Multistage Multiproduct Advertising Budgeting

C. Beltran-Royo* H. Zhang† L. A. Blanco‡ J. Almagro§

September 12, 2012

Abstract

We propose and analyze an effective model for the *Multistage* Multiproduct Advertising Budgeting problem. This model optimizes the advertising investment for several products, by considering cross elasticities, different sales drivers and the whole planning horizon. We derive a simple procedure to compute the optimal advertising budget and its optimal allocation. The model was tested to plan a realistic advertising campaign. **We observed that the *multistage* approach may significantly increase the advertising profit, compared to the successive application of the *single stage* approach.**

Key words: Marketing, advertising budgeting, market response models, convex programming, multistage optimization.

1 Introduction

In this article we address the Multistage Multiproduct Advertising Budgeting problem (for short, we will drop one M and call it the MAB problem). More specifically, by advertising budgeting we mean that we wish to decide the capital to be invested on advertising and how to allocate it in an optimal way. By multiproduct we mean that we simultaneously optimize the advertising campaigns of different products within the same company by using different media (television, radio, internet, etc.) and considering *cross product effects* [9]. By multistage, we mean that we optimize the advertising campaigns for the whole planning horizon.

Every year many companies spend thousands of euros to advertise and promote their products. An appropriate optimizing technology can help either to obtain better advertising results for a given budget or to reduce the advertising expenses. We are living the era of *Big Data*, where companies gather and manage huge data bases [13]. By using *market response models* we can transform this raw marketing information into 'ready to use' information [16]. For example, we can model the sales due to advertising as a function of the advertising investment. The direct use of these models is to study different outcomes to take a 'good' decision. A more effective use is to combine them with a utility function to construct an optimization model intended to give a 'best' decision. The relevance of the advertising budgeting problem for the marketing industry is discussed in [13].

The advertising budgeting problem has been addressed in literature from different perspectives. An introductory and interesting paper can be found in [9] where the authors propose a simple formula

*Corresponding author: cesar.beltran@urjc.es, Statistics and Operations Research, Rey Juan Carlos University, Madrid, Spain.

†School of Management, University of Shanghai for Science and Technology, Shanghai, China.

‡Consulting Company Bayes Forecast, Madrid, Spain.

§Consulting Company Bayes Forecast, Madrid, Spain.

for calculating the optimal level of media spending in the case of a single product, single medium and a single stage. The *multiproduct* advertising budgeting problem is analyzed in [10], whereas the *multistage* advertising budgeting problem is studied in [26]. In some cases, due to the complexity of the formulation one has to use heuristic methods which produce good (suboptimal) solutions [29, 22]. In other cases if the complexity of the formulation is moderate, depending on the focus of the model, one can use an optimal control approach [28, 12, 15] or a stochastic optimization approach [1, 11, 2], or a stochastic optimal control approach [5], or a game theory approach [21, 23, 8], or a goal programming approach [3], among others.

As we have mentioned, different aspects of the advertising budgeting problem have been considered. However, as far as we know, a multistage version of the multiproduct advertising budgeting problem has not yet been addressed in literature. Thus, for example, [26] considers a multistage setting but for a single product whereas, [10] considers the multiproduct case but for a single stage. However, as we will illustrate in our case study, relevant savings can be achieved if the planning stages are optimized simultaneously (multistage models) compared to the optimization stage-by-stage (single stage models). Therefore, the main contribution of this paper is to propose and analyze a *multistage version* of the multiproduct advertising budgeting problem.

The objective of this paper is to propose and analyze an effective formulation for the MAB problem. We are interested in calculating the optimal advertising budget and its optimal allocation. We will try to answer the following questions: a) Which is the optimal multiproduct advertising budget for the whole planning horizon? b) Given an advertising budget, how can we optimally allocate it along the planning horizon? c) Is it important to consider multistage models? or on the contrary, is it enough to consider single stage models?

We are also concerned with effectiveness. In [18] it is pointed out that a model that is to be used by a manager should be simple, robust, easy to control, adaptive, as complete as possible and easy to communicate with. In line with this recommendation, the MAB model that we propose is simple but realistic enough to be used in the advertising industry. Furthermore, from a mathematical point of view, it corresponds to a concave maximization problem which is numerically tractable and allows for the computing of a (global) optimal solution with moderate computational effort.

The remainder of the paper is organized as follows. In the second section we will formulate the MAB problem (unconstrained and constrained case). In the third section we will derive a simple procedure to compute the optimal advertising budget and its optimal allocation. In the fourth section a realistic case study will allow us to illustrate the effectiveness of the model as well as the theoretical concepts of the third section. In the last section we will give some conclusions. Appendix A contains the proofs of all the theoretical results of this article. Appendix B contains the data for the case study in section fourth.

2 Problem formulation

In order to formulate the Multistage Multiproduct Advertising Budgeting (MAB) problem, we distinguish between baseline sales (sales that one would expect without advertising) and *sales due to advertising*. The objective of the MAB model is to maximize the profit of the sales due to advertising. Expressed in a different way, we will consider the profit of the sales due to advertising as the measure of the advertising effectiveness. Prior to formulate the MAB problem, we briefly review some key concepts in the advertising industry (see, for example, [16] for more details). *Reach* is the proportion of the target audience exposed to at least one insertion of the advertisement [9]. **We call this proportion the reach audience. *Frequency* is the average number of times a person from the reach audience is**

exposed to an advertisement. *Exposure* to an advertisement involves reach and frequency and can be measured in Gross Rating Points: GRPs $\equiv$ reach $\times$ frequency. For example, a purchase of 100 GRPs could mean that 100% of the market is exposed once to an advertisement or that 50% of the market is exposed twice [16]. According to [2], advertising is measured in GRPs and not in euros, since there are two advantages of using GRPs. First, GRPs provide a more accurate picture of advertising input than advertising expenditures since it is not clear how much advertising exposure can be purchased for a given budget. Second, most media buying is done in terms of GRPs and managers evaluate the effectiveness of their campaigns in terms of demand generated per GRP.

The impact of advertising effort spreads over time and the advertising effort in one stage is cumulated with past advertising efforts. In this respect, we will use the variable so called *adstock* [4, 16], which is a measure of the past and current advertising effort that is effective in the current stage. For brand management, market response models provide a basis for fine tuning marketing mix variables (*marketing mix* for short), such as price, sales promotions, advertising copy, media selection, timing, and other brand-specific marketing factors. Furthermore, the marketing mix has to take into account the *market segmentation,* that is, the distinct consumer groups, each one characterized by the same needs and behaviors [7, 6]. The largest category of empirical response models are those dealing with sales and market share as dependent variables. Companies want to know what influences their sales (the *sales drivers* or, for short, *drivers*). They want to know how to set the marketing mix so that they can control their sales. One of the limitations of the MAB model that we present is that it does not take the product price as a sales driver, that is, as a decision variable (prices are input data). All the other above mentioned drivers (sales promotions, advertising copy, media selection and timing) can be taken into account in our MAB model. Furthermore, by an abuse of terminology, when we talk about sales due to advertising, we refer to sales due to these drivers (analogously with advertising budgeting, advertising allocation, etc.).

2.1 Notation

In our formulation of the MAB problem we consider:

Indexes:

t	Index for stages, $t \in \mathcal{T} = \{1, \ldots, T\}$.	
i	Index for products, $i \in \mathcal{I} = \{1, \ldots, I\}$.	
j	(Auxiliary) index for products, $j \in \mathcal{I}$.	
k	Index for sales drivers for product i, $k \in \mathcal{K}_i = \{1, \ldots, K_i\}$.	$i \in \mathcal{I}$
$\mathcal{TIK}_i$	Stands for $\mathcal{T} \times \mathcal{I} \times \mathcal{K}_i$.	$i \in \mathcal{I}$

Parameters:

a_{tijk}	Sales of product i in stage t induced by one unit of driver jk **where** $j \neq i$ (note that, for simplicity, to refer to the k-th driver of product j, we use the expression 'driver jk'), $a_{tijk} \in \mathbb{R}$.	$tijk \in \mathcal{TIIK}_i$
c_{tik}	Cost of driver ik in stage $t \in \{1, \ldots, T+1\}$, $c_{tik} > 0$.	$ik \in \mathcal{IK}_i$
δ_{ik}	Retention rate of the advertising effort from stage to stage for driver ik, $\delta_{ik} \in \ ]0, 1[$.	$ik \in \mathcal{IK}_i$
p_{ti}	Profit per unit of product i in stage t, $p_{ti} > 0$.	$ti \in \mathcal{TI}$
$\widetilde{x}_{0ik}$	Advertising effort of driver ik previous to the first stage, $\widetilde{x}_{0ik} \geq 0$.	$ik \in \mathcal{IK}_i$

Functions:

R_{tik}	Sales of product i in stage t due to driver ik.	$tik \in \mathcal{TIK}_i$.
L_{tijk}	Sales of product i in stage t due to driver jk **where** $j \neq i$ ('cross product effect').	$tijk \in \mathcal{TIIK}_j$
S_{ti}	Sales of product i in stage t due to advertising.	$ti \in \mathcal{TI}$
P	Profit function (information aggregated by products).	
$\widetilde{Q}$	Profit function (information aggregated by drivers).	

