

Revenue Management under Customer Choice Behaviour with Cancellations and Overbooking

D.D. Sierag^{a,b,c,*}, G.M. Koole^b, R.D. van der Mei^{a,b}, J.I. van der Rest^c, B. Zwart^{a,b}

^a*CWI, Stochastics Department, Amsterdam, The Netherlands*

^b*VU University Amsterdam, Faculty of Exact Sciences, Amsterdam, The Netherlands*

^c*Hotelschool The Hague, Research Centre, The Hague, The Netherlands*

Abstract

In many application areas such as airlines and hotels a large number of bookings are typically cancelled. Explicitly taking into account cancellations creates an opportunity for increasing revenue. Motivated by this we propose a revenue management model based on Talluri and van Ryzin [27] that takes cancellations into account in addition to customer choice behaviour. Moreover, we consider overbooking limits as these are influenced by cancellations. We model the problem as a Markov decision process and propose three dynamic programming formulations to solve the problem, each appropriate in a different setting. We show that in certain settings the problem can be solved exactly using a tractable solution method. For other settings we propose tractable heuristics, since the problem faces the curse of dimensionality. Numerical results show that the heuristics perform almost as good as the exact solution. However, the model without cancellations can lead to a revenue loss of up to 20%. Lastly we provide a parameter estimation method based on Newman et alii [24]. This estimation method is fast and provides good parameter estimates. The combination of the model, the tractable and well-performing solution methods, and the parameter estimation method ensures that the model can efficiently be applied in practice.

Keywords: Revenue Management, Dynamic Pricing, Customer Choice Models, Markov Decision Processes, Cancellations, Overbooking

1. Introduction

Revenue management is the practice of deciding which products to sell to which customers at what price under capacity constraints. Classic examples can be found in the aviation industry and the hotel industry, where a limited number of aircraft seats and hotel rooms are available, respectively (e.g., Vinod [29], Clarke [7]). However, the revenue management models can be applied to a wide range of fields, including but not limited to the car rental industry (Geraghty and Johnson [11]), theatres (Langeveld [17]), golf courses (Kimes [15]), restaurants (Guerriero et alii [12]), and cruise ships (Maddah et alii [21]). All application areas have in common that a limited number of perishable goods is sold, whether it is a hotel room, a seat on an aircraft, a rental car, a seat in the theatre, or a timeslot on a golf course. The revenue management models that have been developed exploit this fact such that the price of a product is dynamically adjusted according to the demand and availability of the products. Littlewood [19] is considered a pioneer in the field. He introduced a capacity control model for two fare-classes on a single flight leg. The natural extension of this model to multiple fare-classes is intractable, but Belobaba [2] [3] [4] introduced two effective heuristics EMSR-a and EMSR-b that are still used in practice. From then several extensions have been made; see McGill and van Ryzin [22] for an overview.

All applications also have in common that customers choose a product among different products. For example, a customer who is looking for a hotel room may compare rooms of different hotels before making a decision. The seminal paper by Talluri and van Ryzin [27] incorporates customer choice behaviour in revenue management models. Their idea lay the foundation for a new theory in revenue management:

*Corresponding author. Postal address: Science Park 123, 1098 XG, Amsterdam, The Netherlands. Tel.: +31(0)20 592 4168. Email address: dirk@cwi.nl

20 customer choice models. Since then the theory of customer choice models has been enriched with among
others better solution methods (Strauss and Talluri [26], Meissner and Strauss [23]), network models to
take into account multiple night stays or multiple flight legs (Liu and van Ryzin [20]), and outstanding
methods to estimate the parameters of the model (Newman et alii [24]). However, an extension to
25 a model including cancellations has not been made. Modelling cancellations by taking into account
customer choice behaviour has been studied by Iliescu et alii [14], but the cancellation process is not
integrated with the decision process. Examples of studies on models that include cancellations but do
not take into account customer choice behaviour are Subramanian et alii [18], Aydin et alii [1], and
Bertsimas and Popescu [6]. [18] provide an analysis of a revenue management model with cancellations
and overbooking, but customer choice behaviour is not considered. Moreover, a satisfactory solution
30 method for their intractable problem is not provided. Recently, [1] proposed a revenue management
model with cancellations, but they also do not take customer choice behaviour into account. Erdely and
Topaloglu [10] and, more recently, Kunnumkal et alii [16] provide customer choice models that take into
account overbooking but no cancellations.

The main contribution of this paper is including cancellations and overbooking in a revenue management
35 model that also takes customer choice behaviour into account. Two other key contributions of this paper
are well-performing tractable solution methods and an accurate parameter estimation method with low
computation time. The combination of these three contributions makes the customer choice cancellation
model very suitable for practitioners.

Taking cancellations into account in the decision making process has a big impact on revenue. Numerical
40 results (see Section 5) show that policies that do not take cancellations into account can lead to a
substantial revenue loss of 20%. So far the customer choice models in existing literature do not take
into account cancellations. However, in practice it is common that for example seats on an aircraft are
cancelled, hotel rooms are cancelled, theatre tickets are cancelled, reservations for rental cars are cancelled,
and reservations for spots on a golf course are cancelled. Therefore we stress that effective revenue
45 management systems should take cancellation into account in the decision making process. Moreover,
overbooking makes more sense when taking cancellations into account.

We model the problem as a continuous-time Markov decision process and solve it using dynamic program-
ming. Some instances of the problem are too large to solve exactly because of the curse of dimensionality.
To overcome this problem we propose three tractable methods, each appropriate under different assump-
50 tion. First we show that if the cancellation rates are equal and linear with respect to the number of
reservations the problem can be reduced to a one-dimensional problem and thus it can be solved exactly.
Second we propose an efficient heuristic which is appropriate in the case that the cancellation rates are
linear with respect to the number of reservations, but not equal. Third we propose a heuristic that can
be applied to general cancellation rates, under the only assumption that cancellations occur independent
55 from each other. We also provide a heuristic which can be applied if we do not assume that all our
resources are identical. This occurs if we have multiple room types in a hotel or multiple seat classes in
a theatre. Numerical results show that the heuristics perform well.

To apply this model in practice it is essential to estimate the parameters of the model accurately. The
estimation of parameters of customer choice models can be challenging. A common attempt is to use
60 variations of the expectation maximisation algorithm. The basic version of Talluri and van Ryzin [27],
but also more sophisticated methods (e.g., Ryzin and Vulcano [28]) have the drawback of long compu-
tation time and bad parameter estimates. Recently Newman et alii [24] proposed a different parameter
estimation method that shows great potential. The parameter estimation method that we propose for
our customer choice cancellation model is based on this model. We estimate the parameters of the
65 continuous-time Markov chain, in contrast to most literature (e.g., van Ryzin and Vulcano [28]), which
estimate the parameters of the discretised Markov chain. Our estimation method has the advantage that
it has low computation time and it gives good estimates. The combination of the three effective solution
methods and the efficient estimation method ensures that the customer choice cancellation model is well
applicable in practice.

70 The remainder of this paper is organised as follows. First we describe the customer choice cancellation
model in Section 2. In this section we also describe several solution methods: a dynamic programming
formulation, which solves the problem exactly but is intractable; two heuristics, which are appropriate
in different settings; and a heuristic to solve the problem for a hotel with different room types. Then in
Section 4 describe methods to estimate the parameters of the model. Numerical results of the model are
75 presented in Section 5. Finally, in Section 7 we make some concluding remarks and propose topics for

further research.

2. Model Description

In this section we introduce the *customer choice cancellation model*. It is an extension of the *customer choice base model* by Talluri and van Ryzin [27], which does not consider cancellations. First we provide the model with no assumptions on the cancellation behaviour. Second we provide a first reformulation based on the assumption that cancellations only depend on the current number of reservations. Third we provide a reformulation based on the additional assumption that cancellation rates are linear. Finally we provide a reformulation based on the additional assumption that all cancellation rates are equal. In Section 3 we provide effective solution methods for the models.

Consider a hotel with $C \in \mathbb{N}$ identical rooms. We sell the rooms for customer check-in at a fixed future time-unit such that revenue is maximised. The customer check-in is in T time units. A common time unit is days, but some hotels sell rooms for a few hours to accommodate for short stays (e.g., at airport hotels, see Hopman [13]). We allow overbooking up to C_{\max} rooms. A *rate* or *fare product* j is a room in combination with a certain price r_j , also called the reward for product j , and certain conditions, e.g. cancellation conditions. Typical conditions in a hotel environment are refundable/non-refundable rooms, where a refundable room can be cancelled free of charge and a non-refundable not; or breakfast could be included or excluded from the fare product. Let N be the set of all fare products. We assume that there is a finite amount n of fare products, such that w.l.o.g. $N = \{1, \dots, n\}$. Let $c_j(t)$ be the costs if product j gets cancelled in time period t . Note that $c_j(t)$ depends on the cancellation conditions.

We assume that cancellations only depend on the current number of reservations. This model can also be used with time of booking dependent cancellations by including more fare products. We model this problem as a finite-horizon continuous time Markov decision process over T time units. Define the state space by

$$X := \left\{ x \in \mathbb{N}^n \mid x \geq 0, \sum_{j \in N} x_j \leq C_{\max} \right\}.$$

Each entry x_j of $x \in X$ corresponds to the pickup (number of reservations) for product j .

There is a penalty $q(x)$ involved if at the arrival day we are in state x and we have more reservations than capacity, i.e.,

$$\sum_{j \in N} x_j > C.$$

Customers arrive according to a Poisson process with parameter λ . Continuously in time we decide which subset $S \subset N$ of products we offer. S is called an *offer set*. If S is offered the probability that an arriving customer buys product j equals $P_j(S)$. Furthermore, the probability that an arriving customer buys nothing under offer set S equals $P_0(S)$. Finally, for all products j the cancellations of reservations for product j are exponentially distributed with parameter $\gamma_j(x)$. In accordance with Markov decision process literature we call a feasible solution π to our problem a *policy*. A policy π that optimises total revenue is called an *optimal policy*. See Figure 1 for a visualisation of this model.

Discretisation

To solve the customer choice cancellation model we consider the discretised Markov decision process. We divide time into T time periods, where the length of the intervals is such that the probability that more than one event occurs is very small. Therefore we assume that only one event occurs per time period. We denote λ as the probability that a customer arrives in a time period; and $\gamma_j(x)$ as the probability that product j is cancelled in state x . In each time period we decide which set S we offer. Note that time has to be scaled such that

$$\lambda + \max_{x \in X} \left\{ \sum_{j \in N} \gamma_j(x) \right\} \leq 1,$$

otherwise the probabilities are not well defined. This is possible as X is finite.

