



Resource allocation procedures for unknown sales response functions with additive disturbances

Daniel Gahler¹ · Harald Hruschka¹ 

Accepted: 7 December 2021
© The Author(s) 2021

Abstract

We develop a modified exploration–exploitation algorithm which allocates a fixed resource (e.g., a fixed budget) to several units with the objective to attain maximum sales. This algorithm does not require knowledge of the form and the parameters of sales response functions and is able to cope with additive random disturbances. Note that additive random disturbances, as a rule, are a component of sales response functions estimated by econometric methods. We compare the developed algorithm to three rules of thumb which in practice are often used to solve this allocation problem. The comparison is based on a Monte Carlo simulation for 384 experimental constellations, which are obtained from four function types, four procedures (including our algorithm), similar/varied elasticities, similar/varied saturations, high/low budgets, and three disturbance levels. A statistical analysis of the simulation results shows that across a multi-period planning horizon the algorithm performs better than the rules of thumb considered with respect to two sales-related criteria.

Keywords Marketing resource allocation · Exploration–exploitation algorithm · Monte Carlo simulation · Optimization

JEL Classification M30 · C61 · C63

1 Introduction

Allocation decisions in marketing refer to decision variables like advertising budgets, sales budgets, sales force sizes, and sales calls that are allocated to sales units like sales districts, customer groups, individual customers, and prospects. Studies using optimization methods and empirical sales response functions provide evidence

✉ Harald Hruschka
harald.hruschka@wiwi.uni-regensburg.de

Daniel Gahler
daniel.gahler@wiwi.uni-regensburg.de

¹ University of Regensburg, Regensburg, Germany

to the importance of such allocation decisions. These studies demonstrate that sales or profits can be increased by changing allocation of budgets, sales force or sales calls (Beswick and Cravens 1977; LaForge and Cravens 1985; Sinha and Zoltners 2001). The average increase of profit contribution across studies analyzed in the review of Sinha and Zoltners (2001) compared to the current policies was 4.5% of which 71% are due to different allocations and 29% are due to size changes. The smaller second percentage can be explained by the well known flat maximum principle (Mantrala et al. 1992). According to this principle, deviations of up to $\pm 25\%$ from the optimal value of a marketing decision variable affect profit only slightly. To our knowledge, Tull et al. (1986) were the first who demonstrate the flat maximum principle for different elasticities, contribution margins and sales response functions.

These studies all require knowledge of the mathematical form of sales response functions that reproduce the dependence of sales on decision variables. In addition, they require that parameter values of sales response functions are available, e.g., determined by econometric methods using historical data or by means of a decision calculus approach that draws upon managers' experiences (Gupta and Steenburgh 2008). Of course, there are situations in which both econometric methods and decision calculus cannot be applied. Lack of historical data (e.g., for new sales units), lack of variation of past allocations, lack of experiences with the investigated or similar markets constitute possible causes.

In such difficult situations the question arises how management may arrive at rational allocation decisions nonetheless. To our knowledge, Albers (1997) developed the only relevant approach. The German version of this paper Albers (1998) has the same content but was published in a German journal, rather than being a conference paper. Albers demonstrates that, in spite of the lack of knowledge on functional form and parameters, the allocation problem for one resource may be solved by a heuristic that he obtains from the optimality conditions (see expression (6) in Sect. 2) by inserting previous period's elasticity estimates and sales.

Albers (1997) applies this heuristic in several iterations (= periods) in each of which he updates elasticity estimates based on allocations and sales of the previous period. Being combined with this updating mechanism, the heuristic works without knowledge of functional forms and parameters of sales response functions. Therefore this heuristic seems to be appropriate in situations in which neither econometric methods nor decision calculus approaches can be used.

In order to compare the performance of the heuristic to several rule of thumbs we generate sales by deterministic simulation, i.e., from response functions that are not subject to random disturbances. Albers (1997) investigates several sales response functions with different parameter values. Note that these functions and their parameters are not used by the heuristic. They only serve to generate sales by deterministic simulation.

To our knowledge the heuristic of Albers (1997) has been applied in two related publications (Fischer et al. 2011, 2013). In contrast to Albers (1997) in both publications the elasticity estimates considered by the heuristic are not updated, but computed from parameters of sales response functions whose functional forms are assumed to be known. In these publications the heuristic serves as iterative method that takes the place of a numerical optimization algorithm.

On the other hand, sales are generated by deterministic simulation in each period and generated sales of the previous period are input into the heuristic just like in Albers (1997).

In Fischer et al. (2011) sales response functions are assumed multiplicative. The heuristic provides solutions close to those obtained by numerical optimization. Fischer et al. (2013) assume that one of two functional forms (multiplicative, modified exponential) is true. These authors compute elasticities based on the assumed form of the response model alternatively from true and erroneous parameters. Sales are generated by deterministic simulation using error-free parameters. In case of erroneous parameters the heuristic performs better than numerical optimization because the former includes sales of previous periods generated by deterministic simulation with error-free parameters, whereas the latter computes sales from erroneous parameters.

Table 1 contains the main properties of the heuristics and simulation techniques applied in previous relevant publications. It also characterizes the iterative algorithm that we introduce here and the type of simulation that we use. We note that the simulations presented in Albers (1997) and Fischer et al. (2011, 2013) generate deterministic sales values, as their sales response models do not include an overall disturbance term. Of course, ignoring disturbances goes against econometrically estimated sales response models that always comprise such a component [for an excellent overview see Hanssens et al. (2001)]. To remove this limitation we perform stochastic simulation based on response models that are subject to additive random disturbances.

We pursue the following research goals in this study. Firstly, we investigate the performance of the heuristic of Albers (1997) with included updating mechanism to solve the allocation problem, if sales are not deterministic but are generated from sales response functions with additive disturbance terms. Secondly, we develop an iterative algorithm for non-deterministic sales. Like the heuristic in Albers (1997) this algorithm has no knowledge about functional forms and parameter values. In addition, the algorithm has no information about the size of the random disturbances. In the first stage of the algorithm, exploration, we extend the heuristic of Albers by smoothing and projecting elasticity estimates into an interval which is based on meta-analytic results. Finally, similar to Albers (1997) we compare the algorithm to three rules of thumb and use several functional forms (i.e., multiplicative, modified exponential, ADBUDG) to generate sales by (in our case: stochastic) simulation.

Table 1 Comparison of relevant publications

Publication	Updating of elasticities	Functional forms and parameters	Simulation to generate sales
Albers (1997)	Yes	Unknown	Deterministic
Fischer et al. (2011)	No	Known	Deterministic
Fischer et al. (2013)	No	Known	Deterministic
This paper	Yes	Unknown	Stochastic

The three rules of thumb allocate a given resource (e.g., a budget) to sales units proportional to a metric. Rule of thumb 1 is based on sales of the previous period, Rule of thumb 2 on sales divided by the allocation of the previous period. Maximum sales observed so far serves as metric in rule of thumb 3 (please also see Table 2). Rules of thumbs 1 and 2 correspond to rules in Albers (1997). Rule of thumb 3 on the other hand replaces the original third rule in Albers (1997), allocation proportional to the saturation level, because saturation levels like all parameters of functions are unknown.

Rule of thumb 1 turns out to be stable, but leads to lower average sales than the developed algorithm. Rule of thumb 2 leads to sales sometimes close to the optimum. However, in several constellations this rule suffers from unusable results. On average sales are lower than for our developed algorithm. Rule of thumb 3 works well within a certain range, but does not scale well with either low or high resources. In other words, there is a higher overall risk that rule of thumb 3 fails under such circumstances.

In Sect. 6 we show that the investigated rules of thumb, which practitioners frequently use, have several shortcomings, which can be overcome by switching to a different procedure. In particular, this is a method that practitioners can not only easily implement, but also modify according to their own particular allocation situation. We also explain how some of the rules of thumb can lead to unusable results (please see Table 2).

The developed alternative algorithm consists of two stages, exploration and exploitation. March (1991) seems to be the first author combining the concepts of exploration and exploitation. The idea is to divide the planning horizon into two stages. In the first stage, data are generated to explore the shape of the response functions. In the second stage, the knowledge of their shape is exploited to find optimal values. In each of several iterations, we approximate the unknown functions by quadratic polynomials and obtain a solution of the allocation problem by quadratic programming.

To understand the problem setting it may be beneficial to keep the following two examples at the back of one's head while reading. A manufacturer sells and advertises in eight different countries. At the beginning of each period (say month or week), the total advertising budget is allocated onto the eight countries, and at the end of the period, the total sales figures from each country are reported. As another example think about the allocation of a given number of total sales calls (= visits by sales representatives) to several sales districts. Decision makers want to optimize sales/profit in both examples. Customers in each country or sales district might react differently, and few data (if any) are known from previous periods.

The managerial implications are twofold: upon entering a new market with no or very little prior information, the developed algorithm leads to stable results better than those provided by any rule of thumb. Furthermore, in any situation it yields better results independent of the circumstances of the market, *in particular* the size of random disturbances.

Section 2 describes the problem from a mathematical point of view. In Sect. 3 we present the necessary preparations for the simulation study. In Sect. 4 we describe the algorithm. Section 5 explains the investigated rules of thumb. We expect that the

Table 2 Comparison of procedures

Procedure	Advantages	Issues
Rule of thumb 1: proportional to sales	Stable results	Lower on average
Rule of thumb 2: proportional sales divided by allocation	May achieve values close to the optimum	Leads to unusable results in some constellations, lower on average
Rule of thumb 3: proportional to maximal sales observed so far	Works well within a certain budget range	Does not scale well with low and high resources
Heuristic of Albers (1997) with elasticity updating	Usually converges for deterministic response functions	Cannot handle stochastic response functions
Smoothing and projection added to the heuristic of Albers (1997) with elasticity updating (exploration stage only)	Useful allocation results for stochastic response functions	Inferior to rules of thumb
Developed exploration–exploitation algorithm	Performs better than the rules of thumb even for stochastic response functions	Possibly first such method, room for improvement

developed algorithm is superior to these rules of thumb. In Sect. 6 we present and discuss results of the simulation study. We investigate the performance of the algorithm under conditions different from those in the simulation study in Sect. 7. We also mention several extensions of the allocation problem, which may be solved by modifications of our algorithm.

We present the algorithm as pseudocode in Appendix A. Further evaluation results not given in the main text can be found in Appendix B. We add supplementary notes on the generation of sales response functions in Appendix C. In Appendix D we explain why higher budgets usually lead to better performance measures for both the algorithm and the rules of thumb.

2 Decision problem

A (scarce) resource B needs to be allocated to $n \in \mathbb{N}_{>1}$ units. We have one sales response function f_i for each unit $i = 1, \dots, n$ that only depends on its allocated input x_i . The sales response function maps the allocated budget to the obtained sales of the same period. Inputs must be non-negative and lower than the resource, i.e., $0 \leq x_i \leq B$.

In addition, the sum of all inputs must not exceed the resource:

$$\sum_{i=1}^n x_i \leq B \quad (1)$$

A n -tuple (x_1, \dots, x_n) satisfying (1) is called an *allocation*. Total sales, i.e., the sum of sales across all units $\sum_{i=1}^n f_i(x_i)$, represent the *objective* of this allocation. The goal is to find an allocation maximizing total sales:

$$\max \sum_{i=1}^n f_i(x_i) \quad (2)$$

We use sales as objective instead of profits for several reasons. Firstly, for the sake of comparability with Albers (1997), who also considers sales. Secondly, if the resource is fixed, sales and profits should be highly correlated, as the only difference is the profit contribution for each unit. Thirdly, it makes calculations and interpretations easier. We doubt that changing the objective to profit will yield more advantages than problems.