Variables:

g_{tik}	Investment in GRPs of driver ik in stage t.	$tik \in \mathcal{TIK}_i$
x_{tik}	Cumulated advertising effort of driver ik in stage t ('adstock').	$tik \in \mathcal{TIK}_i$

2.2 The multiproduct sales response function

In the MAB problem one can maximize different utility functions. One common approach is to maximize the advertising profitability [9]:

$$\text{Advertising profitability} = \text{Profit} \times \text{Sales}_g - \text{Cost}_g$$

where 'Profit' is the profit per unit, 'Sales$_g$' corresponds to the sales due to advertising which is a function of g, the number of GRPs, and 'Cost$_g$' is the cost of buying g GRPs. To model Sales$_g$, sales response models can be constructed [16]. More specifically, for each pair $ti \in \mathcal{TI}$, the multiproduct sales response function S_{ti} accounts for R_{tik}, the sales on product i due to its own drivers, plus L_{tijk}, the sales on product i due to the other product drivers (cross product sales):

$$S_{ti}(x) = \sum_{k \in \mathcal{K}_i} R_{tik}(x_{tik}) + \sum_{jk \in \mathcal{IK}_j,\, j \neq i} L_{tijk}(x_{tjk}), \qquad ti \in \mathcal{TI}, \tag{1}$$

where vector x **accounts for** $(x_{tik})_{tik \in \mathcal{TIK}_i}$.

In general, the single product sales response functions R_{tik} correspond to increasing concave functions which model diminishing returns. According to [16], a typical choice, among others, is the so called 'modified exponential' function

$$R_{tik}(x) = \alpha_{tik}(1 - e^{-\beta_{tik} x}), \qquad tik \in \mathcal{TIK}_i.$$

For more details see the case study in Section 4. On the other hand, the cross product sales response functions L_{tijk} model the cross elasticities among the products due to relationships of complementarity or substitution [10]. In the case of complementarity (positive elasticity), advertising on product j increases sales of product i and this cross effect can be modeled by an increasing concave function. On the contrary, in the case of substitution (negative elasticity), advertising on product j reduces sales of product i (this cross effect is known as *cannibalization* [13]). The cannibalization effect can be modeled by a decreasing convex function. If this function is strictly convex, the resulting multiproduct sales response function S_{ti} may not be concave. As is well known, concavity of the objective function is a desirable property in a maximization problem since it guarantees *global* optimality (assuming a convex feasible domain). On the other hand, the cross product effects are usually small relative to the direct advertising effects modeled by R_{tik}, as in the case study presented in Section 4. Thus, to ensure concavity and assuming that the cross effects are small, we will approximate the cross advertising effects by linear functions:

$$L_{tijk}(x) = a_{tijk}\, x, \qquad tijk \in \mathcal{TIIK}_j, j \neq i. \tag{2}$$

To simplify the notation, we will assume that coefficients a_{tijk} with $j = i$ exist and that they are equal to 0. This will allow us to drop the condition $j \neq i$ in equations that include a_{tijk} and $L_{tijk}(x)$ as for example equations (1) and (2).

2.3 The unconstrained case

In order to calculate the optimal advertising budget we use the previous multiproduct sales response functions in the first MAB model. To distinguish this MAB version from other versions that will appear in the paper, we name it MAB$_{PU}$, where we use P to indicate that it is based on the profit function $P(g)$ that we define below, and we use U to indicate that it is an unconstrained version. Problem MAB$_{PU}$ is defined as

$$\max_{g \in \mathbb{R}^n} P(g),$$

where vector $g = (g_{tik})_{tik \in \mathcal{TIK}_i}$ and $n = T(\sum_{i \in \mathcal{I}} K_i)$.

The profit function is defined as a function of g :

$$P(g) = \sum_{ti \in \mathcal{TI}} p_{ti}\, S_{ti}(x(g)) - \sum_{tik \in \mathcal{TIK}_i} c_{tik}\, g_{tik} + \sum_{tik \in \mathcal{TIK}_i} V_{tik}(x(g)),$$

where

$$\begin{aligned}
S_{ti}(x(g)) &= \sum_{k \in \mathcal{K}_i} R_{tik}(x_{tik}(g)) + \sum_{jk \in \mathcal{IK}_j} L_{tijk}(x_{tjk}(g)), && ti \in \mathcal{TI}, && (3)\\
x(g) &= (x_{tik}(g))_{tik \in \mathcal{TIK}_i}, && && (4)\\
x_{tik}(g) &= \delta_{ik} x_{t-1,ik}(g) + g_{tik}, && tik \in \mathcal{TIK}_i, && (5)\\
x_{0ik} &= \widetilde{x}_{0ik}, && ik \in \mathcal{IK}_i, && (6)\\
V_{1ik}(x(g)) &= -c_{1ik}\, \delta_{ik} \widetilde{x}_{0ik}, && ik \in \mathcal{IK}_i, && (7)\\
V_{tik}(x(g)) &= 0, && 1 < t < T, ik \in \mathcal{IK}_i, && (8)\\
V_{Tik}(x(g)) &= c_{T+1,ik}\, \delta_{ik}\, x_{Tik}(g), && ik \in \mathcal{IK}_i. && (9)
\end{aligned}$$

(3) Accounts for **the** sales of product i in stage t due to advertising.

(4) **Accounts for** the addstock vector.

(5) Expresses the dynamic behavior of variable adstock for driver ik.

(6) **Sets the** initial level of the adstock variable for driver ik.

(7) **Determines the** accounting cost of the initial adstock level for driver ik.

(8) For notational convenience, we set these values to 0.

(9) **Determines the** accounting value of the final adstock level for driver ik.

Note that (5) is the discrete time version of the Nerlove-Arrow continuous time model for the adstock variable [24]. Further details can be found in [19, 17, 30, 20].

In this paper any vector, say g, is assumed to be a column vector and its transpose will be indicated by g^T. Of course, this notation has nothing to do with $\mathbb{R}^T$, the **Euclidean** space of dimension T. Let us now see that each function $x_{tik}(g)$ is a linear function of g. First, we need some definitions. For each driver $jk \in \mathcal{IK}_j$, we define the vector g_{jk} as

$$g_{jk}^T = (g_{1jk}, g_{2jk}, \dots, g_{Tjk}),$$

which accounts for the investment on driver jk at each stage. Then, the allocation vector g can be expressed as:

$$g^T = (g_{tjk})_{tjk \in \mathcal{TIK}_j} = (g_{11}^T,\ g_{12}^T, \dots, g_{1K_1}^T, \dots, g_{I1}^T,\ g_{I2}^T, \dots, g_{IK_I}^T).$$

Furthermore, for each $tjk \in \mathcal{TIK}_j$, we define the vector

$$u_{tjk}^T = (\delta_{jk}^{t-1},\ \delta_{jk}^{t-2}, \dots,\ \delta_{jk},\ 1,\ 0, \dots,\ 0) \in \mathbb{R}^T.$$

Note that, if we denote the τ-th component of vector u_{tjk} by $u_{tjk}(\tau)$, then

$$u_{tjk}(\tau) = \delta_{jk}^{t-\tau}$$

for $1 \leq \tau \leq t$ and $u_{tjk}(\tau) = 0$ in the other cases.

