Christian Schlereth

Tanja Stepanchuk

Bernd Skiera*

Optimization and Analysis of the Profitability of Tariff Structures with Two-Part Tariffs

Christian Schlereth, Dr. Dr. (UA) Tanja Stepanchuk, Prof. Dr Bernd Skiera*, Goethe-University Frankfurt, Faculty of Business and Economics, Department of Marketing, Grueneburgplatz 1, 60323 Frankfurt, Germany, Phone +49-69-798-34638, Fax: +49-69-798-35001, emails: schlereth@wiwi.uni-frankfurt.de, stepanchuk@wiwi.uni-frankfurt.de, skiera@skiera.de

Abstract

Service providers often offer tariff structures with several two-part tariffs that consist of a fixed fee and a usage price, such that consumers may pick the tariff they prefer. Prices of tariffs have significant impacts on service providers' profit, because they simultaneously influence consumers' tariff choices and their usage. The number of tariffs also plays an important role, because more tariffs segment the market better but also increase the administrative burden and require more marketing effort. This article presents a mixed-integer nonlinear programming optimization problem to determine profit-maximizing tariffs; compares several heuristic search methods, in particular, the gradient method, stochastic search, and simulated annealing, to solve this problem; analyses the profitability of different tariff structures; and outlines the factors that drive differences in profitability across various tariff structures. The results show that especially for large samples of more than 100 consumers, simulated annealing performs best and deviates only 0.2% from the optimum. Structures with fewer two-part tariffs are generally sufficient, because additional two-part tariffs only negligibly increase service providers' profit.

Keywords: Marketing, pricing, two-part tariffs, mixed integer nonlinear programming

1 Introduction

By offering tariff structures with several kinds of two-part tariffs, service providers enable consumers to pick the one they prefer. Typically, a two-part tariff consists of a fixed fee and a usage price. For example, the German national railway company, Deutsche Bahn, offers a BahnCard that can be purchased for a fixed fee and entitles the passenger to travel at a discount price for a year. Second-class passengers may choose among the BahnCard 25, BahnCard 50, and BahnCard 100 at yearly fixed fees of 57 €, 230 €, and 3,800 €, respectively. The first two tariffs allow passengers to travel at 25% and 50% lower usage prices, whereas the BahnCard 100 allows free unlimited travel on the German national railway network. The prices of tariffs have significant impacts on the service providers' profit, because they simultaneously influence the consumers' tariff choices and their usage quantities. In the case of the German national railway company, for example, consumers likely travel more when the usage prices are lower, so lowering the fixed fee for the BahnCard 100 should make consumers switch to it and simultaneously increase their travelling, that is, their usage quantity.

In addition to setting prices, providers of such tariffs must determine the number and kind of tariffs in their tariff structures. Apple's iTunes service uses a pay-per-use tariff (e.g., 0.99 € per song), whereas customers of Napster can choose to buy individual songs (also at 0.99 € per song) or pay a monthly flat rate (e.g., 9.95 € per month for unlimited usage). Tariff structures with even more varied tariffs typically are offered by cellular phone operators or by service application providers, who offer software on demand as a service (SaaS); they consist of different two-part tariffs in addition to a pay-per-use tariff and a flat rate. More tariffs allow for better segmentation in the market, but they also increase the administrative burden and require greater marketing effort to explain the differences among the various tariffs (Hui, Yoo, and Tam 2007). Simpler tariff structures, such as a flat rate, may require less marketing effort, but their simplicity also provides less flexibility with regard to segmenting the market.

Empirical research pertaining to the optimal prices, numbers, and kinds of two-part tariffs is scarce; most studies use analytical models to examine the characteristics of optimal two-part tariffs. Leland and Meyer (1976) demonstrate that a profit-maximizing firm always prefers a two-part tariff over a uniform price (i.e., pay-per-use or flat rate). Murphy (1977) also notes that adding another two-part tariff to a tariff structure with different two-part tariffs always increases profits at a decreasing rate, as long as there is no tariff whose usage price equals the marginal cost. In particular, Murphy (1977) shows that offering 2 optional two-part tariffs[1] instead of just 1 tariff, increases

[1] The ”2” indicates the number of optional two-part tariffs, among which a consumer can choose.

profits by 2.88%, whereas 3 optional two-part instead of 2 tariffs increase profits only by an additional 0.65%.

In contrast with the suggestion offered by Maskin and Riley (1984)—namely, that it is optimal to design a two-part tariff for each consumer—various authors, including Huffman and Kahn (1998) and Iyengar and Lepper (2000), state that too many tariffs might confuse consumers. Thus, they may be more likely to purchase when they are confronted by a limited number of tariff options. Bhargava and Choudhary (2008), Hui, Yoo, and Tam (2007), and Sundararajan (2004) all examine the optimal number and kinds of tariffs for information goods (i.e., with near-zero variable costs) analytically, but they cover only a small subset of tariffs. Sundararajan (2004) also analyses the optimality of a pay-per-use tariff, a flat rate, or their combination and shows that sellers of information goods should offer the combination. Hui, Yoo, and Tam (2007) summarize findings that indicate it is optimal to offer as many versions as there are consumer types if the firm has no version-related cost. However, the firm's marginal benefit decreases as it adds more versions to its tariff structure. In addition, these authors note the need for research to analyse the factors that affect the optimal number of tariffs offered.

Several empirical studies also examine consumer behaviour in response to a menu of optional tariffs (Train, McFadden, and Ben-Akiva 1987; Danaher 2002; Iyengar, Jedidi, and Kohli 2008; Lambrecht, Seim, and Skiera 2007; Narayanan, Chintagunta, and Miravete 2007; Wolk and Skiera 2010). In particular, Train, McFadden, and Ben-Akiva (1987) show that households respond to a price adjustment by changing their calling patterns more than their calling plans. This result emphasizes the need to account for the interdependencies between prices and usage. Regarding the number of tariffs, Iyengar and Lepper (2000) offer empirical evidence that too few tariffs have limited appeal to a wide range of customers, but increasing the number of tariffs might demotivate consumers and prevent them from making a choice.

Furthermore, Iyengar, Jedidi, and Kohli (2008), Lambrecht, Seim, and Skiera (2007), and Narayanan, Chintagunta, and Miravete (2007) examine consumers' preferences for two-part tariffs and other tariff structures. Specifically, Lambrecht, Seim, and Skiera (2007) and Narayanan, Chintagunta, and Miravete (2007) develop discrete/continuous models of consumers' tariff choices among tariff structures and use transactional data to investigate empirically their usage behaviour, as well as the provider's revenue. However, they do not vary the tariff structures and perform sensitivity analysis only on current prices, without determining the profit-maximizing prices of tariffs. Consequently, they cannot compare the profitability of different tariff structures. Iyengar, Jedidi, and Kohli (2008) instead use a choice-based conjoint study (e.g., Natter and Feurstein 2002) to develop a model that allows for inferences of consumer behaviour based on tariff choices. They even determine the optimal prices for services, though only for a tariff structure with one tariff. Furthermore, they use a grid search method, which performs well in practice only when there are few consumers and a limited number of tariffs.

With this study, we present a mixed-integer nonlinear programming optimization problem designed to determine profit-maximizing tariffs, compare the performance of three heuristic search methods, analyse the profitability of different tariff structures, and outline the factors that drive differences in profitability across various tariff structures. Our approach differs from analytical models because it does not rely on typical assumptions, such as the single-crossing property (i.e., consumers' utility function can be sorted) or near-zero variable costs. Thus, it responds to the call from Hui, Yoo, and Tam (2007), who emphasize the importance of abandoning such assumptions because of their unknown impacts on the results. In contrast to previous empirical literature, we focus on optimizing tariffs, not estimating consumers’ demand functions. The remainder of our article is organized as follows: In Section 2, we present the tariff optimization problem and the heuristic search methods implemented to solve it. In Section 3, we compare the performance of these heuristic search methods. Section 4 represents a simulation study in which we examine the profitability of different tariff structures, and in Section 5, we summarize our findings.

2 Optimal Prices in Tariff Structures

To determine the optimal prices of tariffs in the tariff structure, we model consumers’ choice and usage behaviour and the corresponding profit for the service provider in Section 2.1. We then formulate the tariff optimization problem in Section 2.2 and present the different heuristic search methods for solving this problem in Section 2.3.