As we assume that all sales response functions are monotonically increasing (thereby excluding effects such as supersaturation, which we briefly discuss in Sect. 7), we conclude that condition (1) is binding and can therefore be rewritten as:

$$\sum_{i=1}^n x_i = B \quad (3)$$

Example 1 We now introduce an example, which we extend throughout this article whenever we present a new concept. Note how an allocation of a budget B on two

units still has one binding, solvable restriction (3) and is therefore a one-dimensional optimization problem. Our example therefore has three units, to make sure the problem is not solvable with elementary one-dimensional methods. Furthermore we round the numbers quite abruptly, indicated by ‘ \approx ’, to make the examples more easily readable.

We assume our company runs advertisements in three different countries, and has a total advertising budget of $B = 6$ that needs to be allocated. If we know nothing about the market, we might choose equal allocations ($x_1 = 2, x_2 = 2, x_3 = 2$) that satisfy (3).

For each of the three countries we have an advertising response function f_i that transforms the input x_i to the sales revenue of the country at the end of that month. In our example, these functions will be

$$f_1(x) = 5\sqrt[3]{x}; \quad f_2(x) = 3\sqrt[8]{x}; \quad f_3(x) = 3\sqrt[8]{x}.$$

We know these functions are monotonically increasing, and therefore the condition (1) is binding and becomes (3). The objective value in this case is

$$\sum_{i=1}^n f_i(x_i) = f_1(2) + f_2(2) + f_3(2) = 5\sqrt[3]{2} + 3\sqrt[8]{2} + 3\sqrt[8]{2} \tag{4}$$

$$\approx 6.3 + 3.3 + 3.3 = 12.9. \tag{5}$$

The decision problem presented so far makes it possible that no (in other words zero) resources are allocated to a unit (e.g., to make no sales calls in certain districts). Of course, decision makers may find it inappropriate to deprive a unit of all resources. To cope with such a situation, one defines a new problem in which a modified total resource $B' := B - \sum_{i=1}^n lb_i$ is allocated to the units with functions $g_i(x_i) := f_i(lb_i + x_i)$ where g_i is f_i shifted by the lower bound. Now the inputs of some (or all) units have a lower bound $lb_i \geq 0$, i.e., the minimal amount of resource allocated to unit i and B' must be zero or positive. One can henceforth allocate onto the functions g_i , however to simplify the notation we will still refer to the functions as f_i . In other words, without loss of generality, we may assume the lower bounds of f_i and hence the lowest possible value of x_i to be 0.

Example 2 Assume now, that the budget is actually $B = 9$, and we have to allocate at least $lb_i = 1$ to every unit. All we have to do is allocate the remaining budget $B' = B - \sum_{i=1}^3 lb_i = 9 - (1 + 1 + 1) = 6$ onto the shifted functions $g_i(x_i) := f_i(lb_i + x_i) = f_1(1 + x_1)$

Albers (1997) derives the following optimality conditions for this decision problem from a Lagrangian:

$$x_i = \frac{f_i(x_i)\epsilon_i}{\sum_j f_j(x_j)\epsilon_j} B \quad \text{for } i = 1, \dots, n \tag{6}$$

ff

ε_i denotes the point elasticity of allocation x_i defined as $\frac{\partial f_i}{\partial x_i} \cdot \frac{x_i}{f_i}$. Equation (6) corresponds to a proportional rule, which makes sure that the x_i do add up to the total resource B . One can also see that the allocation to a sales unit increases with its sales $f_i(x_i)$ and its elasticity ε_i . However, please be aware that these optimality conditions define a fixed-point problem, because each x_i is mapped to itself by the expression on the right hand side of Eq. (6). Note that we get a fixed-point problem even if each elasticity value is constant, which would be true for multiplicative functions with known parameter values.

Example 3 The functions in our example are *multiplicative* and hence have constant elasticity, in this case $\varepsilon_1 = 1/3$, $\varepsilon_2 = \varepsilon_3 = 1/8$. Taking the sales values f_i from the allocation (2,2,2) from Example 1, the denominator of (6) becomes

$$6.3 \cdot 1/3 + 3.3 \cdot 1/8 + 3.3 \cdot 1/8 \approx 2.9,$$

which leads us to the next allocation of

$$x_1 = \frac{6.3 \cdot 1/3}{2.9} B = 0.72 \cdot 6 \approx 4.3, \quad x_2 = x_3 \approx 0.85.$$

Indeed, we see that the new objective value is

$$f_1(4.3) + f_2(0.85) + f_3(0.85) \approx 8.1 + 2.9 + 2.9 = 13.9 > 12.9$$

3 Simulation

To evaluate the allocation procedures, we perform a Monte Carlo simulation. We choose experimental conditions as close as possible to the paper presented by Albers (1997).

The successive steps of the simulation can be characterized as follows:

- Select one of four functional forms,
- Choose one of two budget levels,
- Determine for each sales unit parameters, which depend on (non-) similarity of elasticities and saturation levels,
- select one of four procedures to determine the allocations to sales units,
- Compute deterministic sales given functional form, parameters and allocations,
- Select one of three disturbance levels and add appropriate disturbances to obtain total stochastic sales for each unit.

The simulation comprises 384 constellations, which result from four function types, four allocation procedures, three disturbance levels, and two budget, elasticity and saturation levels. The simulation is replicated twenty times for each constellation. The numerical values for these levels will be given in Sect. 3.2.

Within the simulation, we make the following assumptions:

- Response functions are monotonically increasing, and mostly concave, the only exception being the S-shaped ADBUDG function, which is convex beyond its inflection point.
- A budget is constant across all periods.
- Response units are mutually independent and the objective function is separable, $\frac{\partial f_i}{\partial x_j} = 0 \quad \forall i \neq j$.
- There are no lag-effects, i.e., the response value does not depend on the value of the previous periods
- Error terms added to the response functions are normally distributed with mean zero.

Note that the last assumption is also usual in the literature on nonlinear regression models (see Bates and Watts (1988), Seber and Wild (1989), Cook and Weisberg (1999)).

In the next section we explain which functional forms of sales response functions serve to compute sales, how parameters for sales units are determined based on properties of the functions (with respect to the (non-)similarity of elasticities and saturation levels), and how disturbances are generated.

3.1 Sales response functions

We consider the same four functional forms investigated in Albers (1997), i.e., the multiplicative, the modified exponential, the concave and the S-shaped ADBUDG functions. The deterministic parts of these functions can be written as follows:

$$f_{mul}(x) = ax^b \quad (7)$$

$$f_{exp}(x) = M_{exp}(1 - \exp(-xh)) \quad (8)$$

$$f_{adc}(x) = M_{adc} \frac{x^{\phi_c}}{G_c + x^{\phi_c}} \quad (9)$$

$$f_{ads}(x) = M_{ads} \frac{x^{\phi_s}}{G_s + x^{\phi_s}} \quad (10)$$

where $a, b, M_{exp}, M_{adc}, M_{ads}, h, G_c, G_s, \phi_c, \phi_s$ all are positive parameters.

We now mention properties of these four functional forms (for more details see, e.g., Hanssens et al. (2001)). Each parameter $M_{exp}, M_{adc}, M_{ads}$ symbolizes a saturation level, i.e. maximum sales. Given certain parameter restrictions, the first three functions allow for a concave shape, i.e., they reproduce positive marginal effects that decrease with higher values of x . These restrictions are $0 < b < 1$ for the multiplicative function, $h > 0$ for the modified exponential function and $0 < \phi_c < 1$ for

the first version of the ADBUDG function. The second version of the ADBUDG function f_{adS} leads to an S-shape for $\phi_S > 1$. S-shaped functions consist of two sections. The first section is characterized by increasing positive marginal effects, the second one by decreasing positive marginal effects.

3.2 Determining parameters of sales units

For the simulation we have to assign values to the parameters of the functions described in Sect. 3.1. For each functional form we construct eight sets of parameters (one for each sales unit) which have to fulfill two properties. Firstly, the point elasticity at equal allocations for each unit should have a predetermined value. Secondly, the function's saturation level should take a predetermined value.

To investigate the effects of elasticities and saturation levels, we distinguish two sets of values, which Table 3 presents under the column headings of similar and varied elasticities (saturation levels), respectively. In the two 'similar' columns, values are within a close range, i.e., an elasticity around 0.3 and a saturation level around 6,500,000. In the 'varied' columns, four sales units have a low elasticity (saturation), and four sales units a high elasticity (saturation).

We construct Table 3 from Table 1 in Albers (1997) which is shown in Appendix C.1. In this Appendix we also explain the differences between Table 3 and Table 1 in Albers (1997).

Let us consider the experimental condition 'low budget, multiplicative form, varied elasticity, similar saturation'. Here we determine the parameters of eight multiplicative functions such that given a low allocation budget of 1,000,000, the point elasticity for the equal allocation is one of the eight values in column 3 of Table 3 and the saturation is one of the values in column 4 of Table 3. For a detailed calculation, see Example 4 below.

For multiplicative functions, we can simply set b equal to the elasticity value. The other parameter a equals $M \cdot B^{-b}$, where M is the saturation level.

For modified exponential functions we perform an iterative search to determine the values of parameter h . The elasticity of the modified exponential function is

Table 3 Function properties for parameter generation

Sales Unit	Similar Elasticities	Varied Elasticities	Similar Saturation levels	Varied Saturation levels
1	0.26	0.11	6,100,000	4,500,000
2	0.27	0.12	6,200,000	10,000,000
3	0.28	0.13	6,300,000	4,500,000
4	0.29	0.14	6,400,000	10,000,000
5	0.31	0.47	6,600,000	4,500,000
6	0.32	0.48	6,700,000	10,000,000
7	0.33	0.49	6,800,000	4,500,000
8	0.34	0.50	6,900,000	10,000,000

$$\varepsilon = \frac{\exp(-hx)hx}{1 - \exp(-hx)} \quad (11)$$

This expression can be transformed to a fixed-point formula

$$h = \frac{-\log\left(\frac{\varepsilon(1-\exp(-hx))}{hx}\right)}{x}. \quad (12)$$

Using a starting value $h_0 > 0$ and equal allocations $x = B/8$, fixed-point iterations quickly converge to a positive value for h that satisfies the conditions in Table 3.

The value of the other parameter M of the modified exponential function is directly taken from Table 3. For the ADBUDG-functions, which have an additional parameter, we choose special values that yield reasonable shapes and ensure stability of the iteration process. The values are obtained by trial and error and controlled visually (see Appendix C.2 for a discussion and explanation why this is necessary).

Example 4 In our example so far we used three multiplicative functions, whose parameters were predetermined. In our simulation study, we want the functions *properties* to be fixed, as opposed to the parameters.

Starting from the desired properties of $\varepsilon = 1/3$ and $f(6) \approx 9$, we set the exponent b to be the elasticity $b = \varepsilon = 1/3$ and $a = M \cdot B^{-b} = 9 \cdot 6^{-1/3} = 5$. This way we obtain the parameters of the first function of Example 1.