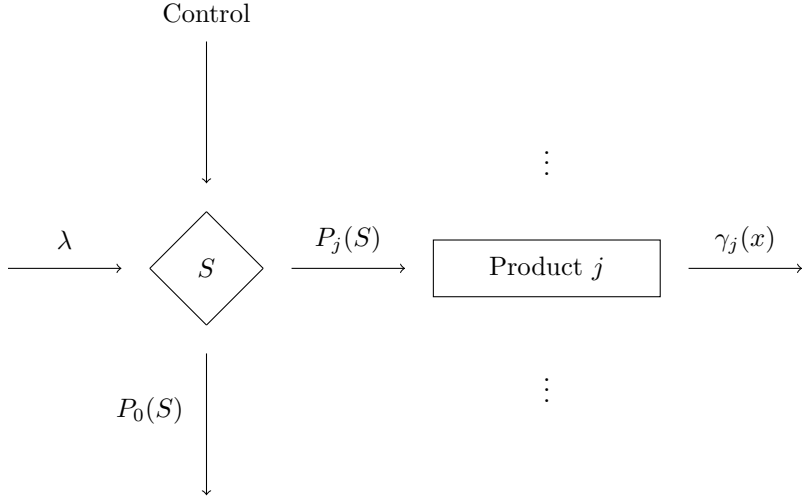


Figure 1: Visualisation of the customer choice cancellation model. The arrival process is Poisson distributed with parameter λ . Then we have control over the offer set S . Under this offer set an arriving customer buys product $j \in S$ with probability $P_j(S)$. With probability $P_0(S)$ the customer buys nothing. Finally, cancellations of product j follow an exponential distribution with parameter $\gamma_j(x)$.

105

Remark. It is possible to incorporate no-shows and walk-ins into this model. Subramanian et alii [18] do this by incorporating it in the penalty $q(x)$. However, it is also possible to incorporate the walk-in probability in the arrival probability and the no-show probability in the cancellation probability in the last time step(s). We choose the latter method.

110

Reformulation under Assumptions on Cancellation Rates

It is reasonable to assume that cancellations occur independent of each other. Assumption 2.1 below states that the probability that product j is cancelled only depends on the current number of reservations x_j for product j . The assumption is equivalent to Assumption 1 in Subramanian et alii [18].

115

Assumption 2.1. *Independence of Cancellations* (Assumption 1 in [18])

$$\gamma_j(x) = \gamma_j(x_j) \text{ for all } x \in X, \text{ for some function } \gamma_j: \mathbb{N} \rightarrow \mathbb{R}^+$$

To emphasize the fact that we work under the independence of cancellations assumption we write $\gamma_j(x_j)$ instead of $\gamma_j(x)$. Note that x is a vector and x_j is a scalar.

120

Assumption 2.2 below is even stronger than Assumption 2.1.

Assumption 2.2. *Linear Cancellation Rates*

1. $\gamma_j(x_j) = \gamma_j x_j$, for $\gamma_j \in \mathbb{R}^+$, for all $j \in N$
2. $q(x^1) = q(x^2)$ for all $x^1, x^2 \in X$ with $\sum_{j \in N} x_j^1 = \sum_{j \in N} x_j^2$

125

The second part of Assumption 2.2 states that the penalty for overbooking only depends on the total number of products that are sold and is independent from the types of products that are sold. This assumption is realistic because we are dealing with identical rooms, so relocation to a similar hotel induces the same costs. This is also a very convenient step if we only want to keep track of the total

number of reservations instead of the number of reservations per fare product. For $x \in X$, let y denote the total number of reservations, i.e.,

$$y = \sum_{j \in N} x_j.$$

Then under the linear cancellation rates assumption we write $q(y) = q(x)$ for all x with $\sum_{j \in N} x_j = y$.

Assumption 2.3 below states that the cancellation probability is independent from the product j , but it still depends on the number of reservations of product j .

130

Assumption 2.3. *Linear and Equal Cancellation Rates*

1. $\gamma_j(x) = \gamma x_j$, for $\gamma \in \mathbb{R}$ and for all $j \in N$
2. $q(x) = q(y)$ with $y = \sum_{j \in N} x_j$

135

Under the linear and equal cancellation rates assumption the state space of the problem is significantly reduced, as is shown in Theorem 3.1 in Section 3. However, this assumption is not realistic. Different fare products are likely to have different cancellation probabilities. For example, a refundable room is more likely to get cancelled than a non-refundable room.

Illustrative Example

140

An illustration of the model can be found in Example 2.1 below. This example is used throughout the paper.

Example 2.1. ¹

Consider a hotel with $C \in \mathbb{N}$ rooms which offers three fare products. The reward for these products is given by

$$r = (r_1, r_2, r_3) = (160, 100, 90).$$

Both fare product 2 and 3 need to be purchased at least 21 days ahead. Furthermore, the cancellation conditions are as follows: if a product of type j is cancelled, the customer receives a refund of

$$c_j = \begin{cases} r_1 & \text{if } j = 1, \\ \frac{1}{2}r_2 & \text{if } j = 2, \\ 0 & \text{if } j = 3. \end{cases}$$

In other words, product 1 is refundable, product 2 only refunds half the price, and product 3 is non-refundable. Assume that overbooking is allowed but is bounded by 20% of the capacity, such that $C_{\max} = \lfloor C \frac{6}{5} \rfloor$. If we have $\sum_{j \in N} x_j - C > 0$ overbookings, then $\sum_{j \in N} x_j - C$ reservations are chosen at random and relocated to another hotel. The costs for relocating a customer is 170. Hence the penalty for overbooking $q(x)$ at the arrival day in state x equals

$$\begin{cases} 170 \left(\sum_{j \in N} x_j - C \right) & \text{if } \sum_{j \in N} x_j - C > 0, \\ 0 & \text{otherwise.} \end{cases}$$

¹This example is based on Example 1 from Talluri and van Ryzin [27]. We scaled the rewards to be representative values for hotels. Moreover, we added cancellation conditions, cancellation probabilities, and an overbooking policy.

S	$P_1(S)$	$P_2(S)$	$P_3(S)$	$P_0(S)$
\emptyset	0	0	0	1
$\{1\}$	0.3	0	0	0.7
$\{2\}$	0	0.4	0	0.6
$\{3\}$	0	0	0.5	0.5
$\{1, 2\}$	0.1	0.6	0	0.3
$\{1, 3\}$	0.3	0	0.5	0.2
$\{2, 3\}$	0	0.4	0.5	0.1
$\{1, 2, 3\}$	0.1	0.4	0.5	0

Table 1: Purchase probabilities under different offer sets.

The hotel wants to sell the rooms in a period of T days. The purchase probabilities $P_j(S)$ are given in Table 1. We assume that the purchase probabilities are the same for all time periods. The probability that a reservation for product j is cancelled is

$$\gamma_j(x) = \gamma_j x_j,$$

with $\gamma_j \in [0, 1]$ for all j .

145

3. Solution Methods

In this section we provide solution methods for the models derived in Section 2. First we provide an exact solution method based on dynamic programming for the model with no restrictions on cancellation behaviour. Unfortunately this problem faces the curse of dimensionality. To overcome this we provide 150 heuristics for each model described in Section 2. In particular we provide a tractable solution method that solves the model under the linear and equal cancellations assumption 2.3 exactly. Finally we provide a heuristic that is applicable to a hotel with multiple room types.

3.1. General Solution Method: No Assumptions on Cancellation Behaviour

Let $V_t(x)$ be the maximal expected revenue from time t to the arrival day. To solve this problem we use 155 the following Bellman equation:

$$V_t(x) = \begin{cases} \max_{S \subset N} \left\{ \lambda \sum_{j \in S} P_j(S) (r_j + V_{t-1}(x + e_j)) \right. \\ \left. + \sum_{j \in N} \gamma_j(x) (-c_j(t) + V_{t-1}(x - e_j)) \right. \\ \left. + \left(1 - \lambda \sum_{j \in S} P_j(S) - \sum_{j \in N} \gamma_j(x) \right) V_{t-1}(x) \right\} & \text{if } t > 0, \\ q(x) & \text{if } t = 0, \end{cases} \quad (1)$$

for all $x \in X$ with where $e_j \in \mathbb{R}^n$ is the j -th unit vector. With dynamic programming we can find an optimal strategy. Unfortunately, this solution methods bears the burden of the curse of dimensionality. The size of the state space X is $O(C^n)$, so it increases exponentially if the number of fare products increases. For a small number of fare products the exact solution can be computed, like in Example 160 2.1. Obviously, in practice the number of fare products is larger than in our small example. Hence this problem can not be solved for all practical purposes. To overcome this problem we propose several options. First we show in Theorem 3.1 that under certain conditions the problem can be solved exactly. Second we propose several heuristics in case the conditions of Theorem 3.1 are not satisfied. It will turn out that each heuristic has its own advantages and disadvantages and is appropriate in a different setting.

165

Remark. In contrast to the dynamic programming formulation of Subramanian et alii [18] the structure of an optimal policy is not clear beforehand. The main difference between the models is that we also consider customer choice behaviour. However, if the purchase probabilities follow an independent model, as described by Talluri and van Ryzin [27], we do get similar admission control policies. More precisely, if the purchase probabilities are given by

$$P_j(S) = \begin{cases} p_j & \text{if } j \in S, \\ 0 & \text{otherwise,} \end{cases}$$

then an optimal policy of Equation (1) will have the form

$$\text{accept a fare-product } j \text{ request if and only if } r_j + V_{t-1}(x + e_j) > V_{t-1}(x).$$

In the remainder of this paper we focus on dependent demand models.

170 Reformulation

The customer choice cancellation model can be reformulated such that the costs are incorporated in the rewards. We do this by adding the expected costs caused by cancellations to the value function. This reasoning is borrowed from Subramanian et alii [18], who use it in their revenue management model which does not take customer choice behaviour into account. We show in Theorem 3.1 below that under certain conditions the customer choice cancellation model is equivalent to a one-dimensional problem.

Let $H(x, t)$ be the expected costs from cancellations from the reservations in state (x, t) . The costs are fixed because we do not have control over the reservations. Then $H(x, t)$ is given by the recursive formula

$$H(x, t) = \begin{cases} \sum_{j \in N} \gamma_j(x)(e_j(t) + H(x - e_j, t - 1)) + \left(1 - \sum_{j \in N} \gamma_j(x)\right) H(x, t - 1) & \text{if } t > 0, \\ 0 & \text{if } t = 0. \end{cases}$$

Define the value function $\tilde{V}_t(x)$ by

$$\tilde{V}_t(x) := V_t(x) + H(x, t). \quad (2)$$

The value function $\tilde{V}_t(x)$ can be interpreted as the maximal expected revenue to go from reservations without costs from current cancellations. These costs are unavoidable and have no influence on future profits. Note that cancellations still occur, but it will turn out that this reformulation alters the cancellation term in a preferable way.

180 Define $\Delta H_j(x, t) := H(x, t - 1) - H(x + e_j, t - 1)$ for all $x \in X$ and for all $t > 0$. Then we can rewrite $\tilde{V}_t(x)$ as follows:

$$\tilde{V}_t(x) = \begin{cases} \max_{S \subset N} \left\{ \lambda \sum_{j \in S} P_j(S)(r_j + \Delta H_j(x, t) + \tilde{V}_{t-1}(x + e_j)) \right. \\ \quad \left. + \sum_{j \in N} \gamma_j(x)(\tilde{V}_{t-1}(x - e_j)) \right. \\ \quad \left. + \left(1 - \lambda \sum_{j \in S} P_j(S) - \sum_{j \in N} \gamma_j(x)\right) \tilde{V}_{t-1}(x) \right\}, \\ q(x) & \text{if } t = 0. \end{cases} \quad (3)$$

Equation (3) is equivalent to Equation (1) in the sense that it leads to the same policy. This follows from Equation (2) and the fact that the policy to go has no influence on $H(x, t)$.