2.1 Modelling Consumers’ Choice and Usage Behaviour and the Service Provider’s Profit

We assume utility-maximizing consumers, each of whom choose one of the two-part tariffs or decides not to use the service at all. The consumers’ choice between offered tariffs from a set J depends on the prices of the tariffs, which influence the usage quantity for each tariff (Iyengar, Jedidi, and Kohli 2008; Lambrecht, Seim, and Skiera 2007; Train, McFadden, and Ben-Akiva 1987). To account appropriately for the interdependency between the tariff choice and usage quantity, we model the willingness-to-pay (WTP) of the i^{th} consumer as a function of the usage quantity q_i, which we refer to as the willingness-to-pay function (Wilson 1993). In line with previous literature (Brown and Sibley 1986; Maskin and Riley 1984), we assume that the consumer's WTP increases with the usage quantity q_i (equation (1)), but the corresponding marginal WTP decreases (equation (2)):

$$(1) \qquad \frac{dWTP_i(q_i)}{dq_i} \geq 0 \qquad (i \in I,\ q_i \geq 0), \text{ and}$$

$$(2) \qquad \frac{d^2WTP_i(q_i)}{dq_i^2} \leq 0 \qquad (i \in I,\ q_i \geq 0).$$

Such behaviour frequently gets captured by a quadratic functional form (e.g., Iyengar, Jedidi, and Kohli 2008; Lambrecht, Seim, and Skiera 2007). The willingness-to-pay function can be described by consumer-specific parameters, a_i, b_i, and c_i, and the usage quantity $q_{i,j}$ that the consumer chooses for the j^{th} tariff:

$$(3) \qquad WTP_{i,j}(q_{i,j}) = \begin{cases} a_i \cdot q_{i,j} - \frac{b_i}{2} \cdot q_{i,j}^2 + c_i & \text{if} \quad q_{i,j} \leq \frac{a_i}{b_i} \\ \frac{a_i^2}{2 \cdot b_i} + c_i & \text{if} \quad q_{i,j} > \frac{a_i}{b_i} \end{cases} \qquad (i \in I,\ j \in J,\ q_{i,j} \geq 0).$$

All parameters a_i, b_i, and c_i are assumed to be continuous and greater than or equal 0 to ensure a quasi-concave functional form (see Fig. 1). The parameter a_i is responsible for the increase in the WTP, the parameter b_i for the decrease in the marginal WTP, and the parameter c_i captures usage-independent WTP for the service. A usage-independent WTP greater than 0 is likely if the service provider does not charge for all services that a consumer might use. For example, in Europe and Australia consumers do not have to pay for incoming calls on their cellular phones. The usage-independent WTP thus captures the WTP for all free-of-charge services. Equation (3) models WTP as a function that increases with usage quantity until it reaches a saturation level for the quantity a_i/b_i, after which the consumer has no additional WTP for more consumption. The parameters a_i, b_i, and c_i of the willingness-to-pay function can be estimated with either transactional data using revealed preferences (e.g., Lambrecht, Seim, and Skiera 2007) or survey data using stated preferences (e.g., Iyengar, Jedidi, and Kohli 2008).

Knowledge about the willingness-to-pay functions permits us to model three consumer decisions: (1) the service purchase decision (to subscribe to a service or not?), (2) the tariff choice decision (which tariff to choose?), and (3) the usage quantity decision (how many units of the service to use?). The last decision is captured in the demand function, which we derive by taking the first derivative of equation (3) with respect to $q_{i,j}$ and substituting for the resulting marginal WTP with the usage price p_j:

$$(4) \qquad q_{i,j}(p_j) = \begin{cases} \frac{a_i}{b_i} - \frac{1}{b_i} \cdot p_j & \text{if } p_j \le a_i \\ 0 & \text{if } p_j > a_i \end{cases} \qquad (i \in I,\ j \in J,\ p_j \ge 0).$$

Note that equation (4) requires that $p_j \le a_i$, which ensures non-negative quantities as presumed in equation (3). In contrast, equation (3) requires that $q_{i,j} \le a_i / b_i$, to guarantee a non-negative marginal WTP for each unit increment, as presumed in equation (4).

The service purchase decision and tariff choice decision of the i^{th} consumer depends on the consumer surplus $CS_{i,j}$ for the j^{th} tariff, which is defined as the difference between the WTP for a usage quantity $q_{i,j}$ and the billing rate $R_{i,j}$:

$$(5) \qquad CS_{i,j}\left(q_{i,j}\right) = WTP_{i,j}(q_{i,j}) - R_{i,j}(q_{i,j}) \qquad (i \in I,\ j \in J,\ q_{i,j} \ge 0).$$

For a two-part tariff j that consists of a fixed fee F_j and a usage price p_j, the billing rate $R_{i,j}$ is:

$$(6) \qquad R_{i,j} = F_j + p_j \cdot q_{i,j}(p_j) \qquad (i \in I,\ j \in J,\ F_j,\ p_j \ge 0).$$

By substituting equations (3), (4), and (6) into equation (5) and rearranging, we obtain the following expression for consumer surplus:

$$(7) \qquad CS_{i,j}\left(F_j, p_j\right) = \begin{cases} \frac{\left(a_i - p_j\right)^2}{2 \cdot b_i} + c_i - F_j & \text{if } p_j \le a_i \\ c_i - F_j & \text{if } p_j > a_i \end{cases} \qquad (i \in I,\ j \in J,\ F_j,\ p_j \ge 0).$$

A consumer makes a service purchase decision and subscribes to a service only if the individual rationality constraint is fulfilled (e.g., Ehtamo, Berg, and Kitti 2010), meaning that he or she realizes a non-negative consumer surplus, $CS_{i,j}(F_j, p_j) \ge 0$. A consumer makes a tariff choice decision only if the incentive compatibility constraint is fulfilled (e.g., Ehtamo, Berg, and Kitti 2010), such that he or she chooses the tariff that maximizes his or her consumer surplus. When two or more tariffs provide the same consumer surplus, we assume the consumer chooses the tariff with the highest usage quantity (Brown and Sibley 1986).

Fig. 1 illustrates the three consumer decisions of the i^{th} consumer, whose willingness-to-pay function is specified by the parameters $a_i = 3.1$, $b_i = 0.1$, and $c_i = 0.1$. In this example, the tariff structure consists of 2 optional two-part tariffs (tariff 1 with $F_1 = 0$ and $p_1 = 1.8$; tariff 2 with $F_2 = 15$ and $p_2 = 0.8$). Under tariff 1 (2), the consumer uses 13 (23) units of the service, pays a billing rate of 23.40 (33.40), and receives a consumer surplus of 8.55 (11.55). The consumer therefore purchases (because of the positive consumer surplus) and chooses tariff 2 (because the consumer surplus for tariff 2 is higher than that for tariff 1).

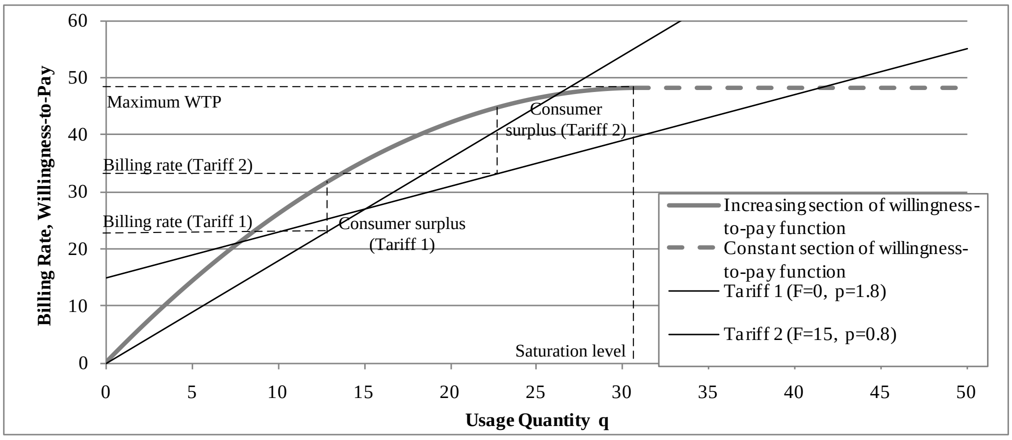

Fig. 1. Illustration of determination of consumer surplus.