3.3 Disturbances

We add normally distributed disturbances $u \sim \mathcal{N}(0, \sigma^2)$ to the deterministic part of each function. Variances σ^2 are set to attain a desired share of explained variance (i.e., R^2 value) for the dependent variable sales. We consider additive normally distributed disturbances, as these are included in the majority of nonlinear regression models. Switching to, say, multiplicative error terms might give the algorithm we propose an unfair advantage as it is based on elasticities. The desired R^2 values amount to 0.9, 0.7, and 0.5. For each function, the error variance σ^2 is set to the value which leads the R^2 -value closest to its desired value in a regression of that function using 2000 values from a discrete uniform distribution of integers as inputs. We choose these values for R^2 as they are easily understood and well known by both practitioners and researchers.

To avoid negative outputs ($f(x) + u$ can be negative when $f(x)$ is too small), the output was defined as $\max(f(x) + u, 0)$. This is a mixed distribution where the discrete value 0 can have positive probability. As this only happens in extreme cases, the distribution of sales is very close to a normal distribution.

Example 5 Extending the situation from Example 3 we add the following standard normally distributed random numbers ($u_1 = -0.6, u_2 = 0.2, u_3 = -0.8$) to the sales function, to obtain

$$f_1(4.3) + u_1 + f_2(0.85) + u_2 + f_3(0.85) + u_3 = 8.1 - 0.6 + 2.9 + 0.2 + 2.9 - 0.8 = 12.7$$

In particular, we see the effect the disturbance has on the objective value, as it is now lower than its value for equal allocations, even though allocation (4.3, 0.85, 0.85) is better than (2, 2, 2) in the deterministic case.

Let us illustrate the process of obtaining the correct variances. We now sample 2000 points between zero and six on the function $f_1(x) = 5\sqrt[3]{x} + u$. We take the data of these 2000 pairs $(x, f(x))$ and perform a non-linear least squares regression $f(x) \sim ax^b$. This regression gives a value for the coefficient of determination $R^2 \approx 0.67$. If we want to find the variance such that $R^2 = 0.5$ we need to increase the variance of the distribution we are sampling from. If we want to find the variance such that $R^2 = 0.7$ we need to decrease it.

4 Developed algorithm

Albers (1997) intends to show that in allocation problems an iterative heuristic along elasticities always outperforms rules of thumb, independent from functional form, other properties of the functions and correlations of starting values with optimal allocations. He considers for each of the four functions discussed in the previous section eight constellations of different values of properties and correlations (for more information see Appendix C.1). Sales are computed based on these functions without adding disturbances though the latter are a component of most econometric models. Results obtained by his iterative algorithm are compared to those computed by the three rules of thumb, which we explain in Sect. 5 below. Overall, his iterative algorithm outperforms all rules of thumb by far.

In our study the iterative algorithm of Albers turns out not to work well for non-deterministic sales response functions with additive disturbances. Disturbances that directly affect sales cause elasticities and new allocation values to fluctuate. In addition to moving allocations far away from the optimal solution, these fluctuations sometimes even lead to negative slopes and hence negative elasticities.

To overcome these problems we develop a modified exploration–exploitation algorithm. In the exploration stage we use a modification of the iterative algorithm (Albers 1997). In the exploitation stage, we start with an approximation to the unknown sales response functions for each unit using a data set generated in the exploration stage. We use these approximations as inputs of a quadratic programming problem whose solution provides new allocations to the units. These new allocations and sales due to these allocations are added to the data. Based on the data set enlarged this way, approximations of sales response functions are updated, for which new allocations are determined by quadratic programming and so on for several iterations (i.e., periods). In each iteration the data are extended by allocations determined by an approximate optimal program and their corresponding sales. Therefore, the fit of the approximate response functions in a neighborhood close to the optimal solution gets more important.

4.1 Exploration

For exploration we modify the iterative algorithm of Albers (1997). Estimated elasticities are heavily influenced by the additive disturbances and often are even negative. That is why as first modification we project each elasticity value into the interval [0.01, 0.50], i.e., if the calculated value is above 0.50 it is set to 0.50 and to 0.01 if it is below 0.01. This value range conforms to the results of meta-analyses of advertising elasticities (Assmus et al. 1984; Sethuraman et al. 2011). Again, in an actual application a practitioner may choose any interval reasonable based on his expectations of the limits of the elasticity.

As second modification we add first order exponential smoothing in the following manner:

$$\tilde{\epsilon}_t := (1 - \beta)\tilde{\epsilon}_{t-1} + \beta\hat{\epsilon}_t \tag{13}$$

with $\beta \in [0, 1]$. We use $\beta := 0.85$ based on a Monte Carlo simulation with β as a regressor with discrete values between 0.05 and 1. However, other values can be used, based on how much influence of the previous period is desired. $\tilde{\epsilon}_t$ is the smoothed elasticity and $\hat{\epsilon}_t$ the estimated elasticity in period t .

Exponential smoothing reduces the fluctuation of elasticities while still emphasizing more recent data points, which as a rule are nearer to the optimum. Exponential smoothing also serves to move elasticities away from the boundary values of 0.01 and 0.50.

Example 6 Returning to the initial Example 1, we might deduce that unit 1 yields higher sales than the other two, and suggest a second allocation of (4, 1, 1), which then (in the deterministic case) gives an objective value of

$$f_1(4) + f_2(1) + f_3(1) = 5\sqrt[3]{4} + 3\sqrt[8]{1} + 3\sqrt[8]{1} \approx 8 + 3 + 3 = 14.$$

As the functional form is invisible to the allocation methods, we now need to estimate the elasticities, using the formula

$$\hat{\epsilon}_i := \frac{\Delta f_i x'_i}{\Delta x_i f'_i}$$

where f'_i and x'_i denote the values in the second period of unit i . We obtain

$$\hat{\epsilon}_1 = \frac{7.94 - 6.3}{4 - 2} \cdot \frac{4}{7.94} \approx 0.41, \quad \hat{\epsilon}_2 = \hat{\epsilon}_3 \approx 0.09$$

which is close enough to the actual elasticities of $\epsilon_1 = 1/3, \epsilon_2 = \epsilon_3 = 0.125$.

For the next allocation we use conditions (6) to obtain

$$x_1 = B \frac{\hat{\epsilon}_1 f'_1}{\sum_{i=1}^3 \hat{\epsilon}_i f'_i} = 6 \cdot \frac{3.27}{3.82} = 5.15, \quad x_2 = x_3 = 0.43,$$

which in turn results in a new objective value of

$$f_1(5.15) + f_2(0.43) + f_3(0.43) = 8.63 + 2.7 + 2.7 = 14.03$$

and we estimate the next elasticities. For example, we obtain

$$\epsilon_1 = \frac{8.63 - 7.94}{5.15 - 4} \frac{5.15}{8.63} = 0.36.$$

Had that value been higher than 0.5, we would have set it to 0.5. Before calculating the next allocation, we perform exponential smoothing by setting

$$\tilde{\epsilon}_1 = (1 - 0.85) \cdot 0.41 + 0.85 \cdot 0.36 = 0.3675.$$

We summarize these deterministic results on the left hand side of Table 4.

However, the deterministic approach is not in line with econometric sales response functions. Therefore, as shown in Example 5, we sample values u from a normal distribution whenever sales f are calculated and add them to obtain $f_u := f + u$.

Observe now the right hand side of Table 4. In the first two periods, very similar data result, however in period three, the error greatly influences the estimation of elasticities with $\epsilon_3^1 = 0.64 > 0.5$ and $\epsilon_3^3 = -0.43 \ll 0.01$. A naive continuation with Alber’s original rule would lead to implausible negative x -values. As explained, we set those elasticities to the interval boundaries of 0.50 and 0.01, respectively. We then apply exponential smoothing, to give them a small boost in the right direction.

Table 4 Deterministic and stochastic examples

Period t	Deterministic			Variable	Stochastic		
	Unit 1	Unit 2	Unit 3		Unit 1	Unit 2	Unit 3
1	2	2	2	x^1	2	2	2
1	6.36	3.3	3.3	f_1	6.36	3.3	3.3
1				u^1	-0.18	0.04	0.32
1				f_u^1	6.12	3.31	3.59
2	4	1	1	x^2	4	1	1
2	8	3	3	f^2	8	3	3
2				u_2	-0.23	-0.02	0.03
2				f_u^2	7.71	2.98	3.03
2	0.41	0.09	0.09	ϵ_2	0.41	0.11	0.19
3	5.15	0.43	0.43	x^3	4.68	0.49	0.83
3	8.63	2.7	2.7	f^3	8.37	2.74	2.93
3				u_3	0.14	-0.05	0.4
3				f_u^3	8.51	2.69	3.33
3	0.36	0.08	0.08	ϵ_3	0.64	0.1	-0.43
3	0.36	0.08	0.08	$proj(\epsilon_3)$	0.5	0.1	0.01
3	0.37	0.08	0.08	$\tilde{\epsilon}_3$	0.48	0.1	0.04

In the following, we show how the algorithm tackles an alternative interval of elasticity values. To this end we assume that the decision maker operates in a specific industry and look up sales force elasticities for pharmaceuticals in a published meta-analysis (Albers et al. 2010). Figure 1C of this publication shows that most elasticities for pharmaceuticals fall into the interval $[0.05, 0.35]$. We therefore change the values of $\varepsilon_3^1, \varepsilon_3^3$ to 0.35 and 0.05, respectively and continue the calculation. The updated smoothed elasticities (0.36, 0.10 and 0.07) are a lot closer to the exact values of $\frac{1}{3}$ and $\frac{1}{8}$. That makes sense intuitively, because we restrict the search to a smaller interval that contains all the sought elasticity values.

4.2 Exploitation

In the exploration stage the algorithm generates data points close to the optimum for a fixed number of periods. Nevertheless, elasticities still jump around too much. Therefore we need a method that dampens disturbances.

The general idea of the exploitation stage can be easily understood when looking at modified exponential functions with varying parameters. Figure 1 shows an example of the modified exponential function for one of the eight sales units upon which the budget of 8,000,000 is to be allocated, where x is the budget allocated to that unit and y its sales value. The optimal allocation to the unit in the example drawn in Figure 1 is around 640,000. Nevertheless, even when disturbances are small, the algorithm and each of the three rules of thumb still fluctuate a lot, showing no sign

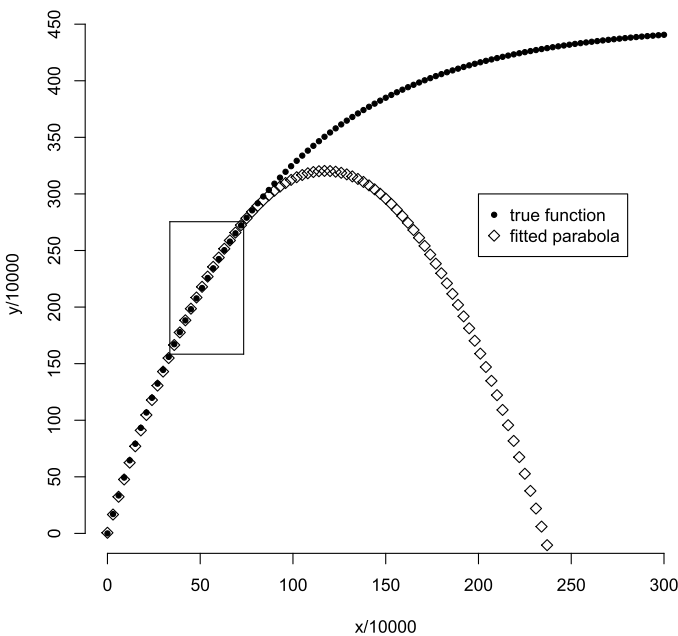


Fig. 1 True function and fitted parabola

of stability, although they do not leave a certain interval of the domain in each variable (and hence of the codomain).