Efficient sets

185 The customer choice base model has some elegant properties described in Talluri and van Ryzin [27] which reduce the action space and therefore the computation time. Unfortunately, these properties do not hold in the customer choice cancellation model. To show this, we introduce the concept of *efficient sets* in Definition 3.1 below (Definition 1 in [27]). We will show later on that some heuristics do satisfy some of the elegant properties described in [27].

190

Definition 3.1. For each $S \subset N$, $x \in X$ and $t \in \{0, \dots, T\}$ define $Q(S)$ and $R(S)$ by

$$Q(S) := \sum_{j \in S} P_j(S), \quad R(S, x, t) := \sum_{j \in S} P_j(S)(r_j + \Delta H_j(x, t)).$$

A set S^* is called *inefficient* at (x, t) , with $x \in X$ and $t \in \{0, \dots, T\}$, if there exist $\alpha(S) > 0$ for all $S \subset N$ with $\sum_{S \subset N} \alpha(S) = 1$ such that both of the following inequalities hold:

$$\begin{aligned} Q(S^*) &\geq \sum_{S \subset N} \alpha(S) Q(S), \\ R(S^*, x, t) &< \sum_{S \subset N} \alpha(S) R(S, x, t). \end{aligned}$$

Otherwise S^* is called *efficient* at (x, t) .

The intuition behind efficient sets is as follows. An offer set $S^* \subset N$ is efficient if there is no randomisation of other offer sets $S \subset N$ such that the expected reward is strictly greater than $R(S^*, x, t)$ and the probability of a purchase is at most $Q(S^*)$. Efficient sets are on the efficient frontier of the trade-off 195 between the expected revenue $R(S^*, x, t)$ and the probability of purchase $Q(S^*)$.

Efficient sets can be identified using the *largest marginal revenue procedure* described by Talluri and van Ryzin [27], which is the following: Initialise with the first efficient set $S_0 = \emptyset$. Then iteratively proceed as follows: Suppose S_i is the i -th efficient set found by the procedure. Then the $(i + 1)$ -th efficient set S_{i+1} is equal to

$$S_{i+1} = \underset{S: \substack{Q(S) \geq Q(S_i) \\ R(S, x, t) \geq R(S_i, x, t)}}{\arg \max} \frac{R(S, x, t) - R(S_i, x, t)}{Q(S) - Q(S_i)}.$$

The complexity of this procedure is $O(m2^n)$, where m is the number of efficient sets, which is at most 2^n . For large instances of n this procedure is computationally intractable, so then we need to use heuristics. One heuristic (proposed by [27]) is to use only a subset of all efficient sets and then use the largest 200 marginal revenue procedure. By looking at the specific instance we can rule out certain offer-sets. For example, if a hotel offers the same room under the same conditions at two different prices €80 and €120 at the same time, the price-sensitive customer would almost surely purchase the room for €80.

In the base model no cancellations exist, such that $\Delta H_j(x, t) = 0$ for all (x, t) . Hence, the set of efficient sets is the same for each (x, t) . In contrast, for the cancellation model the set of efficient sets is not 205 necessarily the same for all (x, t) , as is shown in Example 3.1.

Example 3.1. Consider Example 2.1. Let $C = 5$, $T = 10$, $\gamma = (0.05, 0.0025, 0.001)$. Consider the states (x, t_1) and (x, t_2) , with $t_1 = 1$, $t_2 = 10$, and $x = (0, 0, 0)$. The efficient sets of state (x, t_1) and (x, t_2) are given in Table 2. Offer set $\{1, 3\}$ is efficient in state (x, t_1) but not in state (x, t_2) , and offer set $\{1, 2\}$ is not efficient in state (x, t_1) but it is efficient in state (x, t_2) . 210

An important property of the base model is that an inefficient set is never an optimal solution (see Proposition 2 in [27]). Unfortunately, this property generally does not hold for the cancellation model. To this end consider Example 3.2.

S	$Q(S)$	$R(S, x, t_1)$	$R(S, x, t_1)$	Efficient set for (x, t_1)	Efficient set for (x, t_2)
\emptyset	0	0	0	Yes	Yes
{1}	0.3	48	30	Yes	Yes
{2}	0.4	40	40	No	No
{3}	0.5	45	45	No	No
{1, 2}	0.7	76	69	No	Yes
{1, 3}	0.8	93	75	Yes	No
{2, 3}	0.9	85	85	No	No
{1, 2, 3}	1	101	95	Yes	Yes

Table 2: Efficient sets for Example 2.1.

215

Example 3.2. Consider Example 3.1. In state $x = (0, 0, 4)$ at time T offer set $\{1, 3\}$ is an optimal offer set. However, offer set $\{1, 2\}$ is not efficient in state x at time T , as was shown in Example 3.1.

220 3.2. Tractable Exact Solution: Linear and Equal Cancellations (LEC) Assumption

The model under the linear and equal cancellation Assumption 2.3 turns out to have the elegant property that it can tractably be solved exactly. In order to see this we use some derivations similar to Subramanian et alii [18].

225 First consider the model under the independence of cancellations Assumption 2.1. The statement in Lemma 3.1 is taken from Subramanian et alii [18].

Lemma 3.1. (Lemma 2 in [18])

Under Assumption 2.1 we have that

$$H(x, t) = \sum_{j \in N} H_j(x_j, t),$$

where the function $H_j(x_j, t)$, $1 \leq j \leq n$, satisfy the following recursive equations:

$$H_j(x_j, t) = \begin{cases} (1 - \gamma_j(x_j))H_j(x_j, t - 1) & \text{if } t > 0 \\ +\gamma_j(x_j)(c_j(t) + H_j(x_j - 1, t - 1)) & \text{if } t = 0. \end{cases}$$

for all $0 \leq t \leq T$.

Define $\Delta H_j(x_j, t)$ by

$$\Delta H_j(x_j, t) := H_j(x_j + 1, t - 1) - H_j(x_j, t - 1).$$

230 Lemma 3.2 below states that under the independence of cancellations assumption Equation (3) can be simplified.

Lemma 3.2. Under Assumption 2.1 Equation (3) can be rewritten to

$$\tilde{V}_t(x) = \begin{cases} \max_{S \subset N} \left\{ \lambda \sum_{j \in S} P_j(S) (r_j - \Delta H_j(x, t) + \tilde{V}_{t-1}(x + e_j)) \right. \\ \quad \left. + \sum_{j \in N} \gamma_j(x_j) (\tilde{V}_{t-1}(x - e_j)) \right. \\ \quad \left. + \left(1 - \lambda \sum_{j \in S} P_j(S) - \sum_{j \in N} \gamma_j(x_j) \right) \tilde{V}_{t-1}(x) \right\} & \text{if } t > 0, \\ q(x) & \text{if } t = 0. \end{cases}$$

where $\Delta H_j(x_j, t)$ satisfies the following recursive formula

$$\Delta H_j(x_j, t) = \begin{cases} (\gamma_j(x_j + 1) - \gamma_j(x_j)) c_j(t) & \text{if } t > 1, \\ + (1 - \gamma_j(x_j + 1)) \Delta H_j(x_j, t - 1) \\ + \gamma_j(x_j) \Delta H_j(x_j - 1, t - 1) & \text{if } t = 1. \\ 0 & \end{cases}$$

Now consider the linear cancellations Assumption 2.2. Lemma 3.3 below simplifies the calculations of $\Delta H_j(x_j, t)$ even further. The statement is taken from [18].

Lemma 3.3. (Lemma 3 in [18])

Under Assumption 2.2 $\Delta H_j(x_j, t)$ is independent of the number of reservations x_j such that we can write $\Delta H_j(t) = \Delta H_j(x_j, t)$ for all $j \in N$. Moreover, $\Delta H_j(t)$ satisfies

$$\Delta H_j(t) = \begin{cases} \gamma_j c_j(t) + (1 - \gamma_j) \Delta H_j(t - 1) & \text{if } t > 1, \\ 0 & \text{if } t = 1, \end{cases}$$

for all $j \in N$.

Finally, consider the model under the linear and equal cancellations Assumption 2.3. Theorem 3.1 below shows that the states of the problem is significantly reduced. The statement is an extension of Theorem 1 in Subramanian et alii [18] with customer choice behaviour.

Theorem 3.1.

Under Assumption 2.3 we have that $\tilde{V}_t(x) = \tilde{W}_t(y)$ for all (x, t) , where $y = \sum_{j \in N} x_j$ and $\tilde{W}_t(y)$ is given by the following formula:

$$\tilde{W}_t(y) = \begin{cases} \max_{S \subset N} \left\{ \lambda \sum_{j \in S} P_j(S) (r_j - \Delta H_j(t) + \tilde{W}_{t-1}(y + 1)) \right. \\ \quad \left. + \gamma y \tilde{W}_{t-1}(y - 1) \right. \\ \quad \left. + \left(1 - \lambda \sum_{j \in S} P_j(S) - \gamma y \right) \tilde{W}_{t-1}(y) \right\} & \text{if } t > 0, \\ q(y) & \text{if } t = 0. \end{cases} \quad (4)$$

Proof. We proof that $\tilde{V}_t(x) = \tilde{W}_t(y)$ by induction to t . For $t = 0$ we have that $\tilde{V}(x, 0) = q(x)$ and $\tilde{W}(y, 0) = q(y)$. By Assumption 2.3 we have that

$$\tilde{V}(x, 0) = q(x) = q(y) = \tilde{W}(y, 0),$$

so the statement holds for $t = 0$.

Suppose $t > 0$ and the statement holds for all values smaller than t . By Assumption 2.3, Lemma 3.2, and Lemma 3.3 we have that

$$\begin{aligned} \tilde{V}_t(x) = \max_{S \subset N} & \left\{ \lambda \sum_{j \in S} P_j(S) (r_j - \Delta H_j(t) + \tilde{V}_{t-1}(x + e_j)) \right. \\ & + \gamma \sum_{j \in N} x_j (\tilde{V}_{t-1}(x - e_j)) \\ & \left. + \left(1 - \lambda \sum_{j \in S} P_j(S) - \gamma \sum_{j \in N} x_j \right) \tilde{V}_{t-1}(x) \right\}. \end{aligned}$$

Using the induction hypothesis we can rewrite this to

$$\begin{aligned} \tilde{V}_t(x) = \max_{S \subset N} & \left\{ \lambda \sum_{j \in S} P_j(S) (r_j - \Delta H_j(t) + \tilde{W}_{t-1}(y + 1)) \right. \\ & \left. + \gamma y \tilde{W}_{t-1}(y - 1) + \left(1 - \lambda \sum_{j \in S} P_j(S) - \gamma y \right) \tilde{W}_{t-1}(y) \right\}, \end{aligned}$$

and hence $\tilde{V}_t(x) = \tilde{W}_t(y)$. □

Example 3.2 showed that in general inefficient sets can be optimal. However, under Assumption 2.3 inefficient sets are never an optimal solution, as Proposition 3.1 below states. The set of efficient sets still depends on (x, t) , so unfortunately these efficient sets are not of much use to us. 250

Proposition 3.1. An inefficient set is never an optimal solution to Equation (4).