2.2 Tariff Optimization Problem

We formulate the tariff optimization problem for a monopolistic service provider for a given number of tariffs in the tariff structure as follows:

(8) $$\max \pi(F_J, p_J) = \sum_{j\in J}\sum_{i\in I}\left[\left(p_j - k_v\right)\cdot q_{i,j}(p_j) + F_j\right]\cdot z_{i,j}$$

(9) s.t. $$CS_{i,j}\left(F_j, p_j\right)\cdot z_{i,j} \geq 0 \qquad (i\in I,\ j\in J),$$

(10) $$\left(CS_{i,j}\left(F_j, p_j\right) - CS_{i,j'}\left(F_{j'}, p_{j'}\right)\right)\cdot z_{i,j} \geq 0 \qquad (i\in I,\ j, j'\in J \ \wedge j > j'),$$

(11) $$\left(CS_{i,j}\left(F_j, p_j\right) - CS_{i,j'}\left(F_{j'}, p_{j'}\right)\right)\cdot z_{i,j} > 0 \qquad (i\in I,\ j, j'\in J \ \wedge j < j'),$$

(12) $$F_j \leq F_{j'} \qquad (j, j'\in J, j < j'),$$

(13) $$p_j \geq p_{j'} \qquad (j, j'\in J, j < j'),$$

(14) $$F_j, p_j \geq 0, \qquad (j\in J), \text{ and}$$

(15) $$z_{i,j} = \{0,1\} \qquad (i\in I, j\in J).$$

The decision variables in model (8)–(15) are F_j, p_j, and $z_{i,j}$. The objective function (8) maximizes profit π. Profit is the sum of profit contributions across all consumers, which consist of two components: the number of units consumed, $q_{i,j}(p_j)$, as specified in equation (4), and the margin per unit, which is the usage price minus the variable costs k_v.[2] The binary variable $z_{i,j}$ indicates the tariff choice ($z_{i,j} = 1$) of the i^{th} consumer among the available set J of tariffs and is specified by the constraints (9)–(11). The first constraint (9) expresses the individual rationality constraint (see Section 2.1), which is fulfilled if the consumer surplus (from equation (7)) times the binary variable $z_{i,j}$ is non-negative. Constraints (10) and (11) express the incentive compatibility constraints and distinguish whether the j^{th} tariff is larger than the j'^{th} tariff, ensuring that consumers choose the tariff with the highest usage quantity if two or more tariffs provide the same consumer surplus. Therefore, the consumer chooses not more than one tariff, so that $\sum_{j \in J} z_{i,j} \le 1, \forall i \in I$.

Constraints (12) and (13) order the tariffs according to the size of the fixed fee and avoid dominated tariffs, which are those with a higher fixed fee and a higher usage price than in one of the other tariffs. Constraint (14) specifies F_j and p_j as continuous, non-negative variables, whereas the binary variable $z_{i,j}$ in constraint (15) equals 1 if the i^{th} consumer chooses the j^{th} tariff and 0 otherwise.

For ease of illustration, for models (8)–(15), we consider a monopolistic service provider, though it is straightforward to account for static competition as well. In such a case, the individual rationality constraint (9) is replaced by constraint (16), which ensures that a consumer only chooses one of the tariffs offered by the firm if it yields a higher consumer surplus than any tariff among the set of tariffs J^C of all competitors:

$$(16) \qquad \text{s.t.} \quad \left(CS_{i,j}\left(F_j, p_j\right) - CS_{i,j^C}\left(F_{j^C}, p_{j^C}\right)\right) \cdot z_{i,j} \ge 0 \qquad (i \in I,\ j \in J,\ j^C \in J^C).$$

In addition, we also need to adjust the incentive compatibility constraints to appropriately capture the decision in the case in which the consumer surplus of the firm's and the best competitor's tariffs are equal:

$$(17) \qquad \left(CS_{i,j}\left(F_j, p_j\right) - CS_{i,j'}\left(F_{j'}, p_{j'}\right)\right) \cdot z_{i,j} \ge 0 \qquad (i \in I,\ j \in J,\ j' \in \{J; J^C\} \ \wedge j > j'), \text{ and}$$

$$(18) \qquad \left(CS_{i,j}\left(F_j, p_j\right) - CS_{i,j'}\left(F_{j'}, p_{j'}\right)\right) \cdot z_{i,j} > 0 \qquad (i \in I,\ j \in J,\ j' \in \{J; J^C\} \ \wedge j < j').$$

The tariff optimization problem contains both continuous variables (fixed fees F_j and usage prices p_j), as well as discrete variables ($z_{i,j}$), and thus represents the mixed-integer nonlinear programming (MINLP) class of problems. Nonlinearities appear in the objective function and

[2] Although not considered here, it is straightforward to incorporate other types of costs, such as fixed costs for each consumer served.

constraints, which make the class nonconvex and imply multiple local solutions (Floudas 1995). Determining a global solution for this nonconvex MINLP is NP hard (Murty and Kabadi 1987); even the global optimization of constrained nonlinear programming problems can be NP hard (Pardalos and Vavasis 1991). Furthermore, an increasing number of consumers or tariffs requires a large number of binary variables $z_{i,j}$, which induces a large combinatorial problem, whose complexity is characterized as NP complete (Nemhauser and Wolsey 1988). In the next section, we examine appropriate heuristic search methods to tackle this problem.

2.3 Heuristic Search Methods

No one heuristic search method works best for every kind of nonlinear optimization problem, so we examine the capabilities of three methods: the gradient method, stochastic search, and simulated annealing. The gradient method is a very well-known, powerful, general-purpose method, available within many commercial as well as non-commercial software packages (e.g., LINDO), and it uses the steepest ascent approach (Himmelblau 1972). Ehtamo, Berg, and Kitti (2010) apply the same method to solve a nonlinear pricing problem that involves two consumers with asymmetric and incomplete information. Stochastic search is more efficient for many NP-hard combinatorial problems, because it uses the multilevel single linkage procedure (Rinnooy-Kan and Timmer 1987), which is easy to implement and easy to parallel and thus can increase the number of practicable, solvable instances. However, stochastic search can become stuck in local optima and be misguided by the objective function. Finally, simulated annealing (Crama and Schyns 2003; Meiri and Zahavi 2006; Mika, Waligóra, and Weglarz 2005) was originally developed to solve large combinatorial optimization problems and exploit the theoretical asymptotic convergence of Markov chains (Laarhoven and Aarts 1987). It is more sophisticated to implement and requires many test runs to set the parameters of the algorithm properly; however, it worked well for pricing and portfolio selections problem that are also mixed-integer, nonlinear and NP-hard optimization problems (Mika, Waligóra, and Weglarz 2005; Crama and Schyns 2003).

To ensure better convergence among the heuristic search methods applied to solve this tariff optimization problem in a finite number of iterations, we reduce the search space of the problem as much as possible and set appropriate bounds for all the variables in the model. We use the non-negativity constraint (14) for F_j and p_j as the lower bounds. By considering the usage quantity $q_{i,j}$ of the i^{th} consumer under the j^{th} tariff in equation (4), we can derive the upper bound for the usage price:

(19) $$p_j^U = \max_{i \in I}\left(a_i\right) + \zeta \qquad (j \in J).$$

The small number ζ allows for tariffs with prices p_j that are too high to be chosen by any of the consumers. Thus, the usage price p_j of the j^{th} tariff in the tariff structure is constrained to the maximum WTP for the first usage quantity (i.e., parameter a_i) of all consumers, whereas the fixed fee F_j is bound by the highest maximum WTP:

$$(20) \qquad F_j^U = \max_{i \in I}\left(\frac{a_i^{\,2}}{2 \cdot b_i} + c_i\right) + \zeta \qquad (j \in J).$$

Each heuristic search method is initiated with the same starting points. Specifically, we set the starting points for the fixed fees F_j^0 and the usage prices p_j^0 in the two-dimensional search space, such that the distances between them are consistently equal. We then combine the fixed fees F_j and usage prices p_j of the j^{th} tariff, such that higher fixed fees correspond to lower usage prices ($F_j < F_{j'} \Rightarrow p_j > p_{j'}$). To ease the search process, we employ the individual rationality constraint (9) to compute the upper bound of the fixed fee F_j^U for a particular usage price p_j^0 of the j^{th} tariff, which attracts at least one consumer. Therefore, we set the expression for consumer surplus in equation (7) equal to 0 and solve it with respect to the fixed fee F_j:

$$(21) \qquad F_j^U(p_j^0) = \max_i \begin{cases} \dfrac{(a_i - p_j^0)^2}{2 \cdot b_i} + c_i, & \text{if } p_j^0 < a_i \\ c_i, & \text{if } p_j^0 \geq a_i \end{cases} \qquad (j \in J).$$

For all F_j^0, we ensure the values are less than $F_j^U(p_j^0)$.