This area, highlighted by a rectangle in the figure, can be easily approximated by a parabola, i.e., a polynomial of degree two. Assuming the functions were actually polynomials of degree two, quadratic programming gives an exact solution, because the total resource restriction (3) is linear.

Exploitation consists of two steps. In the first step, for each unit i a quadratic regression of the form

$$y_i = \gamma_{i,0} + \gamma_{i,1} \cdot x_i + \gamma_{i,2} \cdot (x_i)^2 \quad (14)$$

is performed, where y_i and x_i are the sales and allocation values for unit i and $\gamma_{i,j}$ are the coefficients to be estimated. For each regression, the data consist of all periods observed so far (using only a smaller number of the most recent values does not improve results).

In the second step, an allocation that is optimal for these approximations obeying the total resource restriction is determined by quadratic programming using the method of Goldfarb and Idnani (1983). This allocation and its corresponding sales value constitute an additional data point for the next quadratic regression. This two-step process is repeated until a total of 40 periods is reached, where 40 is the planning horizon within our study.

A problem arises when $\gamma_{i,2}$ is estimated as a positive number for any sales unit, as the matrix in the quadratic program is no longer positive definite. In particular, this means that the sampled area suggests a convex shape of the sales function, which is problematic. If this estimated function is monotonically increasing this problem can be remedied by setting $\gamma_{i,1}$ to the slope of the regression line, and setting $\gamma_{i,2}$ to a negative value close to zero (we choose -10^{-15}). This way the quadratic program is forced to approximate a straight line by a parabola. Substantial arguments justify this procedure as well. One argument is the predominance of empirical evidence for concave sales response functions with decreasing marginal returns. Moreover, decision makers typically choose values in the region of decreasing marginal returns if they think that sales response is S-shaped (Hanssens et al. 2001; Simon and Arndt 1980). Also please note that the functional forms underlying our simulation are concave with the only exception of the ADBUG function for a parameter $\phi_S > 1$.

In the worst-case scenario, in addition the slope of the regression line may be negative. This case is very rare and the allocation to this unit will almost certainly be zero. One should remember however, what this actually means: The shape of the data points resembles a monotonically decreasing, convex (!) function and would hence arise either from a few very unfavorable disturbances in a row or an exogenous influence that cannot be explained by additive disturbances. Surely, in this case, the function should be thoroughly analyzed instead of continuing the application of *any* algorithm. We will briefly deal with this situation in Sect. 7.

It is furthermore reasonable to ask whether exploration and exploitation are both necessary, and what the optimal period is to switch from exploration to exploitation. Figure 2 shows the average means of total sales and optimality (defined as ratio of total sales and optimal total sales) for all constellations of the

exploration–exploitation procedures, for switches from exploration to exploitation at $t = 3, 5, 10, 20, 30$ periods.

Switching after three periods represents the case of *only exploitation* and switching after 30 periods is essentially *only exploration*. Figure 2 implies visually, that neither of these extremes are useful, and a balance between them must be found. For this simulation, we therefore switch from exploration to exploitation at period 10.

Of course, in a practical problem setting one cannot use such a simulation to decide about the switching period. This decision can be based on observed changes of total sales between consecutive periods. We recommend switching as soon as these changes become small.

Example 7 Focusing on the key components of the right hand side of Table 4 from Example 6, we see that the sum of Sales of the three units, i.e. the objective, increases. In period 1 we have $6.12 + 3.31 + 3.59 = 13.02$, period 2 has 13.72 and period 3 yields 14.53. However, naively continuing this process gives the following results for the next two periods in Table 5.

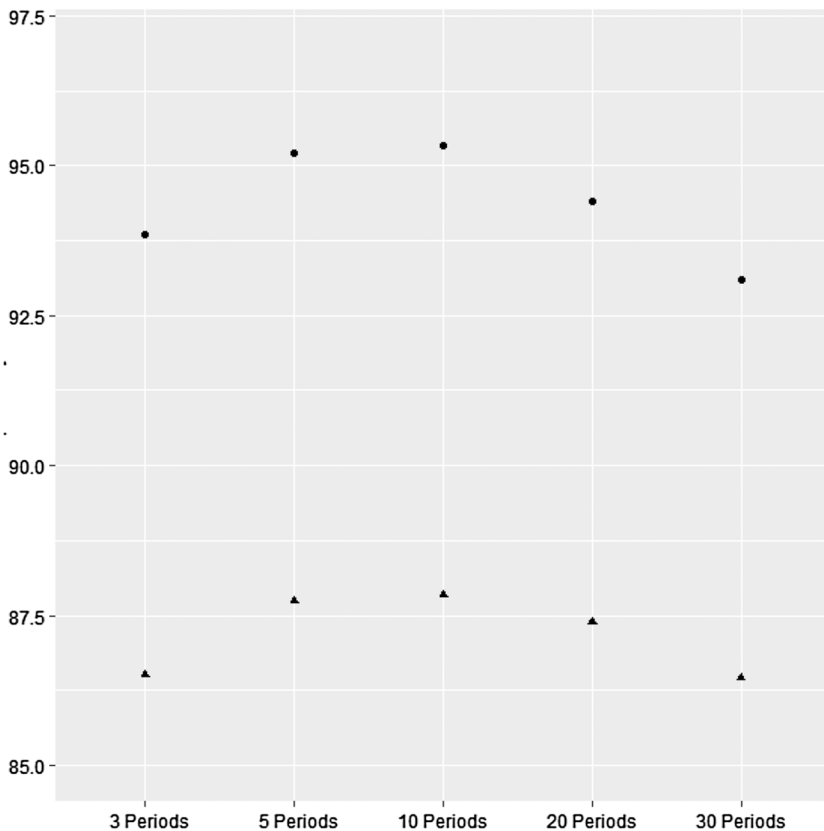


Fig. 2 Average means of sales and optimality for switching after t periods

In particular, the sales value in period 4 is $8.78 + 2.72 + 2.59 = 14.09$ and in period 5 we have 13.9. In other words, the objective has decreased in the last two periods. This means that the exploration phase can only go that far, and in particular, is still rather unstable.

However, we now have sufficient data to estimate parabolas for the individual sales response functions. For unit 1 we obtain the parabola

$$y_1 = \gamma_{1,0} + \gamma_{1,1} \cdot x_1 + \gamma_{1,2} \cdot (x_1)^2 \approx 3.848 + 1.25825 \cdot x_1 - 0.06491 \cdot (x_1)^2$$

which is sketched in Fig. 3

We now use the obtained coefficients and solve the approximate allocation problem described in Sect. 4.2 by quadratic programming to determine the allocation onto the parabolas of $x_6 = (4.56, 0.91, 0.53)$ and its sales values $f_6 = (8.29, 2.96, 2.77)$, bringing us back to a good objective value of 14.02. We can now add this allocation to the dataset, update the estimates for the coefficients of the parabolas, get a new allocation by quadratic programming of $x_7 = (4.71, 0.82, 0.47)$, sales of $f_7 = (8.38, 2.93, 2.73)$ and an objective of 14.04. We obtain consistently good values that still are improving. It may first seem that the improvement is rather small, but we will immediately see the reason for that.

Let us take a look at the optimal solution of the decision problem described in Section . By iterating the rule from Eq. (6) with the true elasticities (i.e., the exponents) $\epsilon = (1/3, 1/8, 1/8)$, we obtain $x^* = (4.8, 0.6, 0.6), f^* = (8.43, 2.81, 2.81)$ with an optimal objective of 14.05. In other words, in period 3, the exploration algorithm

Table 5 Example: switch from exploration to exploitation

Period t	Variable	Unit 1	Unit 2	Unit 3
1	x^1	2	2	2
1	f_u^1	6.12	3.31	3.59
2	x^2	4	1	1
2	f_u^2	7.71	2.98	3.03
3	x^3	4.68	0.49	0.83
3	$\tilde{\epsilon}_3$	0.48	0.1	0.04
4	x^4	4.47	0.37	0.16
4	f^4	8.81	2.65	2.39
4	u_4	- 0.03	0.08	0.2
4	f_u^4	8.78	2.72	2.59
4	ϵ_4	0.22	- 0.04	0.07
4	$proj(\epsilon_4)$	0.22	0.01	0.07
4	$\tilde{\epsilon}_4$	0.26	0.02	0.06
5	x^5	5.47	0.16	0.37
5	f^5	8.81	2.37	2.65
5	u_5	- 0.08	- 0.21	0.36
5	f_u^5	8.73	2.16	3.01

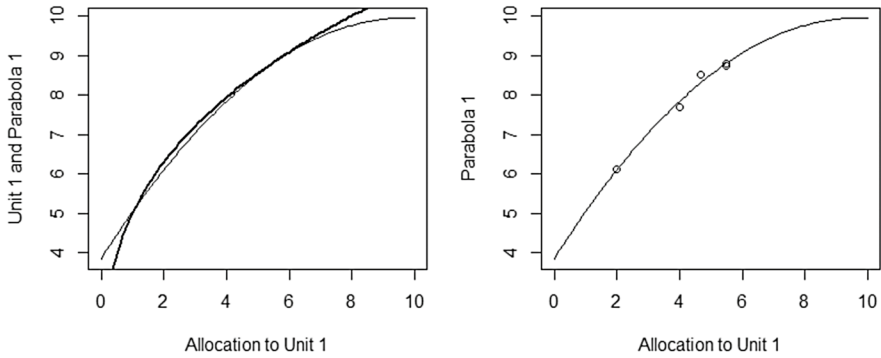


Fig. 3 Estimated parabola of unit 1, original function of unit 1 (bold), points of 5 periods

got lucky to get that close to the optimal solution, but deviated from that in the following two periods. The exploitation brought us back on track towards the optimal solution.

5 Applying the rules of thumb

We continue with an example to demonstrate how the three rules of thumb introduced in Sect. 1 are applied.

Example 8 Starting from the allocation (4, 1, 1) from Example 6, the first rule suggests to use as next allocation $x'_i = B \cdot f_i / (\sum_{i=1}^3 f_i)$ leading to

$$x'_1 = 7.9 / (7.9 + 3 + 3) = 3.42, \quad x'_2 = x'_3 = 3 / 13.9 = 1.29$$

and an objective value of

$$7.53 + 3.1 + 3.1 = 13.73$$

Rule 2 suggests using $x'_i = B \cdot \frac{f_i}{x_i} / (\sum_{i=1}^3 \frac{f_i}{x_i})$ leading to

$$x'_1 = 1.5, \quad x'_2 = x'_3 = 2.25$$

and an objective value of

$$5.7 + 3.3 + 3.3 = 12.3.$$

The third rule uses the maximal observed values of each unit. For the first unit, that is 7.94, for the second and third unit that is 3.27. So we allocate

$$x_1 = B \cdot \frac{7.94}{7.94 + 3.27 + 3.27} = 3.29, \quad x_2 = x_3 = B \cdot \frac{3.27}{7.94 + 3.27 + 3.27} = 1.35$$

and obtain an objective value of

$$7.44 + 3.11 + 3.11 = 13.66.$$

6 Evaluation of procedures

For the sake of comparing the developed algorithm, we conduct a simulation study with 384 constellations, which result from four function types, four procedures, similar/varied elasticities, similar/varied saturations, two budget-levels, three disturbance levels and generate twenty replications for each constellation.