Proof. See Appendix A. □

255 3.3. Independence of Cancellations Heuristic (IOC)

The *Independence of Cancellations* heuristic is appropriate for the model under the independence of cancellations Assumption 2.1. The goal of this heuristic is to find a solution under a model with general assumptions. In this approach we reduce the state space by only keeping track of the total number of reservations rather than the number of reservations per product. Again we make the assumption that the probability that product j is cancelled depends only on the number of reservations for product j , i.e., $\gamma_j^t(x) = \gamma_j^t(x_j)$ (see the independence of cancellations Assumption 2.1), but we make no further assumptions. The advantage of this heuristic is that it can be applied to problems with general cancellation probability functions. 260

Define the one-dimensional state space Y by

$$Y := \{y \in \mathbb{N} \mid 0 \leq y \leq C_{\max}\}.$$

Each state $y \in Y$ corresponds to the total number of reservations. Let $\gamma(y, t)$ be the conditional probability that a reservation is cancelled in state (y, t) . Let $c(y, t)$ be the expected costs for a cancelled reservation in state (y, t) . The Bellman equation corresponding to this problem is 265

$$W_t(y) = \begin{cases} \max_{S \subset N} \left\{ \lambda \sum_{j \in S} P_j(S) (r_j + W_{t-1}(y + 1)) \right. & \text{if } t > 0, \\ \left. + \gamma(y, t) (-c(y, t) + W_{t-1}(y - 1)) \right. \\ \left. + \left(1 - \lambda \sum_{j \in S} P_j(S) - \gamma(y, t) \right) W_{t-1}(y) \right\}, & \\ q(y) & \text{if } t = 0. \end{cases} \quad (5)$$

The problem with using this one-dimensional state space is that we do not know the expected costs $c(y, t)$ and the conditional cancellation probability $\gamma(y, t)$. This follows from the fact that we do not know the mix of reservations on the books, since if we are in state $y \in Y$ we could be in any state $x \in X$ with

$$\sum_{j=1}^n x_j = y.$$

The current mix of reservations on the books depends on the past strategy and therefore $c(y, t)$ and $\gamma(y, t)$ depend on the past strategy. To overcome this problem we propose the following procedure to estimate $c(y, t)$ and $\gamma(y, t)$.

270 Let π be an arbitrary strategy. Let $\bar{x}(y, t) \in \mathbb{R}^n$ be the expected mix of reservations on the books in state y at time t . Let $p(y, t)$ be the probability that we are in state y at time t . Let $p_t^y(x_j = k)$ be the probability that we have k reservations for product j in state (y, t) . For strategy π we will explain a recursive procedure to calculate $\bar{x}(y, t)$, $p(y, t)$, $p_t^y(x_j = k)$, $\gamma(y, t)$, and $c(y, t)$.

275 Let y and t be arbitrary and suppose we know $\bar{x}(y', t')$, $p(y', t')$, $\gamma(y', t')$, $p_{t'}^{y'}(x_j = k)$, and $c(y', t')$ for all $t' < t$, for all k and j , and for all y' . Then we can evaluate their values as stated in Lemma 3.4 and Lemma 3.5 below.

Lemma 3.4. Define $\hat{p}_t^y(x_j = k)$ by

$$\begin{aligned} \hat{p}_t^y(x_j = k) = & p(y+1, t+1) \left[p_{t+1}^{y+1}(x_j = k+1) \gamma_j^{t+1}(k+1) \right. \\ & \left. + p_{t+1}^{y+1}(x_j = k) \sum_{i \neq j} \sum_{k'=0}^{y+1-k} p_{t+1}^{y+1}(x_i = k') \gamma_i^{t+1}(k') \right] \\ & + p(y, y+1) p_{t+1}^y(x_j = k) \left[1 - \lambda \sum_{i \in S} P(S, i) \right. \\ & \left. - \sum_{i \neq j} \sum_{k'=0}^{y-k} p_{t+1}^y(x_i = k') \gamma_i^{t+1}(k') - \gamma_j^{t+1}(k) \right] \\ & + p(y-1, t+1) \left[p_{t+1}^{y-1}(x_j = k-1) \lambda P_j(S) \right. \\ & \left. + p_{t+1}^{y-1}(x_j = k) \lambda \sum_{i \in S \setminus \{j\}} P_j(S) \right], \end{aligned}$$

for all k . Then $p_t^y(x_j = k)$ is given by

$$p_t^y(x_j = k) = \frac{\hat{p}_t^y(x_j = k)}{\sum_{k'=0}^y \hat{p}_t^y(x_j = k')}.$$

Proof. See [Appendix A](#). □

Lemma 3.5. $\bar{x}(y, t)$, $p(y, t)$, $\gamma(y, t)$, and $c(y, t)$ can be found via the following formulas:

$$\begin{aligned}\bar{x}_j(y, t) &= \sum_{k=0}^y p_t^y(x_j = k)k, \\ p(y, t) &= p(y+1, t+1)\gamma(y+1, t+1) + \lambda \sum_{j \in S} P_j(S)p(y-1, t+1) \\ &\quad + \left(1 - \lambda \sum_{j \in S} P_j(S) - \gamma(y, t+1)\right) p(y, t+1), \\ \gamma(y, t) &= \sum_{j=1}^n \sum_{k=0}^y p_t^y(x_j = k)\gamma_j^t(k), \\ c(y, t) &= \frac{\sum_{j=1}^n \sum_{k=0}^y p_t^y(x_j = k)\gamma_j^t(k)r_j(t)}{\gamma(y, t)}.\end{aligned}$$

Moreover, under the linear and equal cancellation rates Assumption 2.3 we have that

$$\begin{aligned}\gamma(y, t) &= \sum_{j=1}^n \gamma_j^t \bar{x}_j(y, t), \\ c(y, t) &= \frac{\sum_{j=1}^n r_j(t) \gamma_j^t \bar{x}_j(y, t)}{\gamma(y, t)}.\end{aligned}$$

Proof. See Appendix A. □

We can use the equations from Lemma 3.4 and Lemma 3.5 to iteratively calculate the conditional cancellation probability $\gamma(y, t)$ and the expected costs $c(y, t)$ of a cancelled product in state (y, t) under a policy π .

The Independence of Cancellations algorithm iterates the following procedure until we hit a stopping criterion, for example after a fixed number of iterations. First we start with an arbitrary expected costs $\hat{c}(y, t)$ and cancellation probability $\hat{\gamma}(y, t)$ for each state (y, t) . A good initial solution would be to use the linear and equal cancellation rates Assumption 2.3 and evaluate $\hat{c}(y, t)$ and $\hat{\gamma}(y, t)$ according to the obtained strategy. Then we solve the problem using dynamic programming with Equation (5), where we use $\hat{c}(y, t)$ as an approximation of the expected costs $c(y, t)$ and $\hat{\gamma}(y, t)$ as an approximation of the conditional cancellation probability $\gamma(y, t)$. Then we find an optimal strategy π which we use to calculate the expected costs $\hat{c}(y, t)$ and the c cancellation probability $\hat{\gamma}(y, t)$ corresponding to π , using the equations in Lemma 3.4 and Lemma 3.5. See Algorithm 3.1 for a summary of this heuristic.

Algorithm 3.1. Independence of Cancellations Heuristic (IOC)

1. Start with initial expected costs $\hat{c}(y, t)$ and cancellation probability $\hat{\gamma}(y, t)$.
2. Find an optimal strategy π with dynamic programming on Equation (5) using $\hat{c}(y, t)$ as an approximation for $c(y, t)$ and $\hat{\gamma}(y, t)$ as an approximation for $\gamma(y, t)$.
3. Use π and the equations in Lemma 3.4 and Lemma 3.5 to calculate $\hat{c}(y, t)$ and $\hat{\gamma}(y, t)$.
4. Go to step 2 unless we hit a stopping criterion.

We now investigate the complexity of Step 2 of this algorithm.

Efficient Sets

Another elegant property of the Independence of Cancellations (IOC) algorithm is that only efficient offer sets are potential optimal sets and the others are never optimal. Moreover, this set of potential offer-sets

is the same for all states (y, t) . In Proposition 2 of Talluri and van Ryzin [27] it is shown that in the customer choice base model inefficient sets are never optimal. We show in Proposition 3.2 that inefficient sets are never optimal for the IOC algorithm and that the set of efficient sets is the same for all (y, t) .

310

Proposition 3.2. An inefficient set is never an optimal solution to Equation (5). Moreover, the set of efficient sets is the same for all (x, t) .

315

Proof. In Equation (5) we have that $\Delta H_j(y, t) = 0$ for all j and (y, t) , so the set of efficient sets is the same for all (y, t) .

The proof that inefficient sets are never an optimal solution to Equation (5) is analogous to the proof of Proposition 3.1. \square

3.4. Linear Cancellation Rates Heuristic (LCR)

The goal of the second heuristic is to be fast and efficient. To accomplish this we do the following:

320

– *We work under the linear cancellation rates Assumption 2.2.* This way the model is restricted, but we gain computation time.

– *We only keep track of the total number of reservations.* We loose the information that we have if we keep track of the number of reservations per product, but we gain computation time.

325

– *We only consider efficient sets in our policy.* Then the computation time is reduced because we need to consider less solutions, but we loose some potential better solutions.

– *We use the same set of efficient sets in each state.* We gain computation time because we do not have to evaluate the efficient sets for each state, but we loose some potential better solutions.

In the second heuristic we work under the linear cancellation rates, i.e., we assume that $\gamma_j(x_j) = \gamma_j x_j$ for all j . The heuristic can also be applied if we work under the linear and equal cancellation rates Assumption 2.3, since this assumption is a special case of the linear cancellation rates assumption. We solve a one-dimensional problem, where we only keep track of the total number of reservations and not of the reservations per product. We reformulate the problem such that the costs of cancellations are incorporated in the maximisation term of the Bellman equation, as described in Theorem 3.1. Only this time we only consider offer sets that are efficient. In particular, we choose the efficient sets that result from taking $R(S, x, t) = 0$ for all (x, t) . The motivation for choosing only these efficient sets is that the numerical results in Section 5 show that (a) leaving out inefficient solutions does not have a great impact on performance; and (b) choosing among the efficient sets is better than choosing among an equal number of arbitrary sets. The Bellman equation that we have to solve is given by

$$\tilde{W}_t(y) = \begin{cases} \max_{\substack{S \subset N \\ \text{efficient}}} \left\{ \lambda \sum_{j \in S} P_j(S) (r_j - \Delta H_j(t) + \tilde{W}_{t-1}(y+1)) \right\} & \text{if } t > 0, \\ + \gamma y \tilde{W}_{t-1}(y-1) + \left(1 - \lambda \sum_{j \in S} P_j(S) - \gamma y \right) \tilde{W}_{t-1}(y) \} & \\ q(y) & \text{if } t = 0. \end{cases} \quad (6)$$

However, we assume that the cancellation probabilities differ per product. We do not know the cancellation probability γ if we are in state y , since we do not know what the mix of reservations on the books is. This problem is similar to the problem we described in Section 3.3. We could use Lemma 3.5 to approximate γ iteratively, but this is computationally expensive. Moreover, numerical results show that simpler approximations outperform this method (see also Section 5). An example of a simple approximation of γ is taking the average of all γ_j 's:

$$\gamma := \frac{1}{n} \sum_{j \in N} \gamma_j. \quad (7)$$

The advantage of using this method over the Independence of Cancellations algorithm is that it is computationally faster and performs better under the linear cancellation rates assumption. However, the

330 downside of this method is that it can not be applied to general cancellation probabilities. See Algorithm
 3.2 for an overview of the heuristic.