The quadratic dimension of the tariff choice decision variable $z_{i,j}$ and the nonlinearities in the constraints (9)–(11) increase the complexity of the tariff optimization problem. However, the tariff optimization problem can still be solved effectively by applying the all-feasible approach to these constraints (e.g., Crama and Schyns 2003), which ensures that no time is lost by investigating infeasible solutions. This approach makes use of the convenient property of the tariff optimization problem that the $z_{i,j}$ of the i^{th} consumer is uniquely identified for any given tariffs. We vary only F_j and p_j in each of the iterations of the heuristic search and subsequently use constraints (9)–(11) to determine the tariff choice decisions $z_{i,j}$ for each consumer.

We subsequently detail the implementation of the three heuristic search methods. It is well known that the performance of heuristic search methods might be sensitive to the particular choice of the parameters (e.g., Kouvelis, Chiang, and Fitzsimmons 1992) and that the problem of fine-tuning parameters of each method becomes a complex optimization task, which is of combinatorial nature. For the purpose of this research, we tried to find a standard setting of the parameters that allows for determining the parameters automatically, i.e. without additional input from the user. Therefore, we studied the implication of parameter settings on the heuristic search methods' performance and computational time through extensive numerical experiments for a large set of problems. For each heuristic search method, we highlight the conclusion we got from our experiments and provide a guideline for carefully tuning parameters, which typically requires a trade-off between the quality of the results and the computational time. Further insights are provided by a sensitivity analysis that is based on a small subset of these experiments and which is decribed in more detail in the supplementary Web Appendix.

2.3.1 Gradient Method

The gradient method begins from a starting point and finds a directional derivative (i.e., the gradient) to proceed in the direction of the steepest ascent or descent. The search for the steepest derivative repeats several times from different starting points and returns the best local optimum. This method converges to the global optimum if the number of starting points approaches infinity.

We perform the optimization using the commercial LINDO-api, in which we apply a C# program in all iterations to evaluate the objective value. The implementation of the LINDO gradient method relies on the classical Newton Raphson algorithm. To ensure that the constraints in equations (12) and (13) hold, we include them using LINDO's array representation technique.

The gradient method requires the specification of two parameters: the step length for computing the derivatives using finite differences and the number of new starting points (drawn by Lindo). A high number of starting points allows for finding better solutions but requires more computational time. The relation between the increase in the number of starting points and the computational time is expected to be linearly. Because we aim to find good search results in subsequent sections, we set it to the highest possible number allowed by LINDO, i.e. 100. Choosing an adequate step length clearly is problem specific (e.g., Carpenter and Smithi 1975). Typical values are in the range between 5.0e-05 and 5.0e-09. The experimental results for the tariff optimization problem indicate that different step lengths have only minor influence on the performance of the gradient method. Therefore, we use LINDO's default value of 5.0E-07.

2.3.2 Stochastic Search

The stochastic search first determines a starting point that represents a starting combination of the tariffs' prices (fixed fees and usage prices). The algorithm constantly examines the search space and randomly generates N candidates (uniformly distributed set of neighbours of the starting point) by changing the fixed fees and usage prices of all tariffs. The candidate generation process determines the size of the next step Δx for each variable by drawing from a uniform, continuous distribution r (between –1 and 1) which is then multiplied by a fixed neighbourhood range NR, $\Delta x = r \cdot NR$. In all iterations, the algorithm computes the objective value for each candidate and moves to the "best candidate"—that is, the combination of tariff prices that ensures the highest profit for the service provider. The optimization procedure repeats until the termination criterion is fulfilled. The termination criterion is that either a maximum number of iterations is reached or the relative increase in profit compared to the previous iteration is lower than a specified termination tolerance threshold (i.e., a percentage between 0% and 100%). The difference between the stochastic search and the gradient method is that the former does not stay at the first optimum that is found and that might be locally optimal, but can find other solutions if the neighbourhood range is large enough.

We implement the stochastic search in C#. Assigning a high value for the number of candidates and the number of iterations as well as a low value for the termination tolerance requires more computational time but provides better solutions. When setting the optimal number of candidates N, the analyst should also consider the dimensionality of the problem, i.e., the number of tariff prices to be optimized. For two and three two-part tariffs, we found N=1,000 to be suitable, but recommend increasing this number for more tariffs. The neighbourhood range is problem specific and should depend on the expectations about the differences between local optima. In our study, we set it to 5% of the difference between the upper and lower bounds of the search space. The values of the termination tolerance and the maximum number of iterations must be carefully chosen in combination to ensure that not one of them dominates the termination of the search process. In the subsequent study, they are set to 0.01% and 350.

2.3.3 Simulated Annealing

In contrast with the gradient method or stochastic search, simulated annealing randomly accepts solutions with a worse objective functional value to overcome local solutions. Because it is based on the theory of Markov processes, the solution depends only marginally on the starting points and is rather flexible and robust (Aarts and Korst 1989). For a description of the algorithm, see Mahlke, Martin, and Moritz (2007) or Michalewicz and Fogel (2000). We also implement simulated annealing in C#. Subsequently, we detail the specification of the constraint handling, the definition of the neighbourhood structure, and the parameterization of simulated annealing, especially the adequate parameterization of the cooling schedule.

In each iteration, we first select randomly the tariff to be changed. Similar to the stochastic search, we subsequently use the step-size selection approach (Miki, Hiroyasu, and Fushimi 2003) and determine independently the change in the fixed fee and usage price through the equation $\Delta x = r \cdot NR$ in the fixed neighbourhood range NR, where r is a uniform continuous random variable, drawn between -1 and 1. We only accept the new candidate if it fulfils constraints (12) and (13) (Michalewicz and Fogel 2000). Otherwise, we reject the new candidate and draw another new candidate, until we find a valid one.

To design an adequate cooling schedule, we must carefully choose a starting point T_0 for the control parameter T, define a decrement function to be applied to the control parameter T, and finally restrict the number of iterations with a well-defined stop criterion (Dekkers and Aarts 1991; Laarhoven and Aarts 1987). The starting point should contain approximated information about the optimal objective value. Thus, we calculate T_0 according to the profit of the result of the gradient optimization run (configured with a small number of different starting points, which can be performed very quickly), times a constant t between 0 and 1. That is, $T_0 = t \cdot \pi^{gradient}(F_j, p_j)$. We decrease the control parameter according to the expression $T_{k+1} = \alpha \cdot T_k$, where α regulates the speed of the decrease of the control parameter to 0 and is a fixed number between 0 and 1.

Finally, we terminate the search after L iterations, which is the product of the problem dimension n and a constant L_0 by $L = n \cdot L_0$ (Dekkers and Aarts (1991)) . The sum of the number of fixed fees and usage prices determines the problem dimension, such that a tariff structure that consists of a pure flat rate has the dimension $n = 1 \cdot 1 = 1$, whereas a tariff structure that consists of 2 optional two-part tariffs has the dimension $n = 2 \cdot 2 = 4$. Simulated annealing terminates if either the iteration counter reaches the predetermined finite number L or there is fixed number of nonimproving, consecutive iterations—here 100,000 (Laarhoven and Aarts 1987).

The selection of the appropriate parameter values becomes especially crucial for simulated annealing, because its parameters are higher in number than for gradient method or stochastic search and they are sensitive to the solution space, the objective function, and the neighbourhood structure (e.g., Kouvelis, Chiang, and Fitzsimmons 1992). Despite of the vast literature, there are no generally accepted rules for selecting the best values (Pirlot 1996). To make simulated annealing generally applicable to different tariff optimization problem settings, we use initial calculation–obtained from the gradient method onfigured with only 10 starting points–to determine the control parameter of the cooling schedule. Furthermore, we set the maximum number of iterations L that terminate simulated annealing equal to the number of tariff prices in the optimization times a constant L_0, here 50,000.

The sensitivity analysis for the remaining parameters t, α, and NR shows that their variation only leads to marginal differences (see the supplementary Web Appendix). Substantial differences mainly exist for neighbourhood ranges that are too large (i.e., 10%), which also require additional computational time, such that we set NR to 5%. It is important to choose not too high values for the constant t to determine the starting point of the control parameter T_0. Values that are too small also may be problematic, because they cannot overcome gaps between local maxima, which is particular relevant in the case of fewer consumers, so that we employ t = 0.3. Yet, the value of the parameter α, which is responsible for the speed of the decrease, has a negligible effect and we set it to 0.9999. In our experience, the chosen parameters of all three heuristic search methods provide a proper trade-off between performance and computational time for a large set of problems; however, they cannot be considered optimal for every application and a careful parameter tuning is always advisable.

