameter. Both the Nerlove-Arrow model and ADBUG model do not consider that different products compete each other because of the limited advertising budget. Therefore, the decision is limited to a single product because sum of the optimal budget from each product sometimes exceeds the total budget. While the Nerlove-Arrow model assumed the growth rate of goodwill increases linearly with the advertising spending and ignores the saturation points, Naik, Mantrala, and Sawyer (1998) proposed a more flexible advertising response model defined by two differential equations to take into account the dynamics of advertising quality. In their model, the growth rate of goodwill can be linear, concave, or S-shaped so that it is possible to provide whether advertising budget dispersion strategy over time is superior to advertising concentration strategy over time. Naik, Mantrala, and Sawyer (1998) concluded that dispersion strategy (called pulsing) generated higher return on advertising than continuous advertising. However, as a stand-alone advertising response model, this study treats advertising as a random variable as in other advertising models. An alternative approach to assess lagged effects that decay over time is Adstock (Broadbent, 1984; Huang & Sarigöllü, 2014). The Adstock model is based on the assumption that each advertising effort adds to a pre-existing stock of advertising goodwill and the stock decays at a constant rate in the absence of any current advertising.

Appendix 2. Supply Side Model

1) Allocation over products

At the product-level, the allocation task is to divide the total budget for a financial year across multiple products (here, events). The product-level advertising share (product share) is the advertising spending for a given event divided by total advertising spending for the financial year. The product share for each event *i* can be specified as follows:

$$0 \leq \text{product share}_i = \frac{\sum_m \sum_t AD_{imt}}{\sum_i \sum_m \sum_t AD_{imt}} \leq 1 \qquad \text{where i = 1,2,…,N} \qquad \text{(i)}$$

where AD_{imt} is the observed advertising spending for event *i* via medium *m* at time *t*. While the individual event may obtain different product shares, the events belonging to the same group are expected to have similar levels of product share. Because product share is a fraction of the total advertising budget, this study re-parameterises the product share in terms of *u* as follows (Schweidel & Kent, 2010):

$$\text{product share}_i = \frac{\exp(u_i)}{\sum_{i=1}^{N} \exp(u_i)} = \frac{\exp(u_i)}{1+\sum_{i=1}^{N-1} \exp(u_i)} \qquad \text{(ii)}$$

where u_i is the re-parameterised product share for product *i*, and the product share for the base product N is then given by the following:

$$\text{product share}_N = 1 - \sum_{i=1}^{N-1} product\ share_i = \frac{1}{1+\sum_{i=1}^{N-1} \exp(u_i)} \qquad \text{(iii)}$$

From equations (ii) and (iii), u_i can be written as follows:

$$u_i = LN(\frac{product\ share_i}{product\ share_N}) = B\gamma + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2) \qquad \text{(iv)}$$

where LN is natural log, B is a vector of *Gain* groups, and γ captures the relative weights for each group.

2) Allocation over media

Advertising decisions over media are treated similarly to the product-level allocation. Media-level allocation divides the given budget for event *i* into different types of media based on the manager's expectations about each medium. Just as the events in the higher *Gain* group obtain higher product shares, it is possible to anticipate that the medium for which the manager expects to generate more revenue receives a higher media share. Media-level allocation for the highest *Gain* group may be different from that for the lowest *Gain* group, as events in the highest *Gain* group may merit investment in a more expensive medium such as TV while events in the lowest *Gain* group do not. Analogous to the product share, the observed advertising spending for each medium *m* for event *i* proportional to the total advertising budget for the event is the media share as follows:

$$0 \leq \text{media share}_{im} = \frac{\sum_t AD_{imt}}{\sum_m \sum_t AD_{imt}} \leq 1 \qquad \text{where m = 1,2,…,M} \qquad \text{(v)}$$

Then, as in the product share, the re-parameterised media share can be given by the following:

$$w_{im} = LN(\frac{media\ share_{im}}{media\ share_{iM}}) = D\eta + \varepsilon_{im}, \quad \varepsilon_{im} \sim N(0, \Omega) \qquad \text{(vi)}$$

where LN is natural log, D is a vector of the Group for each product *i*, and η captures the relative weights for each medium. If there is no advertising in a medium for a certain product, the log transformation is unavailable. To handle this problem, the media share in equation (v) is modified as follows.

$$0 \leq \text{media share}_{im} = \frac{\max(\sum_t AD_{imt,} 1)}{\sum_m \sum_t \max(AD_{imt,}\ 1)} \leq 1 \qquad \text{(vii)}$$