Algorithm 3.2. Linear Cancellation Rates Heuristic (LCR)

1. Estimate $\hat{\gamma}$ by using (γ_j) , for example by Equation (7).
- 335 2. Find an optimal strategy of Equation (6).

3.5. Multiple Room Types Heuristic (MRT)

In practice hotels often have multiple room types. However, we made the assumption that the hotel has C rooms of the same type. Expanding the solution space with multiple room types leads to an intractable
 340 problem. To solve this problem we combine the Linear Cancellation Rates heuristic of Section 3.4 with *one-step improvement*, i.e., one step of the well know policy iteration (see also Puterman [25]). In this approach we first find an approximation of the value function, and then use one maximisation step of policy iteration to find a better solution.

First we consider the different room types separately. Let \mathcal{I} be the set of room types. For each room
 345 type $i \in \mathcal{I}$ we have a certain capacity C_i . For each room type we only keep track of the total number of reservations for that room type. Under the linear and equal cancellation rates Assumption 2.3 this gives the exact solution; see Theorem 3.1. Let S_i be the offer set of products for room type i and let N_i be the set of products for room type i . We solve the problem per room type, as if it would be the only room type we have, using the Linear Cancellation Rates heuristic described in Section 3.4. The Bellman
 350 equation per room type $i \in \mathcal{I}$ is given by

$$\tilde{W}_t^i(y) = \begin{cases} \max_{\substack{S_i \subset N_i \\ \text{efficient}}} \left\{ \lambda \sum_{j \in S_i} P_j(S_i) (r_j - \Delta H_j(t) + \tilde{W}_{t-1}^i(y+1)) \right\} & \text{if } t > 0 \\ + \gamma^i y \tilde{W}_{t-1}^i(y-1) + \left(1 - \lambda \sum_{j \in S_i} P_j(S_i) - \gamma^i y \right) \tilde{W}_{t-1}^i(y) & \\ q(y) & \text{if } t = 0. \end{cases} \quad (8)$$

Let u be the vector containing the number of reservations per room type, such that u_i represents the number of reservations for room type $i \in \mathcal{I}$. Then we want to solve the following recursive formula:

$$U_t(u) = \begin{cases} \max_{S \subset N} \left\{ \lambda \sum_{i \in \mathcal{I}} \left[\sum_{j \in S_i} P_j(S_i) (r_j - \Delta H_j(t) + U_{t-1}(u + e_i)) \right] \right\} & \text{if } t > 0, \\ + \sum_{i \in \mathcal{I}} \left[\sum_{j \in N_i} \gamma_j u_j U_{t-1}(u - e_i) \right] & \\ + \left(1 - \lambda \sum_{j \in S} P_j(S) - \sum_{j \in N} \gamma_j u_j \right) U_{t-1}(u) & \\ q(u) & \text{if } t > 0. \end{cases}$$

However, this problem is still too large to solve, even if we use a small number of room types. Therefore we use *one-step improvement* to solve this problem. Suppose u is the current state and we wish to decide which products to offer. We approximate $U_{t-1}(u)$ by the following formula:

$$\tilde{U}_{t-1}(u) := \sum_{i \in \mathcal{I}} \tilde{W}_{t-1}^i(u_i). \quad (9)$$

$\tilde{U}_{t-1}(u + e_i)$ and $\tilde{U}_{t-1}(u - e_i)$ can be approximated similarly. Then we do one maximisation step to find an approximation $\tilde{U}_t(u)$ for $U_t(u)$:

$$\bar{U}_t(u) = \begin{cases} \max_{S \in \mathcal{N}} \left\{ \lambda \sum_{i \in \mathcal{I}} \left[\sum_{j \in S_i} P_j(S_i) (r_j - \Delta H_j(t) + \tilde{U}_{t-1}(u + e_i)) \right] \right\} & \text{if } t > 0, \\ + \sum_{i \in \mathcal{I}} \left[\sum_{j \in N_i} \gamma_j x_j \tilde{U}_{t-1}(u - e_i) \right] & \\ + \left(1 - \lambda \sum_{j \in S} P_j(S) - \sum_{j \in \mathcal{N}} \gamma_j x_j \right) \tilde{U}_{t-1}(u) & \\ q(u) & \text{if } t = 0. \end{cases} \quad (10)$$

The multiple room type heuristic is summarised in Algorithm 3.3 below.

Algorithm 3.3. Multiple Room Types

1. Fix state u .
2. Evaluate $\tilde{W}_{t-1}^i(u_i)$, $\tilde{W}_{t-1}^i(u_i + 1)$, and $\tilde{W}_{t-1}^i(u_i - 1)$ using Equation (8) for all $i \in \mathcal{I}$.
3. Define $\tilde{U}_{t-1}(u)$, $\tilde{U}_{t-1}(u + e_i)$, and $\tilde{U}_{t-1}(u - e_i)$ using Equation (9).
4. Use Equation (10) to find an improved approximation $\bar{U}_t(u)$ of $U_t(u)$.

Remark. The action space of this problem is still large, even though we only use efficient sets per room type. The concept of efficient sets is only defined for single-resource problems. In the context of multiple room types it is worth looking into a generalisation of efficient sets to multiple-resource problems. However, this is beyond the scope of this paper.

4. Estimating Parameters

In this section we propose a method to estimate the parameters of the model from real data. Ideally we use an estimation model under perfect knowledge, i.e., we assume we know (a) which products were offered at what time, (b) when customers arrived, (c) when they made a purchase, and (d) when they cancelled their reservation. A maximum-likelihood estimation model under perfect knowledge is described in AppendixB. In practice not all information is available, unfortunately. For example, the no-purchase option is rarely observed. To overcome this problem we need to consider methods that deal with censored data. One method is the well-known expectation-maximisation algorithm, which is described in AppendixB. However, the expectation-maximisation algorithm is unattractive for practical purposes for two reasons. First, the computation time is notoriously long. Second, the method estimates the parameters of the discretised model instead of the continuous model, which may lead to estimation errors. To overcome these problems we propose an estimation model that is based on the excellent estimation method proposed by Newman et alii [24]. The proposed method assumes the linear cancellation rates Assumption 2.2 and that the purchase probabilities are modelled by the *multinomial logit model* (MNL), which is briefly explained in AppendixB.

In what follows the purchase probabilities and the offer set depend on the time period. Therefore we introduce the following notation. Let Z_{tj} be the variable containing values of the attributes of product j in time period t and let $Z_t = \{Z_{tj}\}_j$. Let S_t be the set of products that is offered in time period t . Let α be the utility of the no-purchase alternative. For convenience we introduce some notation. Let $P_{tj}(\alpha, \beta, S_t, Z_t)$ be the probability that product j is purchased in time period t under α , β , S_t , and Z_t , i.e.,

$$P_{tj}(\alpha, \beta, S_t, Z_t) := \frac{\exp(\beta^\top Z_{jt})}{\exp(\alpha) + \sum_{i \in S_t} \exp(\beta^\top Z_{it})}.$$

Define the probability $P_{t0}(\alpha, \beta, S_t, Z_t)$ that the no-purchase option is chosen as

$$P_{t0}(\alpha, \beta, S_t, Z_t) := \frac{\exp(\alpha)}{\exp(\alpha) + \sum_{i \in S_t} \exp(\beta^\top Z_{it})}.$$

It is convenient to use the probability $P_{t*}(\alpha, \beta, S_t, Z_t)$ that a product is purchased, i.e., the probability that the no-purchase alternative is not chosen. Hence $P_{t*}(\alpha, \beta, S_t, Z_t)$ is given by

$$\begin{aligned} P_{t*}(\alpha, \beta, S_t, Z_t) &:= \sum_{j \in S_t} P_{tj}(\alpha, \beta, S_t, Z_t) \\ &= \frac{\sum_{j \in S_t} \exp(\beta^\top Z_{jt})}{\exp(\alpha) + \sum_{i \in S_t} \exp(\beta^\top Z_{it})} \\ &= 1 - P_{t0}(\alpha, \beta, S_t, Z_t). \end{aligned}$$

Finally, let $P_{tj|*}(\beta, S_t, Z_t)$ be the probability that product j is purchased under the condition that the customer made a purchase, i.e.,

$$P_{tj|*}(\beta, S_t, Z_t) := \frac{\exp(\beta^\top Z_{jt})}{\sum_{i \in S_t} \exp(\beta^\top Z_{it})}.$$

We divide time into T time periods and assume that the parameters to be estimated are constant within these periods. Note that we allow more than one event to happen in one time period. The parameters that we need to estimate are: the no-purchase utility α ; the weights β of the attributes of the MNL model; the parameters γ_j of the exponential distributions that model the cancellations; and the parameter λ of the Poisson process that models the arriving customers. We do this by using maximum likelihood estimation.

Assume that the offer set S_t does not change within this time period. Suppose we observe z_{jt} purchases for product j in time period t . Let $\zeta_j(y)$ be the total number of cancellations for product j when there were y reservations for product j , measured over all time periods. Suppose that we observe that for a period of $t_j(y)$ there were y reservations for product j , again measured over all time periods. Note that $t_j(y)$ might be 0.

The maximum likelihood function $L(z, \zeta | \alpha, \beta, \gamma, \lambda)$ is the probability that we observe z and ζ under given α, β, γ , and λ . Since the arrival process is independent from the cancellation process we can write

$$p(z, \zeta | \alpha, \beta, \gamma, \lambda) = p(z | \alpha, \beta, \gamma, \lambda) p(\zeta | \alpha, \beta, \gamma, \lambda),$$

with $p(z | \alpha, \beta, \gamma, \lambda)$ the probability that we observe z given α, β, γ , and λ ; and $p(\zeta | \alpha, \beta, \gamma, \lambda)$ the probability that we observe ζ given α, β, γ , and λ . Since the cancellations are only dependent on γ we can even write

$$p(\zeta | \alpha, \beta, \gamma, \lambda) = p(\zeta | \gamma),$$

with $p(\zeta | \gamma)$ the probability that we observe ζ given γ . This also follows from the derivations below.