Lemma 1 *For any $tjk \in \mathcal{TIK}_j$, adstock function $x_{tjk}(g)$, can be expressed as an affine function of g as follows:*

$$x_{tjk}(g) = v_{tjk} + u_{tjk}^T g_{jk},$$

where $v_{tjk} = \delta_{jk}^t\, \widetilde{x}_{0jk}$.

The proof of this Lemma and the proof of all the remaining theoretical results can be found in Appendix A.

Proposition 1 *If functions $R_{tik}(x)$ are concave for all $tik \in \mathcal{TIK}_i$ then, $P(g)$ is concave and problem MAB$_{PU}$ is a concave maximization problem (concave objective function and convex decision domain). Therefore, every local optimal solution is also global.*

2.4 The constrained case

Companies often need to allocate advertising budgets combined with other investment limitations. In this case we have to impose some constraints to problem MAB$_{PU}$. The constrained version of MAB$_{PU}$ is the following problem that we name MAB$_{PC}$:

$$\begin{aligned} \max \quad & P(g) \\ \text{s.t.} \quad & g \in \mathcal{D}_g \end{aligned}$$

where $\mathcal{D}_g$ is the feasible set for vector g.

Proposition 2 *If functions $R_{tik}(x)$ are concave for all $tik \in \mathcal{TIK}_i$ and if $\mathcal{D}_g$ is convex, then MAB$_{PC}$ is a concave maximization problem (concave utility function and convex decision domain). Therefore, every local optimal solution is also global.*

A typical (convex) domain only considers variable bounds and linear constraints, that is,

$$\mathcal{D}_g = \left\{ g \in \mathbb{R}^n \quad | \quad Ag \leq b, \qquad \underline{g} \leq g \leq \overline{g} \right\},$$

as for example

$$\mathcal{D}_g = \left\{ g \in \mathbb{R}^n \quad | \quad \sum_{tik \in \mathcal{TIK}_i} c_{tik}\, g_{tik} \leq b, \qquad \underline{g} \leq g \leq \overline{g} \right\}, \tag{10}$$

where we have to allocate the total advertising budget b for all the drivers and along all the stages. This is the situation analyzed in the case study of Section 4.3. Of course, many other kind of constraints could be considered: the company could be interested in limiting the advertising budget within each planning stage or it could be interested in imposing a threshold for the adstock variables at the end of the planning horizon, or it could impose a budget for each driver, etc.

Note that for the constrained problem MAB$_{PC}$, it is not assured the existence of an optimum. Taking into account that the objective function $P(g)$ is concave, therefore continuous, by the Weierstrass Theorem an optimal solution to problem MAB$_{PC}$ will always exist, if the constraint set $\mathcal{D}_g$ is compact (closed and bounded) [27].

3 The optimal advertising budget

In this section we derive a simple procedure to compute the optimal advertising budget and its optimal allocation. Allocation rules for the constrained Multistage Multiproduct Advertising Budgeting (MAB) problem depend on the domain and therefore, in general, one cannot derive a closed formula. In contrast, for the unconstrained case, one can derive the optimal allocation rule and, as a byproduct, the optimal advertising budget. In formulations MAB$_{PU}$ and MAB$_{PC}$ we aggregated the information by products and defined, at each stage, one multiproduct sales response function per product $S_{ti}(g)$. Another possibility is to aggregate the information by drivers and define, at each stage, one sales response function per driver, say $Q_{tjk}(g)$. This reformulation turns out to be useful to calculate the optimal advertising allocation g^* as we will see in this section. With this objective in mind, we rewrite problem MAB$_{PU}$ into the equivalent problem MAB$_{\widetilde{Q}U}$ (see Lemma 2):

$$\max_{g \in \mathbb{R}^n} \widetilde{Q}(g) \tag{11}$$

where we use the following functions

$$\begin{aligned}
\widetilde{Q}(g) &= \sum_{jk \in \mathcal{IK}_j} \widehat{Q}_{jk}(g), \\
\widehat{Q}_{jk}(g) &= \sum_{t \in \mathcal{T}} Q_{tjk}(g), && jk \in \mathcal{IK}_j, \\
Q_{tjk}(g) &= p_{tj}\, R_{tjk}(x_{tjk}(g)) + \sum_{i \in \mathcal{I}} p_{ti}\, L_{tijk}(x_{tjk}(g)) - c_{tjk}\, g_{tjk} + V_{tjk}(x_{tjk}(g)), && tjk \in \mathcal{TIK}_j, \\
x_{tjk}(g) &= \delta_{ik} x_{t-1,jk}(g) + g_{tjk}, && tjk \in \mathcal{TIK}_j, \\
x_{0jk}(g) &= \widetilde{x}_{0jk}, && jk \in \mathcal{IK}_j.
\end{aligned}$$

Analogously, we can rewrite problem MAB$_{PC}$ into the equivalent problem MAB$_{\widetilde{Q}C}$:

$$\begin{aligned}
\max \quad & \widetilde{Q}(g) \\
\text{s.t.} \quad & g \in \mathcal{D}_g.
\end{aligned} \tag{12}$$

Lemma 2 *$\widetilde{Q}(g) = P(g)$ for all $g \in \mathbb{R}^n$.*

It is clear that problem $\mathrm{MAB}_{\widetilde{Q}U}$ is decomposable by drivers (we have one subproblem per driver jk):

$$\max_{g\in\mathbb{R}^n} \widetilde{Q}(g) = \max_{g\in\mathbb{R}^n} \left(\sum_{jk\in\mathcal{IK}_j} \widehat{Q}_{jk}(g_{jk}) \right) = \sum_{jk\in\mathcal{IK}_j} \left(\max_{g_{jk}\in\mathbb{R}^T} \widehat{Q}_{jk}(g_{jk}) \right).$$

For each driver jk we can define subproblem $\mathrm{MAB}_{\widehat{Q}Ujk}$ as

$$\max_{g_{jk}\in\mathbb{R}^T} \widehat{Q}_{jk}(g_{jk}) = \max_{g_{jk}\in\mathbb{R}^T} \sum_{t\in\mathcal{T}} Q_{tjk}(g_{jk}).$$

The remainder of the section will be devoted to solve $\mathrm{MAB}_{\widehat{Q}Ujk}$. Since all the subproblems have the same structure, we will drop indexes jk in order to lighten the notation. For example, instead of $\mathrm{MAB}_{\widehat{Q}Ujk}$ we will write $\mathrm{MAB}_{\widehat{Q}U}$ defined as

$$\max_{g\in\mathbb{R}^T} \sum_{t\in\mathcal{T}} Q_t(g).$$

Note that with some abuse of notation, in subproblem $\mathrm{MAB}_{\widehat{Q}U}$ vector g has dimension T and in problem $\mathrm{MAB}_{\widetilde{Q}U}$, vector g has dimension n.

Let us see how we can write Q_t in a more compact way. First, we define functions

$$r_{tjk}(x_{tjk}(g)) = p_{tj} R_{tjk}(x_{tjk}(g)) + \sum_{i\in\mathcal{I}} p_{ti} L_{tijk}(x_{tjk}(g)), \qquad tjk \in \mathcal{TIK}_j.$$

Second, we can drop the jk subindex and suppose that the reference product j is product 1:

$$r_t(x_t(g)) = p_{t1} R_t(x_t(g)) + \sum_{i\in\mathcal{I}} p_{ti} L_{ti}(x_t(g)), \qquad t \in \mathcal{T}.$$

By combining this definition with the definition of $Q_t(g)$ we can write:

$$Q_t(g) = r_t(x_t(g)) - c_t\, g_t + V_t(x_t(g)) \qquad t \in \mathcal{T},$$

where $x_t(g) = v_t + u_t^T g$.

Assumption 1 We say that function $R(x)$ satisfies Assumption 1 with constant C, if:

1. $R(0) = 0$ and $R(x) > 0$ for all $x \in J =]0, +\infty[$.
2. $\lim_{x\to 0,\ x>0} R'(x) > C$.
3. $\lim_{x\to +\infty} R'(x) < C$.
4. For all $x \in J$, we have that $R''(x)$ exists and $R''(x) < 0$ (therefore $R(x)$ is strictly concave).