We therefore define the following dummy-coded variables for our models:

“Proc” corresponds to the three rules of thumbs and the developed algorithm, “Form” corresponds to the four types of functions, “Dist” corresponds to the R^2 values of 0.9, 0.7 and 0.5, “Elas” and “Satu” correspond to “similar” and “varied” elasticities and saturations, respectively. “Budg” corresponds to the two budget levels of 1,000,000 and 8,000,000.

The allocation problem of each constellation can be written as

$$\begin{aligned} \max \quad & \sum_{t=1}^{40} \sum_{i=1}^8 f_i(x_{t,i}) \\ \text{s.t.} \quad & x_{t,1} + \dots + x_{t,8} = B \quad \forall t \in \{1, \dots, 40\}, \end{aligned} \quad (15)$$

Allocations $x_{t,i}$ may depend on previous sales and allocations, i.e., all values $f_i(x_{\tau,i})$ and $x_{\tau,i}$ with $\tau < t$ may be used to determine the $x_{t,i}$.

The performance of procedures is measured by two different dependent variables, “Sales” and “Optimality”, which both are computed as arithmetic means across 40 periods. These two dependent variables normalize total sales attained by rules of thumb or the algorithm. We want normalizations to differ with respect to the (non-) consideration of additive random disturbances. That is why the denominator in the definition of “Sales” includes random disturbances, which on the other hand are excluded by the denominator in the definition of “Optimality”.

The first dependent variable “Sales” equals total sales divided by the maximal attainable value given a fixed functional form and a fixed budget level over all other constellations. The other constellations result from four allocation procedures, three disturbance levels, two elasticity levels and two saturation levels. The division to compute “Sales” is necessary as the four function types yield quite different values for total sales (which is especially pronounced in the case of ADBUDG-functions) and the maximal attainable sales depend on the budget level.

The second dependent variable “Optimality” is defined as ratio of total sales and optimal total sales. Optimal total sales are determined by optimizing based on the true response functions without disturbances, i.e., assuming knowledge of sales response functions and their parameters. “Optimality” therefore shows to what extent a procedure that lacks knowledge of the response functions attains optimal total sales on average. A value of 1.0 for “Optimality” indicates that average total sales as a rule equal their optimal value.

As mentioned above, the S-shaped ADBUDG-function is not concave, and hence a problem of local and global optima may arise. In the case of this functional form we therefore search for the optimal solution by means of nine algorithms (the heuristic of Albers without adding of disturbances, the three rules of thumb both without and with adding disturbances, the developed algorithm, and generalized simulated annealing from the R package Ganesa). As these algorithms may stuck in local optima, we could obtain optimalities greater than 1.0. To make sure that values of this dependent variable are not greater than 1.0, we rescale optimalities by dividing by their maximum value.

We expect the developed algorithm to lead to higher sales and sales closer to the optimum compared to each of the three investigated rules of thumb. In the following we examine these expectations by descriptive statistics and appropriate statistical tests.

For each of the four procedures Table 6 gives arithmetic means of both dependent variables. The algorithm attains the highest (i.e., best) values for both dependent variables, followed by the third and the first rule of thumb. The second rule of thumb attains the worst values. The third rule of thumb performs fairly well for both dependent variables. Nevertheless, the developed algorithm achieves 7.93% and 20.14% of the maximum improvement left over by the third rule of thumb for “Sales” and “Optimality”, respectively (computed as $100(0.8711 - 0.8600)/(1 - 0.8600)$ and $100(0.9536 - 0.9419)/(1 - 0.9419)$). Moreover, several statistical tests, which we present in the following, clearly show that the developed algorithm is significantly better.

6.1 Results for main effect regression models

We start from the following two linear regression models comprising main effects only to thoroughly compare the procedures:

$$Sales = \beta_{1,0} + \beta_{1,1}Proc + \beta_{1,2}Form + \beta_{1,3}Dist + \beta_{1,4}Elas + \beta_{1,5}Satu + \beta_{1,6}Budg + e \tag{16}$$

$$Optimality = \beta_{2,0} + \beta_{2,1}Proc + \beta_{2,2}Form + \beta_{2,3}Dist + \beta_{2,4}Elas + \beta_{2,5}Satu + \beta_{2,6}Budg + e \tag{17}$$

In these equations for $i = 1, 2$, $\beta_{i,0}$ is the intercept, $\beta_{i,1}, \beta_{i,2} \in \mathbb{R}^3, \beta_{i,3} \in \mathbb{R}^2, \beta_{i,4}, \beta_{i,5}, \beta_{i,6} \in \mathbb{R}$ are coefficient vectors, each multiplied by dummy-coded variables, and e is the usual normally distributed error term. Reference categories of these binary dummy variables are the developed algorithm, the multiplicative function, low disturbances (i.e., a sales response functions with high

Table 6 Arithmetic means of dependent variables for each procedure

Dependent variable	Algorithm	Rules of thumb		
		1	2	3
Sales	0.8711	0.8522	0.7723	0.8600
Optimality	0.9536	0.9305	0.8268	0.9419

R^2 values of about 0.9), the low budget, and similar values for elasticities and saturations, respectively.

Table shows the estimation results of the main effect models both for “Sales” and “Optimality”. Coefficients for procedures are in line with the arithmetic means of Table 6. They reflect that all rules of thumb perform worse than the developed algorithm (the reference category). These coefficients also indicate that allocation proportional to maximal sales of a unit observed so far (Proc4) performs better than the other two rules of thumb. The second rule of thumb (Proc3), allocation proportional to sales of a unit divided by its allocation in the previous period, clearly turns out as overall worst procedure.

The other coefficients of Table 7 indicate whether we obtain higher or lower values of the two dependent variables for a certain category of the respective independent variable. To our opinion most of these results are intuitive. The higher complexity of S-shaped ADBUDG functions (= Form4) leads to lower values than the other functional forms. Both the modified exponential (= Form2) and the concave ADBUDG function are better than the multiplicative function (the reference category). The coefficients for Dist2 and Dist3 reflect that a higher level of disturbances makes it difficult to attain good values of both dependent variables. Varied elasticities are accompanied by lower values “Optimality”, but do not have a significant effect on “Sales”. If saturations are varied, values of “Sales” are higher, but values of “Optimality” are lower. Higher values of “Sales” are simply a consequence of upward-scaled functions. On the other hand, varied saturations lead to higher derivatives and fluctuations, which both make optimization more difficult.

Table 7 Main effect regression models

Variable	Dependent variable: sales		Dependent variable: optimality	
	Coefficient	t-value	Coefficient	t-Value
Intercept	0.81	204.130***	1.014	274.947***
Proc2	- 0.019	- 5.833**	- 0.023	- 7.687***
Proc3	- 0.099	- 30.485***	- 0.127	- 42.108***
Proc4	- 0.011	- 3.42***	- 0.012	- 3.889**
Form2	0.042	12.905***	0.023	7.476***
Form3	0.078	23.936***	0.013	4.406***
Form4	- 0.135	- 41.729***	- 0.096	- 31.827***
Dist2	- 0.011	- 3.795***	- 0.021	- 8.118***
Dist3	- 0.038	- 13.605***	- 0.067	- 25.674***
Elas2	0.004	1.884	- 0.046	- 21.608***
Satu2	0.067	29.382***	- 0.036	- 17.067***
Budg2	0.091	39.491***	0.05	23.640***
$df = 7, 668$	$R^2 = 0.5325, R^2_{adj} = 0.5318$		$R^2 = 0.4485, R^2_{adj} = 0.4478$	

*p-value < 0.05, **p-value < 0.01, ***p -value < 0.001

Decision makers should expect sales to remain below their optimal value, if elasticities or saturation levels vary across sales units. Such an expectation also applies if sales response is S-shaped rather than concave.

Positive and significant coefficients indicate that higher budgets favor higher values of both dependent variables. One can show that the dependent variable “Sales” increases if at the high budget level additional total sales are greater or not much lower than additional maximal attainable sales. This condition is usually fulfilled due to the fact that except for the S-shaped ADBUG function all investigated sales response functions are characterized by decreasing positive marginal effects (see Appendix D). We obtain an analogous result for the dependent variable “Optimality” by replacing maximal attainable sales with optimal total sales.

6.2 Results for regression models with interactions

We also estimate regression models that include certain sets of interaction terms. The most simple of these models add all pairwise interactions to the main effects.

Following suggestions of one anonymous reviewer, we now discuss the models with added pairwise interactions in more detail. We start by providing Cohen’s partial eta squared values (see Table 8). These values give the proportion of variance of the respective dependent variable attributed to an experimental factor or interaction adjusted by all the other factors and interactions (Cohen 1988).

Like Richardson (2011) we set 0.0099, 0.0588, and 0.1379 as thresholds for small, medium and large partial eta squared values, respectively. Six and four pairwise interactions, respectively, are characterized by medium or large partial eta squared values for the dependent variables “Sales” and “Optimality”, respectively. These results provide evidence that models including pairwise interactions are to be preferred over models with main effects only.

High partial eta squared values of a factor or interaction reflect its importance from a managerial point of view. The factor Procedure, which is central to our investigation, turns out to be important. With respect to the dependent variable “Optimality” we even obtain for Procedure the highest partial eta squared value across all factors and interactions. The factor Form together with interactions between Procedure and Disturbance as well as between Form and Budget are important for both dependent variables. Importance also accrues to the factors Saturation and Budget with respect to the dependent variable “Sales”, to the factor Disturbance as well as to the interactions between Procedure and Form with respect to the dependent variable “Optimality”.

Let us focus on the total effects of the levels of the factor Procedure in the models with pairwise interactions. That is why we apply Tukey’s test (Jobson 1991) to assess the performance of procedures. This test shows whether two procedures have different means for any of the two dependent variables “Sales” and “Optimality”. To this end it determines a confidence interval for each of the six pairs of the four investigated procedures using a studentized range distribution. Tukey’s test takes all

Table 8 Partial eta squared values of factors and pairwise interactions

	Models with interactions		Main effect models	
	Sales	Optimality	Sales	Optimality
Procedure	0.2774	0.3703	0.1314	0.2271
Form	0.6208	0.3405	0.3923	0.2051
Disturbance	0.0612	0.1524	0.0251	0.0824
Elasticities	0.0012	0.1086	0.0005	0.0574
Saturation	0.2221	0.0706	0.1012	0.0366
Budget	0.3403	0.1273	0.1690	0.0679
Procedure: form	0.0936	0.1847		
Procedure: disturbance	0.1768	0.2197		
Procedure: elasticities	0.0182	0.0131		
Procedure: saturation	0.0915	0.0848		
Procedure: budget	0.0153	0.0447		
Form: disturbance	0.0170	0.0127		
Form: elasticities	0.0705	0.0273		
Form: saturation	0.0979	0.0207		
Form: budget	0.4478	0.1893		
Disturbance: elasticities	0.0006	0.0008		
Disturbance: saturation	0.0163	0.0203		
Disturbance: budget	0.0088	0.0002		
Elasticities: saturation	0.0068	0.0085		
Elasticities: budget	0.0305	0.0000		
Saturation: budget	0.0065	0.0009		

main effects and pairwise interactions effects into account as the studentized range distribution depends on the error sum of squares of the respective regression model.