3) Allocation over time

The allocation decision is how to divide each product's media share into time periods. The advertising dollars assigned to a certain period for a certain product's specific medium equals total budget for a given year multiplied by the product share, the media share, and the time share. Advertising decisions over time periods are similar to the product- and media-level allocation. For product- and media-level allocation, the manager assigns a larger share to the product and media in which she expects a higher return. Likewise, the time periods that the manager expects to generate more revenue and profit receives a higher time share. In this analysis, this research assumes three time periods (early, middle, and late) because the weekly level time share is sparse. For each event *i*, the total ticket sales period is from the first week till the last week of ticket sales (T_i). Then, the early period is the first third of the total ticket sales period, the middle period is the second third of the total ticket sales period, and the late period is the last third. Because a given event *i*'s time share for each time period *t* for media *m* is a fraction of the total advertising budget for media *m* for the given event *i*, this research models the time share as follows:

$$0 \leq \text{time share}_{imt} \leq \frac{\max(AD_{imt,}\ 1)}{\sum_t \max(AD_{imt,}\ 1)} \leq 1 \quad \text{where t} = 1, 2, \ldots, \text{T} \qquad \text{(viii)}$$

Then, as in the product share, the re-parameterised time share, z_{imt}, can be given by the following:

$$z_{imt} = LN(\frac{time\ share_{imt}}{time\ share_{imT}}) = H\tau + \varepsilon_{imt}, \quad \varepsilon_{imt} \sim N(0, \Lambda) \qquad \text{(ix)}$$

where LN is natural log, H is a vector indicating the media and time periods for each product *i*, and τ captures the relative weights for each time period.

4) Simultaneous estimation of the proposed model

In the simultaneous demand- and supply-side model, the dependent variables are observed proportional sales (compared to event *i*'s total capacity) for event *i* at time *t* (y_{it}) and the advertising allocation over event *i*, media *m*, and time *t* (AD_{imt}). The model can be written in hierarchical form as follows:

$$y_{it} \,\big|\, AD_{imt}, \theta_i, \beta, \delta \qquad \text{observed demand} \qquad \text{(x)}$$

$$AD_{imt} \,\big|\, \{\theta_i\}, \beta, \delta, \gamma, \varepsilon_i, \eta, \varepsilon_{im}, \tau, \varepsilon_{imt} \qquad \text{observed advertising spending}$$

$$\{\theta_i\} \,\big|\, \overline{\theta}, \Sigma_\theta \qquad \text{heterogeneity across events}$$

$\varepsilon_i \mid \sigma$ supply-side error for allocation over events

$\varepsilon_{im} \mid \Omega$ supply-side error for allocation over media

$\varepsilon_{imt} \mid \Lambda$ supply-side error for allocation over time

Observed demand for event *i* at time *t* is dependent on the event specific coefficients (θ_i), the different responsiveness for each medium (β_m), the sales spike in the last week (β_{m+1}), the decayed advertising effect (δ) and the explanatory variable (advertising spending AD_{imt}). Observed advertising spending is determined by the set of event specific coefficients ($\{\theta_i\}$), the different responsiveness for each medium (β_m), the sales spike in the last week (β_{m+1}), the decayed advertising effect (δ), and supply-side advertising decision parameters (γ,η,and τ) and supply-side errors (ε_i, ε_{im}, ε_{imt}).

Given the model hierarchy above, the joint distribution of demand and supply is obtained by multiplying the conditional (on advertising spending) demand density by the marginal density of advertising spending. The joint density of all parameters is then:

$$
\begin{aligned}
&f(\theta_i, \overline{\theta}, \Sigma_\theta, \beta, \delta, \gamma, \sigma, \eta, \Omega, \tau, \Lambda \mid \{y_{it}\}, \{AD\}_m) \\
&\propto \prod_i \prod_t f(y_{it} \mid \theta_i, \overline{\theta}, \Sigma_\theta, \beta, \delta) \\
&\times \pi_1(\theta_i \mid \overline{\theta}, \Sigma_\theta) \\
&\times \pi_2(u_i \mid \theta_i, \beta, \delta, \gamma, \sigma) \\
&\times \pi_3(w_{im} \mid \theta_i, \beta, \delta, \eta, \Omega) \\
&\times \pi_4(z_{imt} \mid \theta_i, \beta, \delta, \tau, \Lambda) \\
&\times \pi_5(\beta, \delta, \gamma, \sigma, \eta, \Omega, \tau, \Lambda)
\end{aligned}
$$

where $f(y_{it} \mid \theta_i, \overline{\theta}, \Sigma_\theta, \beta, \delta)$ is the proportional ticket sales at time *t* compared to event *i*'s total capacity; π_1 is the distribution of heterogeneity; π_2 is the density contribution of supply-side error for allocation over products; π_3 is the density contribution of supply-side error for allocation over media; π_4 is the density contribution of supply-side error for allocation over time; and π_5 is the prior contribution.