Given problem MAB, it is convenient to define the following auxiliary constants:

$$\begin{aligned}
\widetilde{c}_t &= c_t - \delta^{T+1-t}\, c_{T+1}, && t \in \mathcal{T},\\
\widetilde{c}_{T+1} &= 0, \\
K_t &= \frac{1}{p_{t1}} \left(\widetilde{c}_t - \delta \widetilde{c}_{t+1} - \sum_{i\in\mathcal{I}} p_{ti} a_{ti} \right), && t \in \mathcal{T}.
\end{aligned}$$

Lemma 3 *Given $t \in \mathcal{T}$, if function $R_t(x)$ satisfies Assumption 1 with constant K_t, then there exists a unique $x_t^* > 0$ such that*

$$r_t'(x_t^*) = \widetilde{c}_t - \delta\, \widetilde{c}_{t+1},$$

where $r_t(x_t) = p_{t1} R_t(x_t) + \sum_{i \in \mathcal{I}} p_{ti} a_{ti} x_t$.

Proposition 3 *For a given component t_0 of vector $g \in \mathbb{R}^T$ we have that:*

$$\frac{\partial \widehat{Q}(g)}{\partial g_{t_0}} \quad = \quad \sum_{t_0 \le t \le T} r_t'(u_t^T g + v_t) \cdot \delta^{t-t_0} - \ \widetilde{c}_{t_0}.$$

Proposition 4 *Let us consider that functions $R_t(x)$, with $t \in \mathcal{T}$, fulfil Assumption 1 with constant K_t. Then*

1. *The system of non-linear equations on $(x_1, \ldots, x_T)$:*

   $$\sum_{t_0 \le t \le T} r_t'(x_t) \cdot \delta^{t-t_0} - \widetilde{c}_{t_0} = 0, \qquad t_0 \in \mathcal{T},$$

   has a unique solution vector that we denote by $(x_1^*, \ldots, x_T^*)^T$. Furthermore, $x_t^* > 0$ for all $t \in \mathcal{T}$.

2. *The linear system on g*

   $$u_t^T g + v_t = x_t^*, \qquad t \in \mathcal{T}, \tag{13}$$

   has a unique solution vector that we denote by g^*.

Theorem 1 *Let us consider subproblem MAB$_{\widehat{Q}U}$ where we assume that functions $R_t(x)$ fulfil Assumption 1 with constant K_t, for all $t \in \mathcal{T}$. Then:*

1. *There exists a unique vector $g^* \in \mathbb{R}^T$ such that*

   $$\nabla \widehat{Q}(g^*) = 0.$$

   Therefore g^* is the unique global maximizer of $\widehat{Q}(g)$.

2. *Maximizer g^* can be computed as follows.*

   (a) *Calculate the scalars $x_1^*, \ldots, x_T^*$, by solving the scalar equations*

   $$r_t'(x_t) = \widetilde{c}_t - \delta \widetilde{c}_{t+1}, \qquad t \in \mathcal{T}.$$

   (b) *Compute*

   $$g_t^* = x_t^* - \delta\, x_{t-1}^*, \qquad t \in \mathcal{T}.$$

   where $x_0^* = \widetilde{x}_0$.

3. *The optimal budget is $b^* = \sum_{t \in \mathcal{T}} c_t g_t^*$ and its optimal allocation is given by g^*.*

Note that Theorem 1 does not guarantee a positive g^*. In the case where some of the components of g^* were non positive, we should impose sign constraints ($g \ge 0$) to subproblem MAB$_{\widehat{Q}U}$ and solve the corresponding constrained subproblem MAB$_{\widehat{Q}C}$.

4 Case Study

The objective of this section is to show by example the improvements that a multistage formulation can bring to the single stage formulation of the multiproduct advertising budgeting problem. We try to answer the following questions: a) Which is the optimal multiproduct advertising budget for the whole planning horizon? b) Given an advertising budget (optimal or not), how can we optimally allocate it along the planning horizon? c) Is it important to consider multistage models? or on the contrary, is it enough to consider single stage models?

The Multistage Multiproduct Advertising Budgeting (MAB) instances that we present correspond to problems $\text{MAB}_{\widetilde{Q}U}$ and $\text{MAB}_{\widetilde{Q}C}$ (equations (11) and (12) respectively). These instances were derived from a real-life case addressed at the consulting company Bayes Forecast[1], to plan the advertising campaign for a leading fast moving consumer goods company. The instance we present here, considers a twelve months planning horizon ($T = 12$), two products that we denoted by P1 and P2 ($I = 2$) and two sales drivers per product ($K_1 = K_2 = 2$). For each product, the first driver corresponds to TV advertising and the second driver corresponds to in-store promotions. These parameters as well as the profit per unit sold are summarized in Table 11 of Appendix B (note that profits are constant along all the stages). The values of the advertising retention rate δ_{ik} and the values of the initial adstock $\widetilde{x}_{0ik}$, are in Table 4 and Table 5, respectively.

4.1 The 'modified exponential' response function

In our case study we use the single product sales response function denominated 'modified exponential' [16], where sales are modeled as a function of the advertising effort x as follows

$$R_t(x) = \alpha_t(1 - e^{-\beta_t x}), \qquad t \in \mathcal{T}.$$

Note that functions $R_{tik}(x)$ were introduced in Section 2.2, and in this section we drop indexes ik for simplicity of exposition. The positive parameter α_t corresponds to the saturation level. This means that no matter how much marketing effort is expended, the sales due to advertising will not be higher than α_t (Table 6). The positive parameter β_t regulates the diminishing return to scale (Table 7). On the other hand, the cross product sales effects between P1 and P2 are due to substitution, i.e., advertising on, say P1, will increase P1 sales but will reduce P2 sales and viceversa. These effects are known as *cannibalization* [13]. In this case study, the cannibalization effects is small relative to the direct advertising effects modeled by $R_{tik}(x)$. Thus, as mentioned in Section 2.2, to ensure concavity the cannibalization effects can be approximated by linear functions $L_{ti}(x) = a_{ti}x$. Under cannibalization, the cross product effect parameters a_{ti} are negative (Tables 8 and 9).

In Figure 1 we can see a unidimensional profit function based on the *modified exponential* sales response function for $t = 1$ and $\widetilde{x}_0 = 0$ (in this case $x_1 = g_1$):

$$\begin{aligned} Q_1(x) &= p_{11}\, R_1(x) + p_{12}L_{12}(x) - c_1 x \\ &= p_{11}\, \alpha_1(1 - e^{-\beta_1 x}) + p_{12}a_{12}x - c_1 x \end{aligned}$$

with $p_{11} = 1.75, p_{12} = 1.40, c_1 = 480, \alpha_1 = 345,000, \beta_1 = 0.010$ and $a_{12} = -0.0001$.

Next lemma shows that under cannibalization and some mild assumptions, the modified exponential response function fulfils Assumption 1. This implies that the optimal allocation rule stated in Theorem 1 can be applied for subproblem $\text{MAB}_{\widehat{Q}U}$ based on the modified exponential model.

[1]Bayes Forecast S.L., Madrid (Spain), www.bayesforecast.com

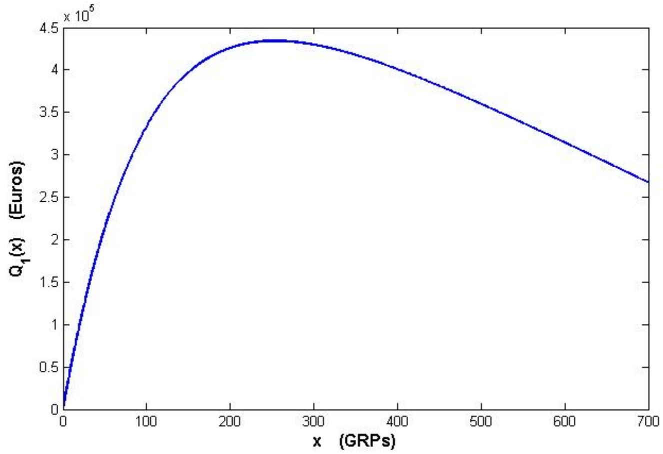

Figure 1: Unidimensional profit function $Q_1(x)$ based on the 'modified exponential' sales response function.

Lemma 4 *Let us assume that $\widetilde{c}_t - \delta\, \widetilde{c}_{t+1} > 0$, that $\alpha_t \beta_t > K_t$ for all $t \in \mathcal{T}$ and that $a_{ti} < 0$ for all $ti \in \mathcal{TI}$. Then, for all $t \in \mathcal{T}$, function*

$$R_t(x) = \alpha_t(1 - e^{-\beta_t x})$$

with $\alpha_t, \beta_t > 0$, fulfils Assumption 1 with constant K_t.