We consider the means of a dependent variable for two procedures to be different if the respective confidence interval does not include the value zero. Table 9 contains the results of Tukey's tests for each of the two dependent variables. All differences shown in this table are greater than zero at a confidence level of 95%. For both dependent variables the third rule of thumb (Proc4) beats the other two rules of thumbs (Proc2 and Proc3), but the developed algorithm (Proc1) attains higher values than any of the three rule of thumbs.

In other words, we obtain the same overall performance ranking of the procedures as for the main effect models. However, main effect models and models with pairwise interactions disagree with respect to the importance of factors. As Table 8 shows partial eta squared values of the factor Procedure are much higher for models with pairwise interactions. The models with interactions also attribute higher importances for the other factors.

Table 10 contains the estimation results of the two regression models with main effects and pairwise interactions. We select four interactions because of the absolute size of their pairwise interaction coefficients. Figure 4 shows plots

Table 9 Tukey's tests

Difference	Dependent variable: sales			Dependent variable: optimality		
	Mean	95% confidence interval		Mean	95% confi- dence interval	
Poc2–Proc3	0.080	0.075	0.085	0.104	0.098	0.109
Proc4–Proc3	0.088	0.082	0.093	0.115	0.110	0.121
Proc1–Proc3	0.098	0.094	0.104	0.127	0.121	0.132
Proc4–Proc2	0.008	0.003	0.013	0.011	0.006	0.017
Proc1–Proc2	0.019	0.014	0.024	0.023	0.018	0.029
Proc1–Proc4	0.011	0.006	0.016	0.012	0.006	0.017

of these four interactions for each dependent variable. Clearly, rule of thumb 2 (Proc3) performs worst with respect to both dependent variables. The disadvantage of this rule of thumb increases with the disturbance level (i.e., for Dist2 and Dist3).

The developed algorithm (Proc1) performs best, followed by rule of thumb 1 (Proc2), if sales response is generated by the S-shaped ADBUDG function. For each of the other functional forms the developed algorithm and rule of thumb 3 (Proc4) turn out to be about equal with respect to the dependent variable “Sales”. On the other hand, rule of thumb 3 attains slightly higher values for the dependent variable “Optimality”. Overall, rule of thumb 2 (Proc3) performs worst, especially for the S-shaped ADBUDG function. These results suggest that managers who expect an S-shaped sales response should prefer the developed algorithm to the investigated rules of thumb.

A higher budget together with sales generated by an ADBUDG function (no matter whether concave or S-shaped) leads to higher values of both dependent variables. For the other functional forms we do not obtain differences between high and low budgets.

To consider additional interaction terms, we also investigate models (one for each dependent variable) with all triple, quadruple, quintuple and sextuple interactions. Quite interestingly, the full models with sextuple interactions not only lead to the best variance explanations (adjusted R^2 values amount to 0.9257 and 0.8705 for “Sales” and “Optimality”, respectively), but F-tests also confirm that the full models outperform all the restricted models. A comprehensive discussion can be found in Appendix B.1.

Because of the superiority of the full models we need a test of differences between the algorithm and each rule of thumb that goes beyond the main effects model by also considering all interactions. We therefore use a method described in Gahler (2020)

to construct a vector (= matrix with one row) of multiplicities (= appearances) for the coefficients in the OLS-estimator. The corresponding matrix (whose product with the coefficient vector is zero under the null hypothesis) is the basis of an F-test. Hence, a positive product indicates that the algorithm performs better than a rule

Table 10 Pairwise interaction effects regression models

Variable	Dependent variable: sales		Dependent variable: optimality	
	Coefficient	t-value	Coefficient	t-value
(Intercept)	0.778	139.155***	1.005	171.887***
Proc2	- 0.001	- 0.165	- 0.003	- 0.446
Proc3	0.017	2.781**	- 0.007	- 1.016
Proc4	- 0.006	- 0.984	- 0.010	- 1.621
Form2	0.08	13.041***	- 0.003	- 0.451
Form3	0.072	11.721***	- 0.046	- 7.146***
Form4	- 0.276	- 44.992***	- 0.113	- 17.635***
Dist2	0.006	1.084	- 0.004	- 0.691
Dist3	0.012	2.141*	- 0.018	- 3.04**
Elas1	0.043	8.904***	- 0.047	- 9.384***
Satu1	0.145	30.387***	0.012	2.484*
Budg2	0.043	9.049***	- 0.026	- 5.168***
Proc2:Form2	- 0.007	- 1.218	- 0.007	- 1.135
Proc3:Form2	0.047	8.068***	0.058	9.574***
Proc3:Form3	0.050	8.652***	0.058	9.528***
Proc2:Form4	- 0.021	- 3.634***	- 0.028	- 4.583***
Proc3:Form4	- 0.076	- 13.155***	- 0.139	- 23.085***
Proc4:Form4	- 0.066	- 11.412***	- 0.081	- 13.393***
Proc3:Dist2	- 0.045	- 9.045***	- 0.049	- 9.403***
Proc3:Dist3	- 0.157	- 31.487***	- 0.187	- 35.682***
Proc4:Dist3	0.008	1.596	0.014	2.717**
Proc2:Elas1	- 0.014	- 3.392***	- 0.016	- 3.666***
Proc3:Elas1	- 0.044	- 10.842***	- 0.044	- 10.197***
Proc2:Satu1	- 0.013	- 3.258**	- 0.014	- 3.277**
Proc3:Satu1	- 0.099	- 24.192***	- 0.1	- 23.31***
Proc4:Satu1	- 0.008	- 1.976*	- 0.009	- 2.088*
Proc3:Budg2	0.036	8.869***	0.072	16.797***
Proc4:Budg2	0.029	7.138***	0.035	8.262***
Form2:Dist2	0.016	2.952**	0.019	3.547***
Form3:Dist2	0.014	2.885**	0.015	2.8**
Form4:Dist2	0.022	4.427***	- 0.006	- 1.084
Form2:Dist3	0.008	1.697	0.026	4.922***
Form3:Dist3	0.030	6.042***	0.041	7.834***
Form4:Dist3	0.050	9.954***	0	0.066
Form2:Elas1	- 0.059	- 14.342***	0.016	3.758***
Form3:Elas1	- 0.011	- 2.575*	- 0.024	- 5.556***
Form4:Elas1	0.039	9.51***	0.037	8.579***
Form2:Satu1	- 0.031	- 7.576***	- 0.017	- 3.897***
Form3:Satu1	- 0.097	- 23.678***	0.031	7.166***
Form4:Satu1	0.009	2.283*	- 0.03	- 7.051***
Form2:Budg2	- 0.022	- 5.282***	- 0.001	- 0.277

Table 10 (continued)

Variable	Dependent variable: sales		Dependent variable: optimality	
	Coefficient	t-value	Coefficient	t-value
Form3: Budg2	0.064	15.614***	0.046	10.75***
Form4: Budg2	0.226	65.131***	0.155	36.357***
Dist2: Elas1	-0.007	-2.112*	-0.011	-2.441*
Dist2: Satu1	-0.017	-4.685***	-0.003	-5.555***
Dist3: Satu1	-0.040	-11.201***	-0.021	-12.54***
Dist2: Budg2	-0.015	-4.113***	-0.046	-0.486
Dist3: Budg2	-0.029	-8.208***	0.002	1.216
Elas1: Satu1	0.021	7.205***	0.004	8.099***
Elas1: Budg2	-0.045	-15.479***	0	0.025
Satu1: Budg2	-0.020	-7.074***	-0.008	-2.683**
	$R^2 = 0.8157, R^2_{adj} = 0.8143$		$R^2 = 0.7245, R^2_{adj} = 0.7224$	

*p-value < 0.05, **p-value < 0.01, ***p-value < 0.001

Contains only pairwise interactions whose coefficients are > 0.005 in absolute size

of thumb, and the F-statistic reveals the significance of better performance. A more detailed explanation can be found in Appendix B.2.

For the dependent variable “Sales” we obtain average differences between the algorithm and each of the three rules of thumb of 1.6093, 9.3232, and 0.8757, respectively (the corresponding F-statistics are 176.4363, 5,922.021, and 52.2406). These differences are all significant at a level (far) below 0.001. For the other dependent variable “Optimality” we obtain differences of 1.9223, 11.8959, and 0.8215, respectively (F-statistics are 198.7278, 7,610.8667, and 36.2925). These differences are all significant at a level (far) below 0.001.

In the analysis above, we use 20 replications as we have a large model and wish to have 20 times as many data points as coefficients in the full regression model. To answer a question from one anonymous reviewer we reduce the number of replications to five. We also perform a Bonferroni correction to take into account that we make multiple comparisons between procedures. This additional analysis confirms the robustness of our results. It shows that the effect differences between the rules of thumb and our procedure as described above keep their correct signs (i.e., the effect of the procedure is higher than the effect of each rule of thumb), the corresponding F-tests remain significant below $\alpha = 0.001$ and the combined significances from the Bonferroni correction remain below 0.001 as well. Therefore, even under these restrictive conditions, the null hypotheses can be rejected, individually and collectively.

Furthermore, issues might be raised concerning the similarity of the results between the developed algorithm and the third rule of thumb. Therefore, we investigate the distributions of the two dependent variables after 40 periods by means of box-plots (see Fig. 5). We see that the third rule of thumb not only has more outliers, but also more extreme outliers. With respect to optimality, there