The cancellation processes are independent from each other, so we can write

$$p(\zeta | \gamma) = \prod_{j \in N} p(\zeta_j | \gamma_j),$$

with $p(\zeta_j | \gamma_j)$ the probability that we observe ζ_j given γ_j . Moreover, if we consider the time intervals $t_j(y)$ we can write

$$p(\zeta_j | \gamma_j) = \prod_{y=0}^{C_{\max}} p(\zeta_j(y) | \gamma_j).$$

Since the cancellations follow a Poisson process with parameter $\gamma_j y t_j(y)$ the probability $p(\zeta_j(y) | \gamma_j)$ that we observe $\zeta_j(y)$ cancellations in time period t conditional on γ_j is equal to

$$p(\zeta_j(y) | \gamma_j) = (\gamma_j y t_j(y))^{\zeta_j(y)} \frac{e^{-\gamma_j y t_j(y)}}{\zeta_j(y)!}.$$

If we take the natural logarithm of $p(\zeta_j(y)|\gamma_j)$ we get

$$\log(p(\zeta_j(y)|\gamma_j)) = \zeta_j(y) \log(\gamma_j) + \zeta_j(y) \log(yt_j(y)) - \gamma_j yt_j(y) - \log(\zeta_j(y)!).$$

Combining all equations we find

$$\log(p(\zeta|\gamma)) = \sum_{j \in N} \sum_{y=0}^{C_{\max}} [\{\zeta_j(y) \log(\gamma_j) - \gamma_j yt_j(y)\} + \{\zeta_j(y) \log(yt_j(y)) - \log(\zeta_j(y)!\}].$$

If we want to maximise this with respect to γ we can solve the following equations separately for each $j \in N$:

$$\sum_{y=0}^{C_{\max}} \zeta_j(y) \log(\gamma_j) - \gamma_j yt_j(y). \quad (11)$$

If we take the first derivative with respect to γ_j we find

$$\frac{1}{\gamma_j} \sum_{y=0}^{C_{\max}} \zeta_j(y) - \sum_{y=0}^{C_{\max}} yt_j(y).$$

The optimum can be found by setting this equation to zero. This gives

$$\gamma_j = \frac{\sum_{y=0}^{C_{\max}} \zeta_j(y)}{\sum_{y=0}^{C_{\max}} yt_j(y)}.$$

Let $\zeta_j := \sum_{y=0}^{C_{\max}} \zeta_j(y)$, i.e., ζ_j is the total number of cancellations for product j . Then we have that

$$\gamma_j = \frac{\zeta_j}{\sum_{y=0}^{C_{\max}} yt_j(y)}. \quad (12)$$

By taking the second derivative of Equation (11) we find

$$-\frac{1}{\gamma_j^2} \sum_{y=0}^{C_{\max}} \zeta_j(y),$$

which is negative, such that the optimum found in Equation (12) is a maximum.

For completeness we below present the derivations of the results by Newman et alii [24], which can directly be applied to our model.

The expression for $p(z|\alpha, \beta, \gamma, \lambda)$ does not depend on γ so we can write $p(z|\alpha, \beta, \lambda)$. Conditioning to the number of reservations m_t in time period t gives

$$p(z|\alpha, \beta, \lambda) = p(z|m, \beta, \lambda)p(m|\alpha, \beta, \lambda),$$

where $m = (m_1, \dots, m_T)$. Note that $p(z|m, \beta, \lambda)$ does not depend on the no-purchase alternative parameter α . The probability $p(z|m, \beta, \lambda)$ that we observe z under the conditions that we observe m_t reservations with parameter β and λ is equal to

$$p(z|m, \beta, \lambda) = \prod_{t=1}^T \binom{m_t}{z_{1t}, \dots, z_{nt}} \prod_{j \in N} [P_{tj|*}(\beta, S_t, Z_t)]^{z_{jt}}.$$

Note that since $\sum_{j \in N} z_{jt} = m_t$ and $z_{jt} = 0$ for $j \notin S_t$ we have that

$$\prod_{j \in N} [P_{tj|*}(\beta, S_t, Z_t)]^{z_{jt}} = \prod_{j \in S_t} \left[\frac{\exp(\beta^\top Z_{jt})}{\sum_{i \in S_t} \exp(\beta^\top Z_{it})} \right]^{z_{jt}} = \frac{\prod_{j \in S_t} [\exp(\beta^\top Z_{jt})]^{z_{jt}}}{[\sum_{i \in S_t} \exp(\beta^\top Z_{it})]^{m_t}}.$$

By taking the natural logarithm we find

$$\log(p(z|m, \alpha, \beta, \lambda)) = \sum_{t=1}^T \log\left(\frac{m_t!}{z_{1t}! \cdots z_{nt}!}\right) + \sum_{j \in S_t} z_{jt} \beta^\top Z_{jt} - m_t \log\left(\sum_{i \in S_t} \exp(\beta^\top Z_{it})\right).$$

If we want to maximise this with respect to β we only have to consider

$$\sum_{t=1}^T \left[\sum_{j \in S_t} z_{jt} \beta^\top Z_{jt} - m_t \log\left(\sum_{i \in S_t} \exp(\beta^\top Z_{it})\right) \right]. \quad (13)$$

On the other hand, m_t follows a Poisson distribution with parameter $\lambda \sum_{i \in S} P_{t*}(\alpha, \beta, S_t, Z_t)$. Therefore the probability $p(m|\alpha, \beta, \lambda)$ that we observe m_t reservations in time period t is given by

$$p(m|\alpha, \beta, \lambda) = \prod_{t=1}^T \frac{[\lambda \sum_{i \in S} P_{t*}(\alpha, \beta, S_t, Z_t)]^{m_t} e^{-\lambda P_{t*}(\alpha, \beta, S_t, Z_t)}}{m_t!}.$$

Taking the logarithm gives

$$\log(p(m|\alpha, \beta, \lambda)) = \sum_{t=1}^T m_t [\log(\lambda P_{t*}(\alpha, \beta, S_t, Z_t))] - \lambda P_{t*}(\alpha, \beta, S_t, Z_t) - \log(m_t!). \quad (14)$$

Note that if we maximise Equation (14) with respect to α and λ we can ignore the last term. The first partial derivative with respect to λ equals

$$\frac{1}{\lambda} \sum_{t=1}^T m_t - \sum_{t=1}^T P_{t*}(\alpha, \beta, S_t, Z_t).$$

The optimum is attained at

$$\lambda = \frac{\sum_{t=1}^T m_t}{\sum_{t=1}^T P_{t*}(\alpha, \beta, S_t, Z_t)}, \quad (15)$$

where $\sum_{t=1}^T m_t$ is the total number of observed reservations. The second partial derivative with respect to λ equals

$$-\frac{1}{\lambda^2} \sum_{t=1}^T m_t,$$

which is negative for $\lambda \neq 0$, so the optimum from Equation (15) is a maximum. We can replace λ in Equation (14) by Equation (15), which gives

$$\begin{aligned} \log(p(m_t|\alpha, \beta, \lambda)) &= \sum_{t=1}^T m_t \log\left(\frac{\sum_{t=1}^T m_t}{\sum_{t=1}^T P_{t*}(\alpha, \beta, S_t, Z_t)} P_{t*}(\alpha, \beta, S_t, Z_t)\right) \\ &\quad - \frac{\sum_{t=1}^T m_t}{\sum_{t=1}^T P_{t*}(\alpha, \beta, S_t, Z_t)} P_{t*}(\alpha, \beta, S_t, Z_t) - \log(m_t!). \end{aligned} \quad (16)$$

If β is given we can find α by maximising Equation (16) with respect to α . Then we can find λ by using Equation (15).

400 The estimation procedure is summarised in Algorithm 4.1.

Algorithm 4.1. Three-Step Continuous

1. Find $\hat{\gamma}$ using Equation (12).
2. Find $\hat{\beta}$ using Equation (13).

3. Use the estimate $\hat{\beta}$ to find $\hat{\alpha}$ using Equation (16). Then use $\hat{\beta}$ and $\hat{\alpha}$ to find $\hat{\lambda}$ using Equation (15).

Theorem 4.1 below states that Algorithm 4.1 provides consistent estimators. Moreover, the estimator for the cancellation parameters γ is an unbiased estimator.

Theorem 4.1. As the sample size increases, Algorithm 4.1 leads to consistent estimators. Moreover, the estimator for γ found by Equation (12) is unbiased.

Proof. Theorem 1 and Theorem 2 in Newman et alii [24] show that the estimators of β and (α, λ) are consistent, respectively.

To show that the estimator $\hat{\gamma}$ of the cancellation parameter found by Equation (12) is a consistent estimator, we consider the cancellation parameters separately per product. Consider a product $j \in N$. Let γ_j^* be the true parameter and let γ_j^k be the parameter obtained using Equation (12) with sample size k . In other words,

$$\gamma_j^k = \frac{\sum_{y=0}^{C_{\max}} \zeta_j^k(y)}{\sum_{y=0}^{C_{\max}} yt_j^k(y)},$$

with $\zeta_j^k(y)$ the observed number of cancellations in state y and $t_j^k(y)$ the observed time spend in state y with k observations. We can think of one observation as an arrival day with all booking history. The expected number of cancellations when the state is y is equal to $\gamma_j^* yt_j^k(y)$. Note that $t_j^k(y)$ is the sum of k i.i.d. random variables $t_j(y)$, where $t_j(y)$ is the time spend in state y for one observation. As the sample size increases, we have that

$$\lim_{k \rightarrow \infty} \frac{yt_j^k(y)}{k} = \gamma_j^* y \mathbb{E}[t_j(y)].$$

On the other hand, $\zeta_j^k(y)$ is the sum of k i.i.d. random variables $\zeta_j(y)$, where $\zeta_j(y)$ is the number of cancellations for one observation. As the sample size increases, we have that

$$\lim_{k \rightarrow \infty} \frac{\zeta_j^k(y)}{k} = \mathbb{E}[\zeta_j(y)] = \mathbb{E}[\mathbb{E}[\zeta_j(y)|t_j(y)]] = \mathbb{E}[\gamma_j^* yt_j(y)] = \gamma_j^* y \mathbb{E}[t_j(y)].$$

Hence we have that

$$\lim_{k \rightarrow \infty} \gamma_j^k = \lim_{k \rightarrow \infty} \frac{\frac{1}{k} \sum_{y=0}^{C_{\max}} \zeta_j^k(y)}{\frac{1}{k} \sum_{y=0}^{C_{\max}} yt_j^k(y)} = \frac{\gamma_j^* \sum_{y=0}^{C_{\max}} y \mathbb{E}[t_j(y)]}{\sum_{y=0}^{C_{\max}} y \mathbb{E}[t_j(y)]} = \gamma_j^*,$$

so Equation (12) provides a consistent estimator. Moreover, using conditional expectation it follows that

$$\mathbb{E}[\gamma_j^k] = \mathbb{E} \left[\mathbb{E} \left[\frac{\sum_{y=0}^{C_{\max}} \zeta_j^k(y)}{\sum_{y=0}^{C_{\max}} yt_j^k(y)} \middle| t_j^k(\cdot) \right] \right] = \mathbb{E} \left[\frac{\sum_{y=0}^{C_{\max}} \gamma_j^* yt_j^k(y)}{\sum_{y=0}^{C_{\max}} yt_j^k(y)} \right] = \gamma_j^*,$$

so Equation (12) provides an unbiased estimator. \square

5. Numerical Results

In this section numerical results are provided to validate the customer choice cancellation model. The four main results that are provided are: 1) the fast computation time of the estimation method described in Section 4; 2) the fast computation time of the heuristics; 3) the good performance of the heuristics; and 4) the performance of the heuristics under estimated parameters. The parameters that are used are based on Simulation Example 2 in Talluri and van Ryzin [27]. Let $n = 10$ be the number of products sold with corresponding price vector

$$r = (240, 220, 190, 160, 120, 112, 96, 80, 74, 70).$$

Demand, cancellation rates, and purchase probabilities are independent from the time period t . To allow the study of the effect of high volume in demand and low volume in demand the *load factor* l is introduced. In our studies we use $l \in \{0.6, 0.8, 1, 1.2, 1.4\}$. Demand λ per time unit is then defined as

$$\lambda(C, l, T) = \frac{Cl}{T}.$$

A low value of l implies a low demand relative to the capacity C and a high value of l implies a high demand relative to the capacity C .