3 Comparison of Heuristic Search Methods' Performance

3.1 Comparison Structure

The comparison aims to identify the heuristic search method that finds the best solution with a varying number of consumers and two-part tariffs. We perform all computations on a 3.2 GHz Pentium Dual Core processor with 2 GB main memory. We randomly create several sets of consumers for a hypothetical service, each represented by the uniformly distributed parameters a_i, b_i, and c_i of their willingness-to-pay functions. We categorize data sets representing 5, 10, 15, and 30 consumers as small samples and data sets with 100, 300, and 5,000 consumers as large samples. For each data set, we determine the optimal prices for tariff structures that consist of 2 optional two-part tariffs and 3 optional two-part tariffs. To account for an increase in the number of two-part tariffs in the tariff structure, we keep the number of consumers constant at 100 and vary the number of two-part tariffs from 1 to 4. We repeat each optimization five times, resulting in 90 optimization runs for each heuristic search method.

The comparison benefits from knowledge about the profit in the global optimum because it allows for a better evaluation of the performance of each heuristic search method. Using BARON/GAMS (Ryoo and Sahinidis 1996), we succeed in determining the global optimum for up to 15 consumers and 2 optional two-part tariffs, which serves as a baseline for our small sample. For more consumers or tariffs, we approximate the global optimum by performing 50 optimization runs for the gradient method and simulated annealing and use the best result as our baseline. We also compute the approximations for 5, 10, and 15 consumers, for which we know the global optimum from BARON/GAMS. The resulting profit is on average only 0.134% lower, which we find acceptably close to the global optimum.

The criteria for comparing the performance of the three heuristic search methods are quality (of the best solution), robustness, and computational time. We assess the quality by comparing the highest profit in the five optimization runs with our baseline. Thus, quality is an indicator of the closeness to the optimal solution, expressed as a percentage value of the optimal solution (i.e., values close to 100% are desired). We assess the robustness of the results by comparing the differences between the best result and the worst result in the five optimization runs. In particular, we divide the differences by the baseline, such that values close to 0% are desired. Thus, robustness is an indicator of the stability of each heuristic search method. Finally, we measure computational time in seconds, taking the average over the five optimization runs.

3.2 Results of the Comparison

Table 1 reports the quality, robustness, and computational time for the small samples (5, 10, 15, and 30 consumers; 2 and 3 optional two-part tariffs). We highlight, in bold, the best heuristic search method for each criterion and case.

Table 1
Quality, robustness, and computational time for small samples

	2 Optional Two-Part Tariffs			3 Optional Two-Part Tariffs		
	Quality	Robustness	Comp. Time (sec)	Quality	Robustness	Comp. Time (Sec)
Gradient method						
5 consumers	**99.5%**	6.1%	7.07	**99.2%**	2.6%	9.89
10 consumers	**99.0%**	3.5%	9.56	**98.0%**	3.4%	11.82
15 consumers	**100.0%**	**2.6%**	9.51	**98.0%**	4.4%	10.83
30 consumers	99.4%	**0.8%**	12.23	96.5%	1.1%	16.85
Average	**99.5%**	3.3%	9.59	**97.9%**	2.9%	12.35
Stochastic search						
5 consumers	98.9%	**0.6%**	4.30	99.0%	**1.7%**	6.96
10 consumers	98.6%	**1.0%**	7.76	97.2%	**1.2%**	9.84
15 consumers	98.5%	3.3%	9.68	96.7%	**0.9%**	9.74
30 consumers	98.9%	1.7%	23.54	95.9%	**0.4%**	**13.74**
Average	98.7%	**1.7%**	11.32	97.2%	**1.1%**	10.07
Simulated annealing						
5 consumers	97.8%	10.2%	**1.29**	95.7%	9.8%	**2.89**
10 consumers	98.1%	6.8%	**2.43**	94.9%	2.8%	**4.83**
15 consumers	99.7%	2.0%	**3.52**	97.6%	2.4%	**7.16**
30 consumers	**99.9%**	4.3%	**7.08**	**99.7%**	3.3%	13.92
Average	98.9%	5.8%	**3.58**	97.0%	4.6%	**7.20**

In the cases in which the global optimum is known (left side for 5, 10, and 15 consumers), the quality is at least 97.8% for simulated annealing and 99.0% for the gradient method. Altogether, for small samples, the gradient method performs best in terms of quality (average 99.5% and 97.9%), whereas stochastic search performs best with regard to robustness (1.7% and 1.1%, compared with

3.3%, 2.9%, 5.8%, and 4.6%). The same pattern appears for the 3 optional two-part tariffs. For both tariff structures and both criteria, simulated annealing performs worst. The value of the robustness of simulated annealing is noticeably high, at 5.8% and 4.6%. However, when we compare the computational time, simulated annealing is the fastest among the three heuristic search methods in nearly every case.

Table 2 reports the results for large samples (100, 300, and 5,000 consumers). In contrast with the optimization results for small samples, simulated annealing clearly outperforms the gradient method and the stochastic search in terms of quality, robustness, and computational time for both tariff structures. Having more consumers leads to a smoother search region, which enables the heuristic search method to traverse among neighbouring local solutions better and discover the relevant search region for the global optimum. For the computational time measure, simulated annealing is about twice as fast as the other heuristic search methods. With simulated annealing, the quality and robustness improve most with a strong increase in the number of consumers; however, we also observe improvements with the gradient method and stochastic search. Fortunately, the computational time for all the heuristic search methods does not increase exponentially with an increase in the number of consumers.

Table 2
Quality, robustness, and computational time for large samples

	2 Optional Two-Part Tariffs			3 Optional Two-Part Tariffs		
	Quality	Robustness	Comp. Time (sec)	Quality	Robustness	Comp. Time (Sec)
Gradient method						
100 consumers	99.7%	1.0%	26.58	99.8%	2.5%	**31.17**
300 consumers	99.7%	1.0%	61.62	99.7%	2.5%	117.23
5,000 consumers	**99.9%**	1.1%	2954.17	99.7%	1.2%	4807.04
Average	99.8%	1.0%	1014.12	99.8%	2.0%	1,651.81
Stochastic search						
100 consumers	99.6%	1.4%	49.28	98.7%	0.7%	50.90
300 consumers	98.9%	1.0%	210.60	98.8%	1.2%	144.96
5,000 consumers	99.4%	1.1%	4353.46	99.2%	1.1%	4,208.80
Average	99.3%	1.2%	1537.78	98.9%	1.0%	1,468.22
Simulated annealing						
100 consumers	**100.0%**	**0.1%**	**22.13**	**100.0%**	**0.2%**	45.78
300 consumers	**100.0%**	**0.1%**	**60.60**	**100.0%**	**0.1%**	**119.65**
5,000 consumers	**99.9%**	**0.1%**	**813.56**	**99.9%**	**0.1%**	**1,654.94**
Average	**99.9%**	**0.1%**	**298.76**	**100.0%**	**0.1%**	**606.79**

Unexpectedly though, we observe a decline in the computational time for the stochastic search as the number of two-part tariffs increases from 2 to 3. This finding may explain the decrease in quality and robustness; as the number of two-part tariffs increases, stochastic search seems to find more or less the same local optimum and is not able to overcome it in subsequent optimization runs. Table 3 confirms this conclusion and reports the results for 100 consumers and different numbers of two-part tariffs.

Table 3
Quality, robustness, and computational time for 100 consumers and a different numbers of two-part tariffs

	Gradient Method			Stochastic Search			Simulated Annealing		
	Quality	Robust-ness	Comp. Time (sec)	Quality	Robust-ness	Comp. Time (sec)	Quality	Robust-ness	Comp. Time (Sec)
1 two-part tariff	100.0%	0.1%	15.55	99.6%	0.0%	10.00	**100.0%**	**0.2%**	**5.42**
2 Optional two-part tariffs	99.7%	1.0%	26.58	99.6%	1.4%	49.28	**100.0%**	**0.1%**	**22.13**
3 Optional two-part tariffs	99.8%	2.5%	**31.17**	98.7%	0.7%	50.90	**100.0%**	**0.2%**	45.78
4 Optional two-part tariffs	**99.9%**	4.8%	**32.40**	98.5%	0.4%	28.02	99.5%	**0.3%**	48.94
Average	**99.9%**	2.1%	**26.42**	99.1%	0.6%	34.55	99.8%	**0.2%**	30.57

For the stochastic search, quality decreases, whereas robustness improves. As the number of two-part tariffs increases, the stochastic search becomes stuck at a local optimum. In contrast, the quality of the gradient method (99.9%) only slightly decreases, though its robustness worsens. Robustness is best for simulated annealing (0.2%). Unlike an increase in the number of consumers, an increase in the number of two-part tariffs has a negative impact on the quality or performance of all heuristic search methods.