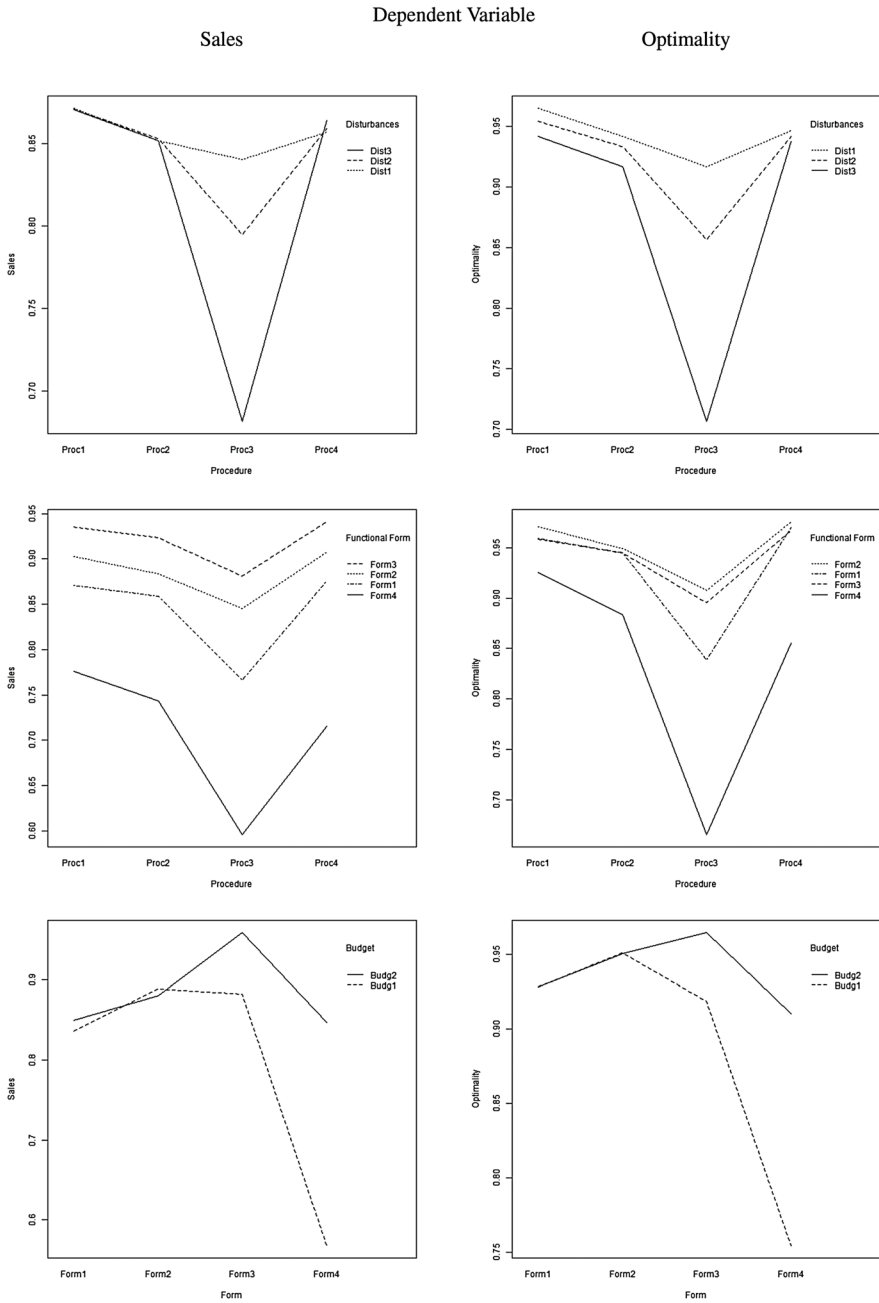


Fig. 4 Pairwise interaction plots

are individual outliers below 0.5, i.e., this rule leads to values over the entire planning horizon that are below half of the optimal possible values and quite a lot of the outliers are below 0.7. For sales, this holds true for all methods, however the third rule of thumb now has many outliers below 0.4.

7 Conclusion

Our study shows that the iterative algorithm of Albers (1997) is not appropriate if sales response functions are not deterministic and include additive disturbances. Disturbances that directly affect sales cause elasticities and new allocation values to fluctuate and sometimes even produce negative elasticities. Let us remind you that elasticities should be positive and lie in a certain interval (see Sect. 4.1). Moreover, performance of this algorithm does not really improve if elasticity estimates are only projected into this interval. Our iterative algorithm does not suffer from such problems.

A conventional approach begins by estimating a sales response function for each unit based on a given data set of sales and allocations. Then numerical optimization using the estimated sales response functions solves the proper allocation problem. Because of its iterative nature, our algorithm is different. In the exploitation stage, we start with an approximation to the unknown sales response functions for each unit using a data set generated in the exploration stage. These approximations serve as inputs of a quadratic programming problem whose solution provides new allocations to the units. These new allocations and sales due to these allocations are added to the data. Based on the data set enlarged this way, approximations of sales response functions are updated, for which new allocations are determined by quadratic programming and so on for several iterations (i.e., periods).

In each iteration the data are extended by allocations determined by an approximate optimal program and their corresponding sales. Therefore, the fit of the approximate response functions in a neighborhood close to the optimal solution gets more

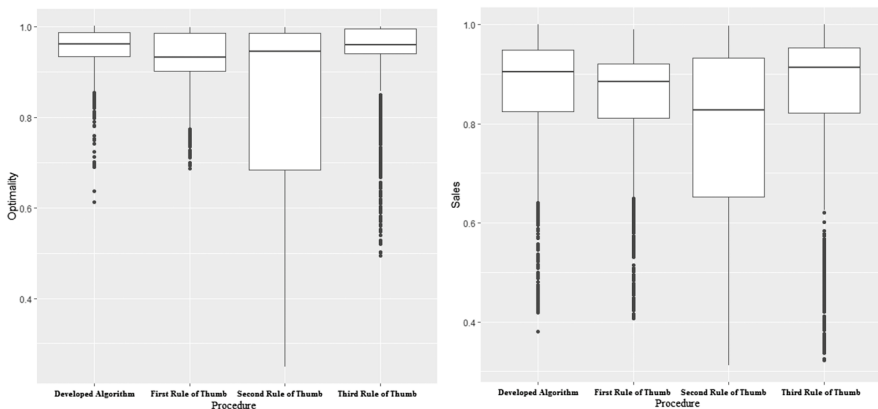


Fig. 5 Box-plot of optimality and sales

important. The conventional approach on the other hand tries to estimate response functions that also fit well for allocations far from the optimal solution.

Summing up, we suggest to prefer the developed algorithm over the investigated rules of thumb and the iterative algorithm of Albers (1997) if marketing allocation problems of the form shown in Sect. 2 must be solved and the underlying sales response functions are unknown and not deterministic. We justify this suggestion by the good performance of the algorithm demonstrated by the simulation study and its stability under changing conditions. Nonetheless, let us point to the trade-off between the better statistical performance of the developed algorithm and its higher complexity compared to rule of thumb 3. Please note that in certain constellations investigated in our simulation this rule of thumb performs equally well.

Modifying the algorithm to solve more general marketing decision problems also seems to be an interesting task of future research. One extended decision problem results if one or several sales functions may change suddenly, for example due to quality problems, successful advertising campaigns, entering another stage of the product life cycle or the addition or eliminations of allocation units. We suspect that under such circumstances the exploration stage will have to start once again. For situations with gradual change on the other hand an appropriate adjustment would be to delete older data points before each iteration. One could also investigate multi-variable generalizations. In one generalization, allocations affect sales of the same unit as well as sales of other units. Another more challenging generalization allows for marketing variables of different types. Examples of such variables are advertising and price or advertising and sales effort, where both variables have an effect on sales of different units. Finally, one might extend the problem by eliminating the monotony condition from the assumptions. In the case of supersaturation a response function has an actual maximum after which its values decrease when increasing the input. However, supersaturation is well outside the usual operation ranges for marketing instruments (Hanssens et al. 2001) as their optimal values are even lower than their values at saturation. Optimal values are further reduced due to the fact that the investigated decision problem deals with a scarce resource.

Appendix

A The developed algorithm in pseudocode

Notation: for two vectors a and b of the same dimension, we denote by $\langle a, b \rangle$ their dot product, and by $a * b$ the componentwise multiplication, i.e. the vector $(a_i b_i)_{i=1}^n$

Data: $x^1 = (x_1^1, \dots, x_n^1)$, x^2 start values, $0 \leq \beta \leq 1$ smoothing parameter, maxit: maximal number of iterations, B: resource, I_n : identity matrix of dimension n , exex=10: number of iterations in the exploration stage, \tilde{f} : multidimensional map consisting of the separate functions in each component, DF: array which will be filled with the data points of all functions

$$y^1 := \tilde{f}(x^1); y^2 := \tilde{f}(x^2);$$

while $i < exex$ **do**

 Test(x^1, x^2, y^1, y^2);

$$\varepsilon := \frac{\frac{y^2 - y^1}{x^2 - x^1}}{x^1}; \text{Epstest}(\varepsilon);$$

if $i \geq 2$ **then**

$$\quad \varepsilon_{sm} := \varepsilon \cdot (1 - \beta) + \varepsilon_{old} \cdot \beta$$

else

$$\quad \varepsilon_{sm} := \varepsilon;$$

end

$$up := \frac{y^1 * \varepsilon_{sm}}{(y^1, \varepsilon_{sm})} \cdot B;$$

$$\varepsilon_{old} := \varepsilon_{sm};$$

$$x^2 := x^1; y^2 := y^1;$$

 Test2(up, x^2);

$$x^1 := up; y^1 := \tilde{f}(x^1);$$

 add (x^1, y^1) to DF;

end

for $i = exex$ to maxit **do**

for $j = 1$ to n **do**

 Perform Regression $y_j = \gamma_{j,0} + \gamma_{j,1}x_j + \gamma_{j,2}(x_j)^2$

end

D: Diagonal Matrix containing values $-2\gamma_{j,2}$

A: n by $n+1$ -Matrix with -1 in the first column followed by I_n

d: Vector containing values $\gamma_{j,1}$

bv: Vector of length $n+1$ containing $(-B, 0, 0, \dots, 0)$

$up := \text{solve.QP}(D, d, A, bv)$

$$x^1 := up; y^1 := \tilde{f}(x^1);$$

 add (x^1, y^1) to DF;

end

Remark: The functions Test and Test2 check if the entries are too small or too close to each other, epstest projects elasticities into the Interval [0.01,0.5].

The function `solve.QP` solves the quadratic program as described in Goldfarb and Idnani (1983) with the notation from the R-Package “quadprog” by Berwin and Turlach (2013).

B Additional regression results

B.1 Comparison of nested models

While Table 7 has entries from which certain statements can be deduced, one should take these values with a grain of salt. The R^2 and adjusted R^2 of these main effect models amount for “Sales” to 0.5325 and 0.5318 and for “Optimality” to 0.4485 and 0.4478, respectively.

For $1 \leq k \leq 6$ we therefore define the k -fold interaction model (or short, k -fold model) as the linear multiple regression model that contains all interactions up to degree k , whereby the 1-fold interaction model is just the main effect model and the 6-fold interaction model is the full model. The R^2 and adjusted R^2 of these full models amount for “Sales” to 0.9294 and 0.9257 and for “Optimality” to 0.877 and 0.8705, respectively.

To be confident that the full model is better at estimating effects, we perform statistical tests. For $1 \leq k < k' \leq 6$, the k -fold model is nested within the k' -fold model. We therefore perform F-tests between the 6-fold model and each k -fold model nested within, for both dependent variables. For both dependent variables, the 6-fold model is superior to the 5-fold model at a significance level of 0.05. In both cases for all $1 \leq k < 5$, the 6-fold model is superior to the k -fold model at a significance level (far) below 0.001. Note that these tests are in fact necessary, since the F-tests of several nested models do not always behave in a transitive manner (see Cameron and Trivedi (2005) Sect. 7.2.7).

B.2 F-tests in the full models

We test the full model by the method explained in Gahler (2020) and use the terminology introduced there. In our case the procedure variable is the primary variable, and the other predictors are secondary variables.

The conditions necessary for using the method are as follows:

1. The multiple regression model contains categorical predictors only.
2. The regression includes all interaction terms of every possible degree.
3. There are no empty cells, i.e. the constructed design matrix is a generic design matrix in the sense of Definitions 1 and 4 in Gahler (2020).

The first two conditions are satisfied by construction, the third is satisfied since our design is balanced.

Furthermore, it only makes sense to use the method if the explanatory power of the full interaction model is higher than for all nested models. This condition is satisfied as well, as detailed in Appendix B.1.

We could use Theorems 2 and 3 from Gahler (2020) to determine the estimated cell means from the coefficients. Thankfully, we can skip this rather tedious procedure by directly applying the result from Section 2.3 in Gahler (2020), which states that the number of appearances (= the multiplicity) of a coefficient in the null hypothesis is zero if it does not describe the non-reference category one wishes to compare to, and equal to the product of the number of categories which according to the coefficient are in the reference category otherwise.