4.2 Determining the optimal budget

When solving the MAB problem one can compute the optimal budget considering all the stages simultaneously (multistage optimization). Alternatively, one can compute the optimal budget stage-by-stage (we call it stage-by-stage optimization). This corresponds to solving a sequence of 12 problems MAB with $T = 1$. As is well known, only the multistage optimization gives an optimal solution for the MAB problem. The stage-by-stage optimization usually gives a solution which is good but suboptimal. One of the main objectives of our case study is to show by example that relevant losses can be incurred by implementing the stage-by-stage optimal allocation. In Table 10 we display the cost of one GRP depending on the driver type. Note that these costs define alternating low and high price periods of two months. This cost structure will allow us to assess the ability of the two optimization approaches to cope with dynamic GRP prices.

4.2.1 Multistage approach

Proposition 5 *Let us consider subproblem MAB$_{\widehat{Q}U}$ based on the modified exponential sales response function, i.e.,*

$$R_t(x) = \alpha_t(1 - e^{-\beta_t x}), \qquad t \in \mathcal{T}.$$

Let us assume that $R_t(x)$ fulfils Assumption 1 with constant K_t, for all $t \in \mathcal{T}$. Then, the optimal multistage budget allocation g^* can be computed as follows. For each $t \in \mathcal{T}$ compute:

$$g_t^* = x_t^* - \delta\, x_{t-1}^*,$$
$$x_t^* = \frac{1}{\beta_t} \ln\left(\frac{p_{t1}\alpha_t\beta_t}{\widetilde{c}_t - \delta\, \widetilde{c}_{t+1} - \sum_{i\in\mathcal{I}} p_{ti} a_{ti}}\right),$$

where $x_0^* = \widetilde{x}_0$.

Table 1: Optimal budget (results in euros): The multistage approach gives the MAB optimal budget (3,818,334 euros) which increases by 5.52% (1,218,625 euros) the profit given by the stage-by-stage approach.

	Stage-by-stage	Multistage	Variation
Budget	3,929,416	3,818,334	-2.83%
Cannibalization	-401,809	-399,039	-0.69%
Profit	22,058,084	23,276,709	5.52%

By using Proposition 5 we computed the optimal multistage budget (Table 1). As a by-product of the optimal budget, we obtained the optimal budget allocation. The optimal budget allocation corresponding to driver 1 of P1, can be seen in Figure 2 (solid line). In the next section we will compare this approach with the stage-by-stage approach.

4.2.2 Stage-by-stage approach

Proposition 6 *Let us consider subproblem MAB*$_{\widehat{Q}U}$ *based on the modified exponential sales response function, i.e.,*

$$R_t(x) = \alpha_t(1 - e^{-\beta_t x}), \qquad t \in \mathcal{T}.$$

Let us assume that $R_t(x)$ fulfils Assumption 1 with constant

$$K_t' = \frac{1}{p_{t1}} \left(c_t - \delta c_{t+1} - \sum_{i \in \mathcal{I}} p_{ti} a_{ti} \right)$$

for all $t \in \mathcal{T}$. Then, the optimal stage-by-stage budget allocation g^* can be computed as follows. For each $t \in \mathcal{T}$ compute:

$$\begin{aligned} g_t^* &= x_t^* - \delta\, x_{t-1}^*, \\ x_t^* &= \frac{1}{\beta_t} \ln \left(\frac{p_{t1}\alpha_t\beta_t}{c_t - \delta\, c_{t+1} - \sum_{i \in \mathcal{I}} p_{ti} a_{ti}} \right), \end{aligned}$$

where $x_0^* = \widetilde{x}_0$.

Note that in the multistage case we used the same formulas but with $\widetilde{c}_t$ instead of c_t.

By using Proposition 6 we computed the 'optimal' stage-by-stage budget. In Table 1 we compare the results obtained by the two approaches: Stage-by-stage versus multistage. This table shows that the MAB optimal budget for the 12 months is 3,818,334 euros which produces a profit 5,52% (1,218,625 euros) higher than the suboptimal budget proposed by the stage-by-stage approach.

The multistage approach gives a better profit than the stage-by-stage approach, as one would expect. Of course, the multistage improvement observed in this case study (5,52%) does not guarantee this level of improvement for all the MAB instances. It only shows that the multistage approach may obtain significatively better results than the single stage approach. The message of this case study is that it is worthy to use the multistage approach.

Although a thorough sensitivity analysis is out of the scope of this article, in Table 2 we show the results of a very basic sensitivity analysis. Given that the parameters are grouped in vectors, e.g.

Table 2: Sensitivity analysis (results in %): In column 'Variation –25%' we have the optimal profit improvement of the multistage approach compared to the stage-by-stage approach when we reduced by 25% de reference values of the corresponding vector parameter. Analogously for column 'Variation +25%'.

Vector parameter	Symbol	Variation –25%	Variation +25%
Profit	p	5,96	5,27
Cost	c	5,21	5,85
Saturation	α	5,91	5,29
Diminishing return to scale	β	5,91	5,29
Cannibalization	a	5,55	5,50
Retention rate	δ	5,38	-

$\alpha = (\alpha_{tik})_{tik \in \mathcal{TIK}_i}$, we considered appropriate to simultaneously perturb all the components of a vector parameter by the same percentage, either -25% or +25% (one vector parameter each time). As reference values we took the parameter values listed in Appendix B (the ones used to obtain the 5,52% improvement). In column 'Variation –25%' we have the optimal profit improvement (in %) of the multistage approach compared to the stage-by-stage approach, when we reduced by 25% de reference values of the corresponding vector parameter. Analogously for column 'Variation +25%'. **The entry for the bottom right corner could not be calculated since in this case functions $R_t(x)$ do not fulfill condition 3 of Assumption 1 with constants K_t and K'_t for some $t \in \mathcal{T}$ and therefore Propositions 5 and 6 could not be applied.**

Finally, as previously mentioned, the advertising investment corresponding to driver 1 of P1 can be seen in Figure 2. The optimal advertising allocation proposed by the multistage approach (solid lines) is better adapted to price changes, since it allocates more GRPs at low price periods and less GRPs at high price periods.

4.3 Determining the optimal allocation for a given budget

To compute the optimal budget one assumes that there is no limit on the available budget, as we did in the previous section (unconstrained optimization case). However, very often companies need to allocate a reduced advertising budget (constrained optimization case). For example, if we were limited to 50% of the optimal budget computed in the previous section, we should solve problem MAB$_{\widehat{Q}C}$ with $b = 1,909,167$ euros, $\underline{g} = 0$ GRPs and $\overline{g} = \infty$ GRPs in (10). In Table 3 we compare the results obtained by the two approaches: unconstrained versus constrained. This table shows that by reducing 50% the optimal budget for the 12 months, we reduce by 6.18% (1,437,892 euros) the optimal profit. In Figure 3 we compare the budget allocation corresponding to driver 1 of P1 obtained by the two approaches. We observe that the reduced budget allocation proposes greater relative reductions in advertising investment at high price stages.

The computations have been conducted on a laptop under Windows XP, with a processor Intel Core Duo 2.40GHz and with 3.48 GB of RAM. Programs have been written in Matlab (R2008b) and problem MAB$_{\widehat{Q}C}$ (constrained concave maximization problem) has been solved by function `fmincon` from the Matlab Optimization Toolbox (V4.1) with default parameters. `fmincon` uses a (sequential quadratic programming) SQP method based on the (Broyden-Fletcher-Goldfarb-Shanno) BFGS formula [14, 25]. In our case the optimal solution of our MAB$_{\widehat{Q}C}$ instance (12 stages, 2 products and 2 drivers per product) has been obtained after 42 SQP iterations (1.4 seconds).