Recapitulating our analysis, the examined heuristic search methods can find fairly robust solutions to the tariff optimization problem in equations (8)–(15). For small samples, the gradient method performs best; however, we recommend several optimization runs to avoid poor solutions. In addition, an increase in the number of two-part tariffs decreases the quality and robustness of the heuristic search methods. Considering more consumers in the tariff optimization problem instead ensures more robust and better results. For large samples that contain more than 100 consumers, simulated annealing clearly outperforms the gradient method and stochastic search with regard to quality, robustness, and computational time.

4 Comparison of the Profitability of Different Tariff Structures

More tariffs allow for better segmentation of the market. Thus, they should increase the service provider's profits (Maskin and Riley 1984; Murphy 1977). Yet more tariffs also increase the administrative burden and require more marketing effort. These effects are not captured by our model, nor are they represented in previous models (e.g., Danaher 2002; Iyengar, Jedidi, and Kohli 2008;

Lambrecht, Seim, and Skiera 2007; Maskin and Riley 1984; Murphy 1977; Narayanan, Chintagunta, and Miravete 2007; Sundararajan 2004). Therefore, we analyse the profitability of different tariff structures in an attempt to gain a better understanding of the relative deviations in a service provider's profit with different tariff structures. This effort enables us to evaluate whether the increases in profit due to more "complex" tariff structures that contain more tariffs can be justified, despite the additional expenditures required to address the administrative burden and marketing efforts. In our simulation study, we use simulated annealing to determine the optimal prices for our tariff optimization problem (8)–(15). The simulation study enables us to vary systematically the environmental settings, including variables that influence profit, such as variable costs, and thereby study different settings and their impact on the differences in profitability across different tariff structures.

4.1 Comparison Structure

Table 4 shows tariff structures frequently used by prominent companies such as Apple, German Telekom, and Napster. To assess the profitability of these tariff structures, we distinguish among tariff structures that contain (1) tariffs with just one price component (tariff structures 1–2), (2) a combination of tariffs with just one price component (tariff structure 3), (3) a combination of tariffs with just one price component and two price components (tariff structure 4), and (4) between 1 and 4 optional two-part tariffs (tariff structures 5–8).

Table 4
Examples of tariff structures

Number	Tariff Structure	Example
1	Pay-per-use tariff	Apple iTunes charges 0.99 € per song
2	Flat rate	German Telekom DSL 6,000 for 14.99 € per month
3	Pay-per-use tariff + flat rate	Napster light for 0.99 € per song and Napster flat rate of 9.95 € for monthly and unlimited access
4	Pay-per-use tariff + flat rate + 1 two-part tariff	European cellular phone operator O2 offers a pay-per-use tariff of 0.25 €/min for domestic calls, a flat rate of 80 € monthly, or a two-part tariff: 20.00€ monthly + 0.19 €/min for domestic calls
5	1 two-part tariff	German electricity company offers Yellow Strom for a monthly fixed fee of 5.55 € and 0.18 € /KWh electricity
6 7 8	2 optional two-part tariffs 3 optional two-part tariffs 4 optional two-part tariffs	German national railway tickets with BahnCard: 25% discount on railway tickets for 57 € per year, 50% discount for 230 €, or 100% discount for 3,800 €

The tariff optimization problem (8)–(15) shows that the number of consumers, their WTP, and the variable costs represent the parameters of the simulation study. Therefore, we specify the environmental settings as follows: Heterogeneity among consumers is captured by the number of consumers (100 and 1,000), their WTP for the first unit of the service (low and high values for parameter a_i), their saturation level (low and high values for a_i/b_i), and their usage-independent

WTP (low and high values for parameter c_i). We randomly draw values from the symmetric triangular distributions displayed in *Table 5*, and we expect that greater heterogeneity benefits tariff structures that contain more tariffs. We also consider low and high variable costs; for example, information goods frequently have very low variable costs, so we use 0.01 € per unit for this kind of goods. High variable costs apply to goods such as cellular phone services, for which the providers need to buy capacity from network operators. We use a value of 1.00 € per unit to reflect these situations.

By systematically varying these factors, we consider 32 different environmental settings (*Table 5*). For each environmental setting and tariff structure, we solve the tariff optimization problem (8)–(15) three times using simulated annealing and take the best solution. To ensure the robustness of the results, we also conduct 20 replications of each setting. As a result, we consider 640 observations for each of the 8 tariff structures.

Table 5
Design of the simulation study

Factors	Notation	Numbers	Factor Levels
Heterogeneity in max marginal WTP	A_HETERO	2	High: $a_i = \Delta(0;2.5;5)$ Low: $a_i = \Delta(1.9;2.5;3.1)$
Heterogeneity in saturation level	SAT_HETERO	2	High: $a_i/b_i = \Delta(0;125;250)$ Low: $a_i/b_i = \Delta(93.8;125;156.2)$
Heterogeneity in usage-independent WTP	C_HETERO	2	High: $c_i = \Delta(0;2.5;5)$ Low: $c_i = \Delta(1.9;2.5;3.1)$
Number of consumers	N_CONSUM	2	High: $\lvert I \rvert = 1{,}000$ Low: $\lvert I \rvert = 100$
Variable costs	VAR_COSTS	2	High: k_v = 1.00 € Low: k_v = 0.01 €
Total number of environmental settings		$2^5 = 32$	
Number of replications		20	
Total number of observations for each tariff structure		640	
Total number of tariff structures		8	
Total number of observations		8·640 = 5,120	

To compare the differences in profitability across the different tariff structures, we compute the relative deviations from the profit that corresponds to the most profitable tariff structure. In our case, it is the tariff structure that contains 4 optional two-part tariffs. For each specified environmental setting s and tariff structure ts, we compute the relative deviations $\Delta\pi_{s,ts}$ in the service provider's profit from the profit earned with the optimal 4 optional two-part tariffs $\pi^*_{s,8}(F^*_{s,8,J}, p^*_{s,8,J})$, as follows:

(22) $$\Delta\pi_{s,ts} = \frac{\pi_{s,ts}(F_{s,ts,J}, p_{s,ts,J})}{\pi^*_{s,8}(F^*_{s,8,J}, p^*_{s,8,J})} - 1 \qquad (s \in S,\ ts \in TS).$$

We also want to determine how many consumers will be served under the different tariff structures and their average usage quantity, so we compute the share of consumers that chooses any one of the available tariffs in the tariff structure (labelled "consumer share") and the corresponding average usage quantity of those consumers. In line with equation (19), we compare those values with the corresponding values of the 4 optional two-part tariffs. This comparison enables us to analyse the relative deviations of the consumer share and total usage quantity for the different tariff structures.

4.2 Results of the Comparison

Our results, displayed in *Table 6*, show that many tariffs in a tariff structure increase profitability only slightly, such as +2.66% in the case of the 4 optional two-part tariffs compared with just 1 two-part tariff. This result matches Murphy (1977)'s analytical result that 3 optional two-part tariffs instead of just 1 two-part tariff in the tariff structure increases the profit by 3.53%.

Table 6
Relative deviation in profit, consumer share, and usage quantity for different tariff structures from the optimum

Number	Tariff Structure	Profit (Std Dev.)	Consumer Share	Usage Quantity	N
Optimum (4 optional two-part tariffs)		33,840.86	84.56%	34,724.74	640
		Relative deviations from optimum			
1	Pay-per-use tariff	-24.37% (6.33%)	18.29%	-28.26%	640
2	Flat rate	-34.60% (30.12%)	-39.26%	8.90%	640
3	Pay-per-use tariff +flat rate	-14.46% (9.12%)	18.29%	-7.59%	640
4	Pay-per-use + flat rate + two-part tariff	-5.03% (2.34%)	18.29%	-0.40%	640
5	1 Two-part tariff	-2.66% (2.06%)	-16.41%	-1.97%	640
6	2 Optional two-part tariffs	-0.67% (0.66%)	-4.51%	-1.97%	640
7	3 Optional two-part tariffs	-0.10% (0.33%)	-1.13%	-0.13%	640

Our results also show that tariffs with just one price component decrease profit significantly. In the case of a pay-per-use tariff, the service provider's profitability decreases by 24.37%. In contrast with 4 optional two-part tariffs, the number of consumers that use such a tariff is much higher (increase in consumer share), but the average usage quantity of each consumer is lower. Avoiding a fixed fee allows more consumers to use the tariff, but the rather high usage price limits the usage of each of those consumers.