Purchase probabilities are modelled by the MNL model. The only attribute that we consider is price, but we assume that there exist high price-sensitive and low price-sensitive customers which have different parameters $\beta_H = -0.005$ and $\beta_L = 0.0015$, respectively (as in [27]). The no-purchase parameter is set to $\alpha = 0$ such that the MNL model is the same as in [27]. We allow overbooking up to 20% of the total capacity C . Cancellation rates are assumed linear and the parameters $\gamma \in \mathbb{R}^n$ depend on l and T in the following way:

$$\begin{aligned} \gamma &= (0.720, 0.320, 0.560, 0.280, 0.200, 0.360, 0.160, 0.080, 0.008, 0.040) \frac{l}{T} \\ &= (18/25, 8/25, 14/25, 7/25, 1/5, 9/25, 4/25, 2/25, 1/125, 1/25) \frac{l}{T}. \end{aligned}$$

420 The policies that follow from the heuristics need to be compared using the corresponding revenue. For small problem instances the exact solution method can be used, but for larger instances the problem instance is intractable. In that case simulation is used to estimate the revenue corresponding to a policy. The estimation errors are between 0.1% and 0.4%.

5.1. Computation Time of Estimation Methods

425 An estimation method only has value for practical applications if it can finish in reasonable time. The three-step continuous estimation method (3SC) described in Section 4 has this property.

In Table 3 the computation time of 3SC is compared with the computation time of EM for different sample sizes. The results show that the computation time of the EM-algorithm increases drastically when the sample size increases, while the computation time of the 3SC-method does not. This is only
430 for a small example with $C = 20$ rooms, so the results will be worse for practical instances. Fortunately, the 3SC-method can be used also for larger instances, with more capacity. In Table 4 the computation time of 3SC is given for a hotel with $C = 200$ rooms for different sample sizes. The computation time increases but it stays reasonable.

Observation days	Computation time	
	EM	3SC
10	219.357	42.611
20	650.314	34.159
50	2869.260	39.311
100	9446.129	34.075

Table 3: Computation times (in seconds) for estimating parameters using EM and 3SC for $C = 20$, $T = 100$, $n = 10$, and different number of observation days.

Observation days	Computation time
10	50.47
20	55.39
50	64.52
100	81.22
200	112.98

Table 4: Computation time for 3SC estimation method for different number of observation days, using $C = 200$, $T = 1000$, $n = 10$.

5.2. Computation Time: Heuristics vs. Exact solution

435 In the analysis of the cancellation model it was stated that the model suffers from the curse of dimension-
 440 ality. The state space of the Markov decision process is too large, such that solving the exact problem
 takes too much time. In Table 5 the computation times of the solution methods are presented for a small
 example, where $C = 10$, $T = 20$, $l = 1$, and the number of products n varies from 2 to 10. Evaluating
 instances with $n > 6$ using the exact solution method causes system to be out of memory and the exact
 solution can therefore not be computed.

This example shows that when the number of products increases the computation time increases dras-
 tically, while the computation times for the heuristic stay more or less constant. In Table 6 below the
 computation times of the heuristics are presented for a larger instance, where $C = 200$, $T = 1000$, and
 $n = 10$. Even in this case the computation time is very small.

n	Computation Time			
	Exact	IOC	LCR	TvR
2	0.229	0.295	0.013	0.009
3	1.020	0.292	0.013	0.008
4	4.293	0.327	0.011	0.008
5	15.041	0.309	0.011	0.008
6	49.082	0.309	0.010	0.009
7	-	0.318	0.010	0.011
8	-	0.318	0.011	0.009
9	-	0.327	0.012	0.010
10	-	0.333	0.012	0.012

Table 5: Computation times heuristics vs. exact solution method, using $C = 10$, $T = 20$, load factor $l = 1$, $n = 2, \dots, 10$.

5.3. Performance of Solution Methods

445 The cancellation model is only useful if the tractable heuristics outperform the base model. To this end,
 consider a large instance, with $C = 200$ rooms and $T = 1000$ time periods. The load factors l that are
 considered are $\{0.6, 0.8, 1, 1.2, 1.4\}$. In Table 6 the results are presented.

load factor	IOC	LCR	TvR	random	$\Delta_{\text{IOC}/\text{TvR}}$	$\Delta_{\text{LCR}/\text{TvR}}$	$\Delta_{\text{LCR}/\text{random}}$
0.6	16,878	17,014	16,904	11,981	-0.15%	0.65%	42.00%
0.8	21,390	21,691	21,252	15,492	0.65%	2.06%	40.01%
1.0	25,297	25,949	24,169	18,818	4.67%	7.37%	37.90%
1.2	27,732	29,859	26,154	21,972	6.03%	14.17%	35.89%
1.4	29,386	33,225	27,848	24,969	5.52%	19.31%	33.06%
Comp. time	7.10	0.58	0.47				

Table 6: Performance of the heuristics over the base model, using $C = 200$, $T = 1000$, $n = 10$, and $l \in \{0.6, 0.8, 1, 1.2, 1.4\}$.

450 There are several noteworthy observations from Table 6. First and foremost, using the policies found by
 the IOC and LCR heuristics lead to solutions with *higher revenue* compared the TvR solution method
 (except for the case that $l = 0.6$, when LCR still outperforms TvR, but IOC is slightly outperformed

by TvR). Therefore it is reasonable to consider cancellations in the decision making process. Taking cancellation into account leads to a revenue gain of up to 20%.

Second, the IOC algorithm performs better than the TvR algorithm, but worse than the LCR algorithm. This is the motivation for choosing LCR over IOC if we work under the linear cancellation rates Assumption 2.2, but choosing IOC over TvR if we work with general probability rates under the independence of cancellations Assumption 2.1.

Third, the effects of taking cancellations into account increase when the arrival volume increases. This is intuitively clear: if the arrival rate is higher, the gap caused by a cancelled reservation will be filled quickly. If you do not take cancellations into account, the gap caused by the cancelled reservation will not be filled, which implies a loss in revenue.

5.4. Estimation Accuracy vs. Performance of Solution Methods

Estimation errors potentially lead to bad performance of the optimisation model. Depending on the sample size the estimation error might be small or large. In the next simulation the performance of the solution method using estimated parameters is measured. Again a large instance is considered, where $C = 200$, $T = 1000$, and $n = 10$. The load factor l is given different values, namely $\{0.6, 0.8, 1, 1.2, 1.4\}$, to consider the effect of different demand volumes. For each load factor a dataset of size 100 (arrival days) is created, from which the parameters are estimated. The estimated parameters are used by all heuristics to find the corresponding policy and revenue. In Table 8 the results of the estimation are presented. The true values as well as the estimated values are presented.

		Load factor				
		0.6	0.8	1	1.2	1.4
IOC	True	16,843	21,408	25,316	27,748	29,406
	Est.	17,168	20,198	24,277	26,764	30,039
LCR	True	16,996	21,685	25,964	29,905	33,186
	Est.	17,212	20,222	24,356	27,876	33,590
TvR	True	16,872	21,270	24,138	26,173	27,862
	Est.	17,162	20,080	23,151	24,758	28,432

Table 7: Performance heuristics under estimated values. Variables: same as 1, except for load factor = 1. Use estimates from estimation method example. Note the difference between the other table. This is due to the difference in cancellation factors, which is constant in the other table!

In Table 7 the revenues corresponding to load factor l , heuristic, and estimated or true parameter set is given. From this table several observations can be made. Firstly, the heuristics outperforms the TvR method in all cases: for all load factors, for all parameter sets, true or estimated. This fact suggests that choosing a heuristic that takes cancellations into account is preferred over the TvR model, even if the parameters are estimated and thus contain errors.

Second, the model using the estimated parameters might outperform the model with the true parameters, or it might not. It seems that in this instance the extreme load factors ($l = 0.6$ and 1.4) lead to a better performance of the model under estimated parameters for all heuristics, while with the other load factors the model under the true parameters performs better.

6. Discussion

In this paper we analysed the customer choice cancellation model in Section 2, including simplified versions that result from Assumptions 2.1, 2.2, and 2.3. Moreover, we introduced several solution methods in Section 3 and estimation methods in Appendix B. In this section we briefly describe some guidelines for selecting a model, a solution method, and an estimation method. Moreover, we briefly discuss an extension of the cancellation model to a more realistic choice-based network cancellation model.

	0.6		0.8		1		1.2		1.4	
	True	Est.	True	Est.	True	Est.	True	Est.	True	Est.
α	0	-0.0241	0	0.3561	0	0.3096	0	0.2598	0	-0.0336
λ	0.12	0.1144	0.16	0.1666	0.2	0.2055	0.24	0.2447	0.28	0.2627
β_1	-0.005	-0.004551	-0.005	-0.004958	-0.005	-0.004717	-0.005	-0.005352	-0.005	-0.004794
β_2	-0.0015	-0.001392	-0.0015	-0.001480	-0.0015	-0.001513	-0.0015	-0.001659	-0.0015	-0.001466
γ_1	0.000432	0.0004539	0.000576	0.0005507	0.00072	0.0007590	0.000864	0.0008718	0.001008	0.0010052
γ_2	0.000192	0.0002130	0.000256	0.0002587	0.00032	0.0003509	0.000384	0.0004001	0.000448	0.0004650
γ_3	0.000336	0.0003534	0.000448	0.0004241	0.00056	0.0005843	0.000672	0.0006968	0.000784	0.0007589
γ_4	0.000168	0.0001300	0.000224	0.0002453	0.00028	0.0002947	0.000336	0.0003177	0.000392	0.0003684
γ_5	0.00012	0.0001004	0.00016	0.0001717	0.0002	0.0002049	0.00024	0.0002435	0.00028	0.0002671
γ_6	0.000216	0.0002182	0.000288	0.0002809	0.00036	0.0003480	0.000432	0.0003976	0.000504	0.0005031
γ_7	0.000096	0.0001021	0.000128	0.0001085	0.00016	0.0001625	0.000192	0.0001651	0.000224	0.0002265
γ_8	0.000048	0.00002455	0.000064	0.00005066	0.00008	0.00007288	0.000096	0.0001019	0.000112	0.0001115
γ_9	0.0000048	0.000003389	0.0000064	0.000005194	0.000008	0.000007687	0.0000096	0.00001236	0.0000112	0.000009344
γ_{10}	0.000024	0.00002257	0.000032	0.00003663	0.00004	0.00003979	0.000048	0.00004427	0.000056	0.00005813

Table 8: Performance heuristics under estimated values. Variables: same as 1, except for load factor = 1. Use estimates from estimation method example.