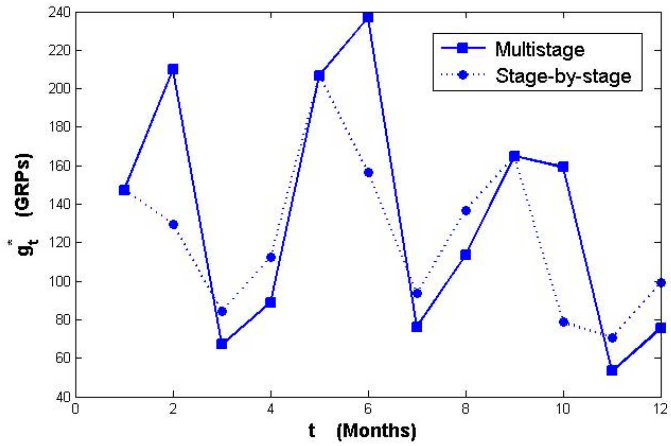

Figure 2: Optimal budget allocation g_t^* for driver 1 of product 1 (unconstrained case): The multistage approach is better adapted to price changes than the stage-by-stage approach. The former allocates more GRPs in low price periods (months 1-2, 5-6 and 9-10) and less GRPs in high price periods (months 3-4, 7-8 and 11-12).

Table 3: Reduction of the optimal budget (results in euros): By reducing 50% the optimal budget we reduce 6% the profit.

	Optimal budget	Reduced budget	Variation
Budget	3,818,334	1,909,167	-50.00%
Cannibalization	-399,039	-231,532	-41.98%
Profit	23,276,709	21,838,817	-6.18%

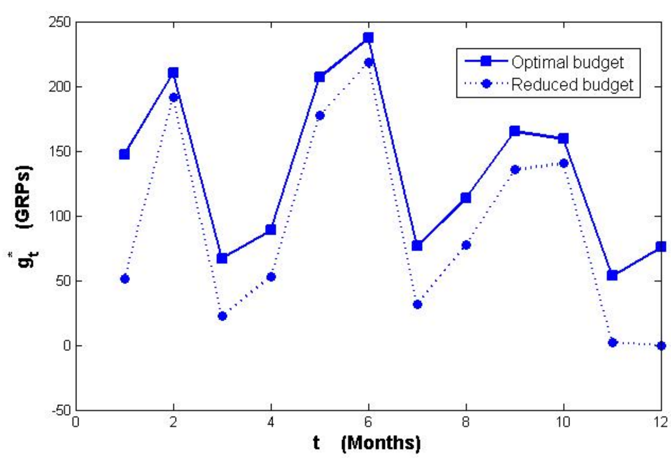

Figure 3: Optimal budget allocation g_t^* for driver 1 of product 1 (optimal budget and reduced budget): The optimal allocation of the reduced budget is attained by allocating less GRPs, especially in high price periods (months 3-4, 7-8 and 11-12).

5 Concluding remarks

The main objective of this paper is to propose and analyze an effective formulation for the multistage multiproduct advertising budgeting (MAB) problem. This model is intended to optimize the advertising investment for several products, by considering ‘cross product effects’, different drivers and the whole planning horizon (multistage). As far as we know, a *multistage multiproduct* version of the advertising budgeting problem has not been considered in literature.

From a theoretical point of view, firstly, we have shown that the MAB model that we propose corresponds to a maximization concave problem (constrained and unconstrained version). Therefore this model is numerically tractable and produces global optimal solutions. Secondly, we have reduced the solution of the unconstrained MAB problem to a set of scalar equations. For the typical sales response functions used in marketing, these equations are easy to solve and give closed formulas to calculate the optimal advertising budget and its optimal allocation.

From a practical point of view, we have seen that the multistage model, in comparison with the single stage model, allows for a better allocation of advertising along the planning horizon. We have solved a MAB instance derived from a realistic case from the advertising industry. Firstly, we have compared the full MAB model to its single stage version. We have observed that the (optimal) budget proposed by the multistage model has increased by 5.52% the profit of the (suboptimal) budget proposed by an equivalent sequence of single stage models. According to the basic sensitivity analysis we have performed, it seems that 5.52% represents fairly well the improvement that one can expect from the multistage approach applied to the type of instance we have analyzed in the case study. Secondly, we have shown by example that the constrained MAB model can be used to allocate more restrictive budgets combined with other investing constraints. For example, we have observed that, by reducing 50% the optimal budget, the advertising profit drops 6.18%.

We would like to point out some limitations of the MAB model presented. The first limitation is that the model does not takes the product price as a sales driver, that is, as a decision variable (prices are input data in the current formulation). The second limitation is that the model is deterministic. That is, it does not take into account the uncertainty of some parameters. The third limitation is that the model does not takes into account the competitive aspects of the problem. As a matter for further research, the authors are planning to improve the MAB model regarding these three aspects.

Acknowledgements

We thank the reviewers for their suggestions and comments that improved this article. Thanks are due to the support of the consulting company Bayes Forecast. We also thank the support of: grant S2009/esp-1594 from the ‘Comunidad de Madrid’ (Spain), grant MTM2009-14039-C06-03 from the Spanish government, grant No.70871081 from National Natural Science Foundation of China, and MAEC-AECID fellowship from the Spanish government.

6 Appendix A: Proofs

Lemma 1

Straightforward.

Proposition 1

This proof is also straightforward since the composition of a concave function with an affine function is concave.

Proposition 2

Analogous to the proof of Proposition 1.

Lemma 2

Straightforward.

Lemma 3

First, if $R_t(x)$ satisfies Assumption 1 with constant K_t and by the previous Lemma, we have

$$\begin{aligned}
\lim_{x\to 0,\ x>0} r_t'(x) &= \lim_{x\to 0,\ x>0} \left(p_{t1} R_t'(x) + \sum_{i\in\mathcal{I}} p_{ti} a_{ti} \right) \\
&> p_{t1} K_t + \sum_{i\in\mathcal{I}} p_{ti} a_{ti} \\
&= \widetilde{c}_t - \delta \widetilde{c}_{t+1}.
\end{aligned}$$

and

$$\begin{aligned}
\lim_{x\to +\infty} r_t'(x) &= \lim_{x\to +\infty} \left(p_{t1} R_t'(x) + \sum_{i\in\mathcal{I}} p_{ti} a_{ti} \right) \\
&< p_{t1} K_t + \sum_{i\in\mathcal{I}} p_{ti} a_{ti} \\
&= \widetilde{c}_t - \delta \widetilde{c}_{t+1}.
\end{aligned}$$

Second, since $r_t''(x) = p_{t1} R_t''(x) < 0$, we have that $r_t'(x)$ is strictly decreasing for all $x \in J$. Third, since $r_t''(x)$ exists on J, we have that $r_t'(x)$ is a continuous function. Putting together these three facts, we have that the value $\widetilde{c}_t - \delta\widetilde{c}_{t+1}$ is attained exactly once by $r_t'(x)$ on J. Therefore, there exists a unique $x_t^* > 0$ such that

$$r_t'(x_t^*) = \widetilde{c}_t - \delta\, \widetilde{c}_{t+1}.$$

Proposition 3

It is not difficult to see that

$$\sum_{t_0\le t\le T} \frac{\partial V_t(g)}{\partial g_{t_0}} = \delta^{T+1-t_0}\, c_{T+1}.$$

Let us prove the statement of this Proposition by using the previous equality. For any component t_0 of vector g we have that

$$\begin{aligned}
\frac{\partial \widehat{Q}(g)}{\partial g_{t_0}} &= \sum_{1\le t\le t_0-1} \frac{\partial Q_t(g)}{\partial g_{t_0}} + \sum_{t_0\le t\le T} \frac{\partial Q_t(g)}{\partial g_{t_0}} \\
&= 0 \quad + \sum_{t_0\le t\le T} \frac{\partial}{\partial g_{t_0}} \left(r_t(v_t + u_t^T\, g) - c_t\, g_t + V_t(g)\right) \\
&= \sum_{t_0\le t\le T} \frac{\partial}{\partial g_{t_0}}\, r_t(v_t + u_t^T\, g) - \sum_{t_0\le t\le T} \frac{\partial}{\partial g_{t_0}}\, c_t\, g_t + \sum_{t_0\le t\le T} \frac{\partial V_t(g)}{\partial g_{t_0}} \\
&= \left(\sum_{t_0\le t\le T} r_t'(v_t + u_t^T\, g)\cdot \delta^{t-t_0}\right) - \widetilde{c}_{t_0}
\end{aligned}$$

where $\widetilde{c}_{t_0} = c_{t_0} - \delta^{T+1-t_0}\, c_{T+1}$.

Proposition 4

• Statement 1: We will proceed by backward induction, i.e., we will start by studying the non-linear equation at stage T, then at stage $T-1$, etc.