In addition, we analyse the factors that drive these differences according to an analysis of variance (ANOVA) and a multiple classification analysis (MCA). *Table 7* reports the results.[3] Heterogeneity in WTP for the first unit explains most of the variance of the relative profit deviations (57.78%) for pay-per-use tariffs. In the case of high heterogeneity, the average profit deviations are lower (see the MCA values), which mirrors Oi (1971) results that show it is easiest to extract all the consumer surplus with a two-part tariff if consumers' WTP are rather homogeneous.

The greatest decrease in profitability occurs for the flat rate, when the tariff just contains a fixed fee; profitability falls on average by 34.60%. The ANOVA shows that the "variable costs" factor explains 80.78% of the variance (*Table 7*), and the MCA values indicate that high variable costs lead to high relative profit deviations (-27.05%). High variable costs frequently create situations in which the consumers' WTP is lower than variable costs, so the provider faces losses, which are responsible for the high profit deviations. The flat rate limits the number of consumers in the market (-39.26% in consumer share, *Table 6*), but each of those consumers has a much higher average usage quantity (+8.90%).

The combination of a pay-per-use tariff and flat rate leads to greater profitability than the pay-per-use tariff or flat rate alone (-14.46% compared with -24.37% and -34.60%, respectively). Nevertheless, the deviation in profitability for the case "pay-per-use tariff + flat rate" is still high compared with the 1 two-part tariff (-14.46% compared with -2.66%). The ANOVA analysis explains that the high variable costs remain responsible for most of those deviations (72.71% of the variance, -7.77% of profitability decreases; see *Table 7*).

---

[3] All tariff structures have a high total percentage of explained variance, with the exception of the 3 optional two-part tariffs. The deviation in profit is very small in the case of 3 optional two-part tariffs (on average, 0.10%), such that the ANOVA mainly detects randomness and no systematics in the deviations.

Table 7
Results of the analysis of variance and multiple classification analysis

		Pay-per-Use Tariff (N=640)		Flat Rate (N = 640)		Pay-per-Use+ Flat Rate (N = 640)		Pay-per-Use + Flat Rate + Two-Part Tariff (N = 640)		1 Two-Part Tariff (N = 640)		2 Optional Two-Part Tariffs (N = 640)		3 Optional Two-Part Tariffs (N = 640)	
Factor	Factor levels	Explained variance	MCA value	Explained variance	MCA value	Explained variance	MCA value	Explained variance	MCA value	Explained variance	MCA value	Explained variance	MCA value	Explained variance	MCA value
Average			-24.37%		-34.60%		-14.46%		-5.03%		-2.66%		-0.67%		-0.10%
A_	high	57.78%**	4.80%	6.70%**	7.79%	11.07%**	3.03%	1.13%**	0.25%	0.13%*	0.08%	0.72%**	-0.06%	0.22%	-0.02%
HETERO	low		-4.80%		-7.79%		-3.03%		-0.25%		-0.08%		0.06%		0.02%
SAT_	high	14.74%**	2.43%	7.62%**	-8.31%	1.24%**	-1.02%	9.81%**	-0.73%	54.25%**	-1.52%	35.95%**	-0.40%	3.96%**	-0.07%
HETERO	low		-2.43%		8.31%		1.02%		0.73%		1.52%		0.40%		0.07%
C_	high	0.05%*	0.15%	0.00%	0.10%	0.03%*	0.16%	0.42%**	0.15%	0.75%**	0.18%	0.57%**	0.05%	0.32%	0.02%
HETERO	low		-0.15%		-0.10%		-0.16%		-0.15%		-0.18%		-0.05%		-0.02%
N_	1,000	4.00%**	1.26%	0.02%**	-0.40%	0.70%**	0.76%	0.03%	0.04%	0.28%**	0.11%	0.02%	-0.01%	2.30%	0.05%
CUSTOM	100		-1.26%		0.40%		-0.76%		-0.04%		-0.11%		0.01%		-0.05%
VAR_	1.00 €	11.55%**	2.15%	80.78%**	-27.05%	72.71%**	-7.77%	76.40%**	-2.05%	1.94%**	-0.29%	3.05%**	-0.12%	0.04%	-0.01%
COSTS	0.01 €		-2.15%		27.05%		7.77%		2.05%		0.29%		0.12%		0.01%
% Factors		88.13%		95.12%		85.76%		87.78%		57.36%		40.32%		6.84%	

**1% level of significance. *5% level of significance.

Notes: % Factors row indicates the share of variance explained by all factors. MCA values are percentage points.

Furthermore, the tariff structure “pay-per-use + flat rate + two-part tariff” performs worse than the tariff structure that contains only 1 two-part tariff (-5.03% versus -2.66%). This result might seem surprising, in that it contradicts prior findings that more tariffs in the tariff structure increase the service provider’s profit (Maskin and Riley 1984). However, the pay-per-use tariff is only a special case of a two-part tariff, and its availability on the market limits the ability of the two-part tariff to subtract consumer surplus. In our case, the pay-per-use tariff attracts many customers who subscribe to it but do not use it intensively. This behaviour corresponds to usage patterns on iTunes: Consumers subscribe to iTunes services but only use it to organize their audio and video files and do not download files that are not free of charge. They thus realize positive consumer surplus with this pay-per-use tariff, which is difficult to increase further with a two-part tariff that has a significant fixed fee. At the same time, the profit of the pay-per-use tariff is limited because consumers do not use it intensively. Thus, the pay-per-use tariff restricts the possibilities of attracting consumers with an additional and more profitable two-part tariff.

5 Conclusion

Many service providers, including software as a services (SaaS), railway services, mobile telephone or Internet services, prescription drug plans, and car rental agencies, use two-part tariffs. Yet the determination of optimal tariff structures remains a challenging task, and prior literature provides only limited guidance. Our study contributes to efforts to answer essential questions: Which tariff structure is optimal to offer? What are the optimal prices of the tariffs in the tariff structure that maximize service providers’ profits? How many and what kinds of tariffs is it profitable to offer?

We propose a tariff optimization problem that maximizes the service providers’ profit by determining the optimal prices (fixed fees and usage prices) of the tariffs in a tariff structure, according to individual rationality and incentive compatibility constraints. Because our mixed-integer nonlinear programming tariff optimization problem belongs to the NP-hard class of problems, we compare the performance of heuristic search methods, such as the gradient method, stochastic search, and simulated annealing, in terms of their quality, robustness, and computational time. The results of the comparison suggest that the best heuristic search method depends on the sample size. For small samples, the gradient method provides the best results with several optimization runs. For large samples, simulated annealing performs best in a reasonable time. Its solutions are very robust, and it deviates only 0.2% from the solution with the highest profit.

To gain an understanding of whether the introduction of new tariffs in a tariff structure is profitable, we also conduct a simulation study and apply simulated annealing to solve the tariff optimization problem. In analysing how service providers’ profit changes with different tariff structures, we find that offering many two-part tariffs instead of a few increases profits only slightly. This result holds for a wide variety of environmental settings, because the influence of heterogeneity in consumers' WTP and providers' variable costs are rather moderate. Among those factors, heterogeneity in the saturation level of consumers' usage quantity is highest but still moderate. A requirement to offer a pay-per-use tariff in addition to two-part tariffs lowers the profitability substantially, because many consumers will be satisfied with the pay-per-use tariff, which implies lower usage, and do not choose other, more profitable tariffs. Furthermore, offering tariffs with just one price component is unprofitable; the high usage price of the pay-per-use tariff hinders significant usage quantities, and the flat-rate tariff is particularly bad if providers face high usage costs.