Therefore, when comparing the algorithm to the first rule of thumb, coefficients of an interaction term that does include the first rule of thumb (for example the coefficient of Proc₃ *Elas₂), have a multiplicity of zero. The multiplicity of all the other coefficients equals the product of the number of categories for those variables that assume a reference value. For example, the cell where Proc = ROT1, Form = multiplicative, Dist = 0.9, Elas = similar, Satu = varied and Budg = low has four variables (Form, Dist, Satu and Budg) in their reference categories. The multiplicity is therefore the product of their numbers of categories: $4 * 3 * 2 * 2 = 48$.

This way the multiplicity of each coefficient can be determined. The product of this vector of multiplicities and the OLS estimated coefficients gives the effect of the algorithm relative to another procedure. The significance of this relative effect can be obtained using these vectors as basis for an F-test.

C Supplementary material

C.1 Comparison to Albers (1997)

Albers (1997) investigates three different starting solutions, one of which consists of equal allocations. The other two starting solutions are either positively or negatively correlated with the optimal solution and differ from the equal allocation by 5%. In our case, the magnitude of the error terms, even in the case of $R^2 = 0.9$ (see Section 3.3), greatly exceeds a 5% boundary. Therefore we do not need different starting conditions and use equal starting allocations with $B/8$ for each unit.

The original table from Albers (1997), reproduced as Table 11, has more columns and slightly different values than Table 3. We explain why we use fewer columns due to a problem with the degrees of freedom for the determination of parameters. Albers specifies three restrictions:

1. The elasticity for an equal allocation corresponds to the elasticity value from Table 11.
2. The contribution (which is equivalent to sales in our decision problem, because contribution margins are constant) for an equal allocation must be the contribution value from Table 11.

Table 11 Original Table from Albers (1997)

No.	Products	Almost equal elasticities	Varianted elasticities	Equal profit contribution at equally allocated budgets	Varianted profit contribution at equally allocated budgets	Almost equal multipliers for saturation levels	Varianted multipliers for saturation levels
1	A	0.26	0.11	2,500,000	1,000,000	6.1	3
2	B	0.27	0.12	2,500,000	1,000,000	6.2	10
3	C	0.28	0.13	2,500,000	4,000,000	6.3	3
4	D	0.29	0.14	2,500,000	4,000,000	6.4	10
5	E	0.31	0.47	2,500,000	1,000,000	6.6	3
6	F	0.32	0.48	2,500,000	1,000,000	6.7	10
7	G	0.33	0.49	2,500,000	4,000,000	6.8	3
8	H	0.34	0.50	2,500,000	4,000,000	6.9	10
	Sum			20,000,000	20,000,000		

3. The function's saturation is defined as sales for an equal allocation multiplied by the saturation multipliers from Table 11.

For the ADBUDG function we have one additional restriction, $\phi_c < 1$ or $\phi_S > 1$, depending on the intended concave or S-shape.

Determining parameters requires that the number of restrictions equals the number of parameters of a sales response function. However, the number of restrictions is higher than the number of parameters for each considered sales response function (three parameters for the ADBUDG function and two parameters for each of the other sales response functions). That is why we eliminate the second restriction and accordingly the columns for equal and varied profit contributions in Table 11.

C.2 Visually testing S-shaped functions

We use visual control of the graphs of the S-shaped ADBUDG-functions to ensure that the functions do not behave badly. The functions shown here are scaled such that they only take values between zero and one, where one is the actual saturation, and the exponent is greater than one, i.e. there is an actual S-shape in the graph.

Functions with graphs such as the four in Figure 6 are rejected for the following reasons:

In the first image, the saturation is not reached within the domain of the function.

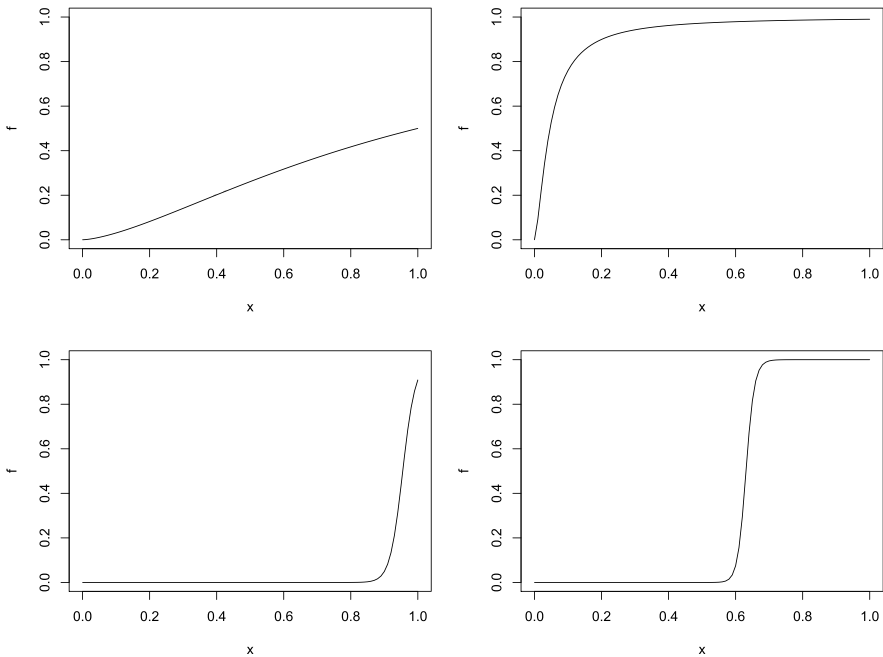


Fig. 6 Bad functions

In the second image, the saturation is reached at an unreasonably small input.

In the third image, the lower bound zero is not exceeded before a reasonable input.

In the fourth image, the non-constant part of the function is found on an unreasonably short interval of the input.

To make sure that these functions do not have such undesirable properties, we test them visually after generating parameter values.

D Dependent variable “Sales” at the high budget level

The dependent variable “Sales” increases at the high budget level if:

$$\frac{S_1}{S_1^{max}} > \frac{S_0}{S_0^{max}} \quad \text{or equivalently} \quad \frac{S_1}{S_0} > \frac{S_1^{max}}{S_0^{max}}$$

S_1, S_0 denote total sales at the high and low budget levels, respectively. S_1^{max} and S_0^{max} are the corresponding maximal attainable total sales. Substituting S_1 by $S_0 + \Delta S_0$ and S_1^{max} by $S_0^{max} + \Delta S_0^{max}$ with ΔS_0 and ΔS_0^{max} as additional total sales and additional maximal attainable sales at the high budget level and rewriting several times we obtain:

$$\begin{aligned} \frac{S_0 + \Delta S_0}{S_0} &> \frac{S_0^{max} + \Delta S_0^{max}}{S_0^{max}} \\ 1 + \frac{\Delta S_0}{S_0} &> 1 + \frac{\Delta S_0^{max}}{S_0^{max}} \\ \frac{\Delta S_0}{S_0} &> \frac{\Delta S_0^{max}}{S_0^{max}} \end{aligned}$$

The last expression shows that the dependent variable “Sales” increases if at the high budget level the relative increase of total sales is greater than the relative increase of maximal attainable sale. Please note that the total sales S_0 are less than the maximal attainable sales S_0^{max} . That is why the dependent variable “Sales” increases if additional total sales are higher or not much lower than additional maximal attainable sales at the high budget level. This condition is usually fulfilled because except for the S-shaped ADBUG function all investigated sales response functions are characterized by decreasing positive marginal effects.

Acknowledgements We thank two anonymous reviewers for their suggestions helping us to improve the paper.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Statements and declarations The authors declare that no funds, grants, or other support were received during the preparation of this manuscript. The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Albers S (1997) Rules for the allocation of a marketing budget across products or market segments. In: *Marketing: progress, prospects, perspectives, proceedings of the 26th EMAC conference, 20th–23rd May 1997*, pp 1–17
- Albers S (1998) Regeln für die Allokation eines marketing-budgets auf Produkte oder Marktsegmente. *Z Betriebswirt Forsch* 50(3):211–235
- Albers S, Mantrala MK, Sridhar S (2010) Personal selling elasticities: a meta-analysis. *J Market Res* 47(5):840–853
- Assmus G, Farley JU, Lehmann DR (1984) How advertising affects sales: Meta-analysis of econometric results. *J Market Res* 21:65–74
- Bates D, Watts D (1988) *Nonlinear regression analysis and its applications*. Wiley, Hoboken
- Berwin A, Turlach R (2013) quadprog: Functions to solve quadratic programming problems, R package version 1.5-5. <https://CRAN.R-project.org/package=quadprog>. Accessed 22 Dec 2021
- Beswick CA, Cravens DW (1977) A multistage decision model for salesforce management. *J Market Res* 14:135–144
- Cameron AC, Trivedi PK (2005) *Microeconometrics: methods and applications*. Cambridge University Press, Cambridge
- Cohen J (1988) *Statistical power analysis for the behavioral sciences*, 2nd edn. Erlbaum, Washington
- Cook R, Weisberg S (1999) *Applied regression including computing and graphics*. Wiley, Hoboken
- Fischer M, Albers S, Wagner N, Frie M (2011) Practice prize winner-dynamic marketing budget allocation across countries, products, and marketing activities. *Market Sci* 30(4):568–585
- Fischer M, Albers S, Wagner N (2013) Investigating the performance of budget allocation rules: a Monte Carlo study. *MSI reports*, pp 13–114
- Gahler D (2020) *Evaluating marketing allocation and pricing rules by Monte-Carlo simulation*. Dissertation, University of Regensburg
- Goldfarb D, Idnani A (1983) A numerically stable dual method for solving strictly convex quadratic programs. *Math Program* 27(1):1–33
- Gupta S, Steenburgh T (2008) Allocating marketing resources. In: Kerin RA, O'Regan R (eds) *Marketing mix decisions: new perspectives and practices*. American Marketing Association, Chicago, pp 90–105
- Hanssens DM, Parsons LJ, Schultz RL (2001) *Market response models: econometric and time-series research*. Kluwer Academic Publishers, Boston
- Jobson J (1991) *Applied multivariate data analysis*. Volume I: regression and experimental design. Springer, Berlin
- LaForge RW, Cravens DW (1985) Empirical and judgment-based sales-force decision models: a comparative analysis. *Decis Sci* 16(2):177–195

- Mantrala MK, Sinha P, Zoltners AA (1992) Impact of resource allocation rules on marketing investment-level decision and profitability. *J Market Res* 29(2):162
- March JG (1991) Exploration and exploitation in organizational learning. *Organ Sci* 2(1):71–87
- Richardson JTE (2011) Eta squared and partial eta squared as measures of effect size in educational research. *Educ Res Rev* 6(2):135–147
- Seber G, Wild C (1989) *Nonlinear regression*. Wiley, Hoboken
- Sethuraman R, Tellis GJ, Briesch RA (2011) How well does advertising work? generalizations from meta-analysis of brand advertising elasticities. *J Mark Res* 48(3):457–471
- Simon JL, Arndt J (1980) The shape of the advertising response function. *J Advert Res* 20(4):25–32
- Sinha P, Zoltners AA (2001) Sales-force decision models: insights from 25 years of implementation. *Interfaces* 31(3-supplement):S8–S44
- Tull D, Wood V, Duhan D, Gillpatrick T, Robertson K, Helgeson J (1986) Leveraged decision making in advertising: the flat maximum principle and its implications. *J Market Res* 23:25–32

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.