1. For the last stage T, that is $t_0 = T$, we have

   $$\begin{aligned}
   \sum_{T\ge t\ge T} r_t'(x_t)\cdot \delta^{t-T} - \widetilde{c}_T &= 0 \\
   r_T'(x_T)\cdot \delta^{T-T} - \widetilde{c}_T &= 0 \\
   r_T'(x_T) &= \widetilde{c}_T - \delta \widetilde{c}_{T+1}
   \end{aligned}$$

   where we have used that $\widetilde{c}_{T+1} = 0$.

   By Lemma 3 there exists a unique scalar $x_T^* > 0$ such that $r_T'(x_T^*) = \widetilde{c}_T - \delta\widetilde{c}_{T+1}$ (in this particular case $r_T'(x_T^*) = \widetilde{c}_T$).

2. For stage $T-1$, that is $t_0 = T-1$, we have

   $$\begin{aligned}
   \sum_{T\ge t\ge T-1} r_t'(x_t)\cdot \delta^{t-(T-1)} - \widetilde{c}_{T-1} &= 0 \\
   r_T'(x_T)\cdot \delta^{T-(T-1)} + r_{T-1}'(x_{T-1})\cdot \delta^{T-1-(T-1)} - \widetilde{c}_{T-1} &= 0 \\
   r_T'(x_T^*)\cdot \delta + r_{T-1}'(x_{T-1}) - \widetilde{c}_{T-1} &= 0 \\
   r_{T-1}'(x_{T-1}) - \widetilde{c}_{T-1} + \delta\widetilde{c}_T &= 0 \\
   r_{T-1}'(x_{T-1}) &= \widetilde{c}_{T-1} - \delta\widetilde{c}_T.
   \end{aligned}$$

   By Lemma 3 there exists a unique scalar $x_{T-1}^* > 0$, such that $r_{T-1}'(x_{T-1}^*) = \widetilde{c}_{T-1} - \delta\widetilde{c}_T$.

3. By induction hypothesis, let us suppose that, for any $\tau \ge 1$ and any t such that $T > t > T-\tau$, there exists a unique x_t^* such that

   $$r_t'(x_t^*) = \widetilde{c}_t - \delta\widetilde{c}_{t+1}.$$

Let us prove that the statement is also true for $t_0 = T - \tau$. For stage $T - \tau$, that is $t_0 = T - \tau$, we have

$$\begin{aligned}
\sum_{T \geq t \geq T-\tau} r'_t(x_t) \cdot \delta^{t-(T-\tau)} - \widetilde{c}_{T-\tau} &= 0 \\
\sum_{T \geq t > T-\tau} r'_t(x^*_t) \cdot \delta^{t-(T-\tau)} + r'_{T-\tau}(x_{T-\tau}) - \widetilde{c}_{T-\tau} &= 0 \\
\widetilde{c}_T \cdot \delta^{\tau} + (\widetilde{c}_{T-1} - \delta\widetilde{c}_T) \cdot \delta^{\tau-1} + \ldots + (\widetilde{c}_{T-\tau+1} - \delta\widetilde{c}_{T-\tau+2}) \cdot \delta + r'_{T-\tau}(x_{T-\tau}) - \widetilde{c}_{T-\tau} &= 0 \\
\widetilde{c}_{T-\tau+1}\delta + r'_{T-\tau}(x_{T-\tau}) - \widetilde{c}_{T-\tau} &= 0,
\end{aligned}$$

which implies

$$r'_{T-\tau}(x_{T-\tau}) = \widetilde{c}_{T-\tau} - \delta\widetilde{c}_{T-\tau+1}.$$

By Lemma 3 there exists a unique scalar $x^*_{T-\tau} > 0$, such that $r'_{T-\tau}(x^*_{T-\tau}) = \widetilde{c}_{T-\tau} - \widetilde{c}_{T-\tau+1}\delta$.

• Statement 2: The system of linear equations (13) can be stated as $Bg = d$ where

$$B = \begin{pmatrix} 1 & & & & \\ \delta & 1 & & & \\ \delta^2 & \delta & 1 & & \\ \vdots & \vdots & & \ddots & \\ \delta^{T-1} & \delta^{T-2} & \ldots & \delta & 1 \end{pmatrix} \tag{14}$$

$$d^T = (x^*_1 - v_1, \ldots, x^*_T - v_T). \tag{15}$$

Given that $\det B \neq 0$, the linear system has a unique solution, say g^*.

Theorem 1

• Statement 1: By definition

$$\nabla \widehat{Q}(g) = \left(\frac{\partial \widehat{Q}(g)}{\partial g_t} \right)_{t \in \mathcal{T}}.$$

By Proposition 3, equations $\frac{\partial \widehat{Q}(g)}{\partial g_{t_0}} = 0$ with $t_0 \in \mathcal{T}$, can be written as:

$$\sum_{t_0 \leq t \leq T} r'_t(x_t) \cdot \delta^{t-t_0} - \widetilde{c}_{t_0} = 0, \qquad t_0 \in \mathcal{T}, \tag{16}$$

where

$$\begin{aligned}
x_t &= u_t^T\, g_t + v_t, && t \in \mathcal{T}, \\
u_t^T &= (\delta^{t-1},\ \delta^{t-2}, \ldots,\ \delta,\ 1,\ 0, \ldots,\ 0) \in \mathbb{R}^T, && t \in \mathcal{T}, \\
v_t &= \delta^t \widetilde{x}_0, && t \in \mathcal{T}.
\end{aligned}$$

By Proposition 4, there exists a unique positive vector $(x^*_1, \ldots, x^*_T)^T$ solution of the system of non-linear equations (16). Also by Proposition 4, there exists a unique vector $g^* \in \mathbb{R}^T$ that solves the following system of T linear equations

$$u_t^T g + v_t = x^*_t, \qquad t \in \mathcal{T}.$$

Therefore, we have proved that g^* is the unique solution of

$$\frac{\partial \widehat{Q}(g)}{\partial g_t} = 0, \qquad t \in \mathcal{T}.$$

i.e., there exists a unique vector g^* such that

$$\nabla \widehat{Q}(g^*) = 0.$$

Since $\widehat{Q}(g)$ is concave, it follows that g^* is the unique global maximizer of $\widehat{Q}(g)$.

• Statement 2: Part (a) directly follows from the proof of Proposition 4.

Part (b) can be proved by taking into account the triangular structure of B in the system of linear equations

$$Bg = d,$$

where B and d are defined by equations (14) and (15), respectively. In this case, we can write

$$\begin{aligned} g_1^* &= x_1^* - \delta\, \widetilde{x}_0, \\ g_t^* &= x_t^* - \delta^t\, \widetilde{x}_0 - \sum_{\tau=1}^{t-1} \delta^{t-\tau}\, g_\tau^*, \qquad\qquad t \geq 2 \end{aligned}$$

These expressions that can be written in a recursive way:

$$g_t^* = x_t^* - \delta\, x_{t-1}^*, \qquad t \in \mathcal{T},$$

where $x_0^* = \widetilde{x}_0$.

• Statement 3: Trivial.

Lemma 4

For all $t \in \mathcal{T}$ the following statements hold:

1. $\alpha_t(1 - e^{-\beta_t 0}) = 0$ and $\alpha_t(1 - e^{-\beta_t x}) > 0$ for all $x \in J =]0, +\infty[$.
2. Since $R_t'(x) = \alpha_t \beta_t e^{-\beta_t x}$, we have that $\lim_{x \to 0,\, x>0} R_t'(x) = \alpha_t \beta_t > K_t$ for all $t \in \mathcal{T}$. Therefore $\lim_{x \to 0,\, x>0} R_t'(x) > K_t$.
3. On the one hand, considering that by hypothesis $\widetilde{c}_t - \delta\, \widetilde{c}_{t+1} > 0$ and $a_{ti} < 0$, we have that

   $$K_t = \frac{1}{p_{t1}} \left(\widetilde{c}_t - \delta \widetilde{c}_{t+1} - \sum_{i \in \mathcal{I}} p_{ti} a_{ti} \right) > 0.$$

   On the other hand, we have that $\lim_{x \to +\infty} R_t'(x) = 0$. Therefore $\lim_{x \to +\infty} R_t'(x) < K_t$.
4. $R_t''(x) = -\alpha_t \beta_t^2 e^{-\beta_t x}$ exists and $R_t''(x) < 0$ for all $x > 0$ (therefore $R_t(x)$ is strictly concave).