Additional studies might investigate the performance of other heuristic search methods, such as genetic algorithms (e.g., Pirlot 1996) or tabu search (e.g., Hedar and Fukushima 2006; Pirlot 1996). Extensive comparisons indicate results comparable to simulated annealing (e.g., Johnson et al. 1989; Johnson et al. 1991), but there is no guarantee of the same outcome for the tariff optimization problem considered herein. In addition, the model that we consider assumes no competition, which is relaxed by introducing static competition, i.e. competitors who do not react to changes in tariffs. Analytical studies with dynamic competition indicate that the quality distortion—here, the number of offered two-part tariffs—decreases with increased competition (Armstrong and Vickers 2001). Furthermore, in their analytical study, Shi et al. (2008) use game theory to point out that service providers likely will start to use additional tariff elements, such as usage allowances, to differentiate their offerings further. Extending our model to cope with such a dynamic competition would provide additional insights.

Acknowledgements

Parts of this manuscript were written during the stay at the Center of the Study of Choices (CenSoC) at the University of Technology in Sydney (UTS). The authors would like to thank Jordan Louviere and Christine Ebling for their support and Debora Malke for providing many helpful comments and suggestions. We would also like to thank the Editor and all the anonymous referees for their valuable remarks, which led to an improved version of this manuscript. This work was created as part of the project PREMIUM-Services, kindly funded by the German Federal Ministry of Education and Research (Project-ID 01IA08003C). We also acknowledge the financial support of the E-Finance Lab.

References

Aarts, E.H.L., J.H.M. Korst. Simulated Annealing and Boltzmann Machines - A Stochastic Approach to Combinatorial Optimization and Neural Computing. Great Britain: Chichester et al.; 1989.

Armstrong, M., J. Vickers. Competitive Price Discrimination. RAND Journal of Economics 2001; 32 (4); p. 579-605.

Bhargava, H.K., V. Choudhary. When is Versioning Optimal for Information Goods? Management Science 2008; 54 (5); p. 1029-1035.

Brown, S.J., D.S. Sibley. The Theory of Public Utility Pricing. Great Britain: Cambridge et al.; 1986.

Carpenter, W.C., E.A. Smithi. Computational Efficiency in Structural Optimization. Engineering Optimization 1975; 1 (3); p. 169-188.

Crama, Y., M. Schyns. Simulated Annealing for Complex Portfolio Selection Problems. European Journal of Operational Research 2003; 150 (3); p. 546-571.

Danaher, P.J. Optimal Pricing of New Subscription Services: Analysis of a Market Experiment. Marketing Science 2002; 21 (2); p. 119-138.

Dekkers, A., E. Aarts. Global Optimization and Simulated Annealing. Mathematical Programming 1991; 50 (3); p. 367-393.

Ehtamo, H., K. Berg, M. Kitti. An Adjustment Scheme for Nonlinear Pricing Problem with Two Buyers. European Journal of Operational Research 2010; 201 (1); p. 259-266.

Floudas, C.A. Nonlinear and Mixed-Integer Optimization: Fundamentals and Applications. Great Britain: Oxford; 1995.

Hedar, A.-R., M. Fukushima. Tabu Search Directed by Direct Search Methods for Nonlinear Global Optimization. European Journal of Operational Research 2006; 170 (2); p. 329-349.

Himmelblau, D.M. Applied Nonlinear Programming. USA: New York; 1972.

Huffman, C., B.E. Kahn. Variety for Sale: Mass Customization or Mass Confusion? Journal of Retailing 1998; 74 (4); p. 491–513.

Hui, W., B. Yoo, K.Y. Tam. The Optimal Number of Versions: Why Does Goldilocks Pricing Work for Information Goods? Journal of Management Information Systems 2007; 24 (3); p. 167–191.

Iyengar, R., K. Jedidi, R. Kohli. A Conjoint Approach to Multi-Part Pricing. Journal of Marketing Research 2008; 45 (2); p. 195-210.

Iyengar, S.S., M.R. Lepper. When Choice Is Demotivating: Can One Desire Too Much of a Good Thing? Journal of Personality and Social Psychology 2000; 79 (6); p. 995–1006.

Johnson, D.S., C.R. Aragon, L.A. McGeoch, C. Schevon. Optimization by Simulated Annealing: An Experimental Evaluation; Part I, Graph Partitioning. Operations Research 1989; 37 (6); p. 865-892.

—. Optimization by Simulated Annealing: An Experimental Evaluation; Part II, Graph Coloring and Number Partitioning. Operations Research 1991; 39 (3); p. 378-406.

Kouvelis, P., W.-C. Chiang, J. Fitzsimmons. Simulated Annealing for Machine Layout Problems in the Presence of Zoning Constraints. European Journal of Operational Research 1992; 57 (2); p. 203-223.

Laarhoven, P.J.M., E.H.L. Aarts. Simulated Annealing: Theory and Applications. Holland: Dordrecht; 1987.

Lambrecht, A., K. Seim, B. Skiera. Does Uncertainty Matter? Consumer Behavior under Three-Part Tariffs. Marketing Science 2007; 26 (5); p. 698-710.

Leland, H., R. Meyer. Monopoly Pricing Structures with Imperfect Discrimination. Bell Journal of Economics 1976; 7 (2); p. 449-462.

Mahlke, D., A. Martin, S. Moritz. A Simulated Annealing Algorithm for Transient Optimization in Gas Networks. Mathematical Methods of Operations Research 2007; 66 (1); p. 99-115.

Maskin, E., J. Riley. Monopoly with Incomplete Information. Rand Journal of Economics 1984; 15 (2); p. 171-196.

Meiri, R., J. Zahavi. Using Simulated Annealing to Optimize the Feature Selection Problem in Marketing Applications. European Journal of Operational Research 2006; 171 (3); p. 842-858.

Michalewicz, Z., D.B. Fogel. How to Solve It: Modern Heuristics. Germany: Berlin; 2000.

Mika, M., G. Waligóra, J. Weglarz. Simulated Annealing and Tabu Search for Multi-Mode Resource-Constrained Project Scheduling with Positive Discounted Cash Flows and Different Payment Models. European Journal of Operational Research 2005; 164 (3); p. 639-668.

Miki, M., T. Hiroyasu, T. Fushimi. 2003. Parallel Simulated Annealing with Adaptive Neighborhood Determined by GA. Paper presented to the IEEE International Conference on Systems, Man and Cybernetics, USA: Washington, D.C., 2003.

Murphy, M.M. Price Discrimination, Market Separation, and the Multi-Part Tariff. Economic Inquiry 1977; 15 (October); p. 587-599.

Murty, K.G., S.N. Kabadi. Some NP-Complete Problems in Quadratic and Nonlinear Programming. Mathematical Programming 1987; 39 (2); p. 117-129.

Narayanan, S., P. Chintagunta, E. Miravete. The Role of Self Selection, Usage Uncertainty and Learning in the Demand for Local Telephone Service. Quantitative Marketing and Economics 2007; 5 (1); p. 1-34.

Natter, M., M. Feurstein. Real World Performance of Choice-Based Conjoint Models. European Journal of Operational Research 2002; 137 (2); p. 448-458.

Nemhauser, G.L., L.A. Wolsey. Integer and Combinatorial Optimization USA: New York; 1988.

Oi, W.Y. A Disneyland Dilemma: Two-Part Tariffs for a Mickey Mouse Monopoly. Quarterly Journal of Economics 1971; 85 (1); p. 77-96.

Pardalos, P.M., S.A. Vavasis. Quadratic Programming with One Negative Eigenvalue is NP-hard. Journal of Global Optimization 1991; 1 (1); p. 15-22.

Pirlot, M. General Local Search Methods. European Journal of Operational Research 1996; 92 (3); p. 493-511.

Rinnooy-Kan, H.G., G.T. Timmer. Stochastic Global Optimization Methods Part II: Multilevel Methods. Mathematical Programming 1987; 39 (1); p. 57-78.

Ryoo, H.S., N.V. Sahinidis. A Branch-and-Reduce Approach to Global Optimization. Journal of Global Optimization 1996; 8 (2); p. 107-139.

Shi, B., Y. Song, C. Liu, Y. Luo. 2008. Competition and evolution of linear and two-part tariff. Paper presented to the Computers & Mathematics with Applications, 2008.

Sundararajan, A. Nonlinear Pricing of Information Goods. Management Science 2004; 50 (12); p. 1660-1673.

Train, K.E., D.L. McFadden, M. Ben-Akiva. The Demand for Local Telephone Service: A Fully Discrete Model of Residential Calling Patterns and Service Choices. Rand Journal of Economics 1987; 18 (1); p. 109-123.

Wilson, R. Nonlinear Pricing. New York et al.; 1993.

Wolk, A., B. Skiera. The Influence of Tariff-Specific Preferences on Tariff Choice and Usage. Goethe University Frankfurt; Business Research; forthcoming; 2010.