Proposition 5

According to Theorem 1, the optimal allocation g^* can be computed as follows: For each $t \in \mathcal{T}$, calculate the scalar x_t^* by solving the scalar equation

$$r_t'(x_t) = \widetilde{c}_t - \delta\, \widetilde{c}_{t+1},$$

which amounts to

$$p_{t1} R_t'(x_t) + \sum_{i \in \mathcal{I}} p_{ti} a_{ti} = \widetilde{c}_t - \delta\, \widetilde{c}_{t+1},$$
$$p_{t1} \alpha_t \beta_t e^{-\beta_t x_t} + \sum_{i \in \mathcal{I}} p_{ti} a_{ti} = \widetilde{c}_t - \delta\, \widetilde{c}_{t+1},$$

from where

$$x_t^* = \frac{1}{\beta_t} \ln \left(\frac{p_{t1} \alpha_t \beta_t}{\widetilde{c}_t - \delta\, \widetilde{c}_{t+1} - \sum_{i \in \mathcal{I}} p_{ti} a_{ti}} \right), \qquad t \leq T.$$

The second statement was proved in Theorem 1.

Proposition 6

Straightforward.

7 Appendix B: Data for the case study

Table 4: Values for δ_{ik} .

$i \setminus k$	1	2
1	0.660	0.552
2	0.588	0.552

Table 5: Values of $\widetilde{x}_{0ik}$ (in GRPs).

$i \setminus k$	1	2
1	300	300
2	50	50

Table 6: Values of α_{tik} (in units of product i).

t	α_{t11}	α_{t12}	α_{t21}	α_{t22}
1	345,000	270,000	86,400	105,600
2	389,850	305,100	83,700	102,300
3	493,350	386,100	113,400	138,600
4	510,600	399,600	137,700	168,300
5	731,400	572,400	199,800	244,200
6	838,350	656,100	251,100	306,900
7	897,000	702,000	278,100	339,900
8	969,450	758,700	259,200	316,800
9	734,850	575,100	224,100	273,900
10	386,400	302,400	191,700	234,300
11	427,800	334,800	189,000	231,000
12	407,100	318,600	218,700	267,300

Table 7: Values of β_{tik} do not depend on t (in GRPs^{-1}).

$i \setminus k$	1	2
1	0.010	0.010
2	0.005	0.005

Table 8: Values of a_{tij1} (in units of product i per unit of driver $j1$).

$i \setminus j$	1	2
1	0	-0.00010
2	-0.00010	0

Table 9: Values of a_{tij2} (in units of product i per unit of driver $j2$).

$i \setminus j$	1	2
1	0	-0.00015
2	-0.00015	0

Table 10: Values of c_{tik} (in euros per GRP).

t	c_{t11}	c_{t12}	c_{t21}	c_{t22}
1	480	528	432	475
2	480	528	432	475
3	640	704	576	634
4	640	704	576	634
5	480	528	432	475
6	480	528	432	475
7	640	704	576	634
8	640	704	576	634
9	480	528	432	475
10	480	528	432	475
11	640	704	576	634
12	640	704	576	634

Table 11: Other parameters.

Parameter	Value	Units
T	12	month
I	2	produt
K_1	2	driver
K_2	2	driver
p_{t1}	1.75	euro/unit of product 1
p_{t2}	1.40	euro/unit of product 2

References

[1] A. Ahmadi-Javid and P. Hoseinpour. On a cooperative advertising model for a supply chain with one manufacturer and one retailer. *European Journal of Operational Research*, 219(2):458–466, 2012.

[2] F. M. Bass, N. Bruce, S. Majumdar, and B. P. S. Murthi. Wearout effects of different advertising themes: A dynamic bayesian model of the advertising-sales relationship. *Marketing Science*, 26(2):179–195, 2007.

[3] U. K. Bhattacharya. A chance constraints goal programming model for the advertising planning problem. *European Journal of Operational Research*, 192:382–395, 2009.

[4] S. Broadbent. One way TV advertisements work. *Journal of the Market Research Society*, 21(3):139–165, 1979.

[5] A. Buratto and L. Grosset. A communication mix for an event planning: A linear quadratic approach. *Central European Journal of Operations Research*, 14(3):247–259, 2006.

[6] A. Buratto, L. Grosset, and B. Viscolani. Advertising a new product in a segmented market. *European Journal of Operational Research*, 175(2):1262 – 1267, 2006.

[7] A. Buratto, L. Grosset, and B. Viscolani. Advertising channel selection in a segmented market. *Automatica*, 42(8):1343 – 1347, 2006.

[8] A. Buratto, L. Grosset, and B. Viscolani. Advertising coordination games of a manufacturer and a retailer while introducing a new product. *Top*, 15:307–321, 2007.

[9] P. J. Danaher and R. T. Rust. Determining the optimal return on investment for an advertising campaign. *European Journal of Operational Research*, 95:511–521, 1996.

[10] P. Doyle and J. Saunders. Multiproduct advertising budgeting. *Marketing Science*, 9(20):97–113, 1990.

[11] R. Du, A. Hu, and S. Ai. Stochastic optimal budget decision for advertising considering uncertain sales responses. *European Journal of Operational Research*, 183:1042–1054, 2007.

[12] F. M. Feinberg. On continuous-time optimal advertising under s-shaped response. *Management Science*, (11):1476–1487, 2001.

[13] M. Garrat, R. Alcaraz, and S. Cohen. From brilliant to actionable: It takes technical brilliance and constant questioning to achieve the truly actionable in marketing ROI. In *Esomar - Congress*, 2011.

[14] P. E. Gill, W. Murray, and M. H. Wright. *Practical optimization*. Academic Press Limited, London, 10th edition, 1995.

[15] L. Grosset and B. Viscolani. Advertising for a new product introduction: A stochastic approach. *Top*, 12(1):149–167, 2004.

[16] D. M. Hanssens, L. J. Parsons, and R. L Schultz. *Market response models: Econometric and time series analysis*. Kluwer academic publishers, second edition, 2001.

[17] J. D. C. Little. Aggregate advertising models: the state of the art. *Operations Research*, 27:629–667, 1979.

[18] J. D. C. Little. Models and managers: The concept of a decision calculus. *Management Science*, 50(12):1841–1853, 2004.

[19] J. D. C. Little and L. M. Lodish. A media planning calculus. *Operations Research*, 17(1):1–35, 1969.

[20] P. Mariel and J. Sandonis. A model of advertising with application to the german automobile industry. *Applied Economics*, 36:83–92, 2004.

[21] H. I. Mesak and J. A. Calloway. A pulsing model of advertising competition: A game theoretic approach, part a - theoretical foundation. *European Journal of Operational Research*, 86:231–248, 1995.

[22] P. A. Naik, M. K. Mantrala, and A. G. Sawyer. Planning media schedules in the presence of dynamic advertising quality. *Marketing Science*, 17(3):214–235, 1998.

[23] P. A. Naik, K. Raman, , and R. S. Winer. Planning marketing-mix strategies in the presence of interaction effects. *Marketing Science*, 24(1):25–34, 2005.

[24] M. Nerlove and K. J: Arrow. Optimal advertising policy under dynamic conditions. *Economica*, 29(114):129–142, 1962.

[25] M. Powell. A fast algorithm for nonlinearly constrained optimization calculations. In G. Watson, editor, *Numerical Analysis*, volume 630 of *Lecture Notes in Mathematics*, pages 144–157. Springer Berlin / Heidelberg, 1978.

[26] S. Sriram and M. U. Kalwani. Optimal advertising and promotion budgets in dynamic markets with brand equity as a mediating variable. *Management Science*, 53(1):46–60, 2007.

[27] R. K. Sundaram. *A first course in optimization theory*. Cambridge University Press, 1996.

[28] S. Swami and A. Dutta. Advertising strategies for new product diffusion in emerging markets: Propositions and analysis. *European Journal of Operational Research*, 204:648–661, 2010.

[29] B. Wierenga. Modelling impact of advertising and optimizing advertising policy: an application in recreation. *European Journal of Operational Research*, 8(3):235–248, 1981.

[30] H. A. Zielske and W. A. Henry. Remembering and forgetting television ads. *Journal of Advertising Research*, 20:7–13, 1980.