6.1. Selecting an Appropriate Model, Solution Method, and Estimation Method

In practice one needs to make a choice which model to use before deciding about the solution method and estimation method. To this end, a balance must be established for the trade-off between a good model, a good estimation method, and a good solution method. If little restriction is put on the cancellation probabilities, heuristics need to be used unless the problem is small. Subsequently, complex cancellation rate functions need to be estimated. On the other hand, under the linear and equal cancellation rates Assumption 2.3 the problem can be solved exactly in an efficient way, if the action space small. However, the assumption that cancellation probabilities are equal for all products is rather strong and not realistic in many cases.

For different model instances different solution methods and estimation methods may be appropriate. There are three dimensions that have to be considered when choosing a solution method: the model choice, the size of the state space, and the size of the action space. When choosing an estimation method we need to consider the model choice and in particular the model choice for the purchase probabilities and cancellation rates.

The efficient three-step estimation Algorithm 4.1 can only be applied to a model that satisfies Assumption 2.2 or 2.3 and the purchase probabilities are modelled by the multinomial logit model. In all other cases the expectation-maximisation algorithm described in AppendixB can be used.

If none of the Assumptions 2.1, 2.2, or 2.3 is satisfied, we can only use the exact solution method to solve the problem. The computation time of this method may be prohibitively long if the state space or the action space is large.

Under the independence of cancellations Assumption 2.1 only (the reformulation of) the exact method described by Equation (3) or the Independence of Cancellations heuristic (IOC) described by Algorithm 3.1 are appropriate. If the state space or the action space is large we use the IOC heuristic and if the state space and the action space are small we can use the exact solution method.

Under the linear cancellation rates Assumption 2.2 we can apply the exact method, the IOC heuristic, and the Linear Cancellation Rates heuristic (LCR). However, numerical results show that LCR outperforms IOC under the linear cancellation rates assumption, so we will not consider IOC here. If both the action space and the state space are small, we will use the exact method. If however either the action space is large or the state space is large we prefer LCR because the exact method is intractable. The action space can be small when the state space is large if n is small but C is large.

Under the linear and equal cancellation rates assumption the problem can be solved exactly using Equation 4. However, if the action space is too large we have to use the LCR heuristic.

6.2. Future Directions: Choice-Based Network Cancellation Model

In practice it is common that fare products consist of multiple night stays/multiple flight legs, so it is worth looking at network revenue management problems. The theory described in this paper is not sufficient to solve these complex problems. However, we provide an outline of a way to approach this problem. The choice-based network model by Liu and van Ryzin [20] could be extended to include cancellations. The network consists of m legs and n fare products, where a fare product is now not only a combination of a price r_j and conditions, but also a combination of legs consumed by the product, e.g. several nights for a hotel stay.

As in the single-resource case we can write the dynamic programming formulation to solve the problem. In the network base model (without cancellations) we have to keep track of the available resources, which is at most a vector of size m . [20] argue that this problem is intractable and provide a choice-based linear programming heuristic (CDLP) to solve the problem in reasonable time. Now if we also consider cancellations we have to keep track of the number of reservations per product. This is a vector of size at most $2^m - 1$, so we need to consider heuristics. One heuristic is to use state space reduction techniques similar to methods described in this paper. Instead of keeping track of purchases of all products we can keep track of the availability of each resource, which reduces the state-space to an m -dimensional vector, as in the network base model. A heuristic similar to CDLP could be used to solve the remaining problem. However, cancellations contribute to the complexity of modelling and solving such heuristic. More research is necessary to investigate an efficient method to model and solve the choice-based network

cancellation model. This topic would enrich the choice-modelling literature, but this is beyond the scope of this paper.

7. Concluding Remarks

540 Cancellations have a big impact on revenue. Our model incorporates cancellations in the customer choice behaviour setting of revenue management. Policies that are optimal in a setting where cancellations are not considered can lead to a revenue loss of up to 20% compared to an optimal policy in our setting where cancellations are considered.

545 The exact solution can not be evaluated for all practical purposes because of the curse of dimensionality. To overcome this problem we introduced three heuristics: (1) the Independence of Cancellations heuristic 3.1, which reduces the state space and can be applied to models with general cancellation rates; (2) the Linear Cancellation Rates heuristic 3.2, which also reduces the state space and is appropriate under the linear cancellations Assumption 2.2; and (3) heuristic 3.3, which is an extension of the Linear Cancellation Rates algorithm and can be used for a hotel with several room types.

550 Two key results in this paper ensure that the customer choice cancellation model can be an effective method for practitioners. First, the Linear Cancellation Rates heuristic from Section 3.4 is fast and gives a solution close to optimal. Second, the parameters of the model can be estimated in a consistent and fast way using Algorithm 4.1.

555 The results of this paper can function as a foundation for several topics of further research. First of all, the effects of cancellation conditions and overbooking policies in combination with customer choice behaviour can now be studied with our customer choice cancellation model. Second, our customer choice cancellation model can be extended to include group bookings and multiple night stays/multiple flight legs, which are also very common in practice. Finally, cancellations can be embedded in other revenue management models to improve performance.

560 8. Acknowledgements

This research was partly funded by the Ministry of Education, Culture and Science (OCW) via the Regional Attention and Action for Knowledge circulation (RAAK), Foundation Innovation Alliance (SIA).

565 We would like to thank the anonymous referees for their constructive input. Their comments on among others the discussion on efficient sets and the estimation method improved the presentation of the manuscript indeed.

AppendixA. List of Proofs

Proof of Proposition 3.1.

First we rewrite Equation (3) by taking all terms that do not depend on S out of the maximisation part. This gives us the following equation:

$$\begin{aligned} \tilde{W}_t(y) = \max_{S \subset N} \{ & \lambda(R(S, y, t) - Q(S)\Delta\tilde{W}_{t-1}(y)) \} \\ & + \gamma y \Delta\tilde{W}_{t-1}(y-1) + \tilde{W}_{t-1}(y) \end{aligned} \quad (\text{A.1})$$

570 with $\Delta\tilde{W}_{t-1}(y) := \tilde{W}_{t-1}(y) - \tilde{W}_{t-1}(y+1)$.

Let $S^* \subset N$ be an inefficient set, i.e., there exist $\alpha(S) \geq 0$ for all $S \subset (N)$ with $\sum_{S \subset N} \alpha(S) = 1$ and

$$Q(S^*) \geq \sum_{S \subset N} \alpha(S)Q(S), \quad R(S^*, y, t) < \sum_{S \subset N} \alpha(S)R(S, y, t).$$

Then we have that

$$\begin{aligned} \lambda \left(R(S^*, y, t) - Q(S^*) \Delta \tilde{W}_{t-1}(y) \right) &< \lambda \left(\sum_{S \subset N} \alpha(S) R(S, y, t) - \sum_{S \subset N} \alpha(S) Q(S) \Delta \tilde{W}_{t-1}(y) \right) \\ &= \sum_{S \subset N} \alpha(S) \lambda (R(S, y, t) - Q(S) \Delta \tilde{W}_{t-1}(y)). \end{aligned}$$

Since $\alpha(S)$ defines a probability, there is at least one $S \subset N$ such that

$$R(S^*, y, t) - Q(S^*) \Delta \tilde{W}_{t-1}(y) < R(S, y, t) - Q(S) \Delta \tilde{W}_{t-1}(y).$$

Hence S^* is not optimal. \square

Proof of Lemma 3.4. The probability that $x_j \in \{0, \dots, y\}$ in state (y, t) is equal to 1. Therefore the probability $p_t^y(x_j = k)$ that $x_j = k$ for some $k \in \{0, \dots, y\}$ if we are in state (y, t) is equal to the probability $\hat{p}_t^y(x_j = k)$ that $x_j = k$ and we are in state y if we are at time t divided by the probability that we are in state y at time t . The latter is equal to

$$\sum_{k=0}^y \hat{p}_t^y(x_j = k),$$

so we only have to find $\hat{p}_t^y(x_j = k)$ for all $k \in \{0, \dots, y\}$.

There are three ways to get to state (y, t) . First, we could come from state $(y+1, t+1)$ and a product was cancelled. The probability that we get to state $(y+1, t+1)$ is $p(y+1, t+1)$. There are two possible values for x_j at time $t+1$ that can lead to $x_j = k$ at time t , namely $x_j = k+1$ at time $t+1$ and product j is cancelled or $x_j = k$ at time $t+1$ and product $i \neq j$ is cancelled. The probability that the former event occurs is equal to

$$p(y+1, t+1) p_{t+1}^{y+1}(x_j = k+1) \gamma_j^{t+1}(k+1).$$

The probability that the latter event occurs is equal to

$$p(y+1, t+1) p_{t+1}^{y+1}(x_j = k) \sum_{i \neq j} \sum_{k'=0}^{y+1-k} p_{t+1}^{y+1}(x_i = k') \gamma_i^{t+1}(k').$$

The second possible state that we could come from to end up in state (y, t) is state $(y, t+1)$ and neither a product was cancelled nor a product was bought. The probability that we get to state $(y, t+1)$ is equal to $p(y, t+1)$. The only possibility to get to $x_j = k$ in state (y, t) if nothing happens is that $x_j = k$ in state $(y, t+1)$. The probability for this event is $p_{t+1}^y(x_j = k)$. The probability that nothing happens in state $(y, t+1)$ under the condition that $x_j = k$ is then equal to

$$1 - \lambda \sum_{j \in S} P_j(S) - \sum_{i \neq j} \sum_{k'=0}^{y-k} p_{t+1}^y(x_i = k') \gamma_i^{t+1}(k') - \gamma_j^{t+1}(k).$$

The third possible state that we could come from to end up in state (y, t) is state $(y-1, t+1)$ and a product was bought. The probability that we get to state $(y-1, t+1)$ is equal to $p(y-1, t+1)$. There are two possible values for x_j at time $t+1$ that can lead to $x_j = k$ at time t , namely $x_j = k-1$ at time $t+1$ and product j is bought or $x_j = k$ at time $t+1$ and product $i \neq j$ is bought. The probability that the former event occurs is equal to

$$p_{t+1}^{y-1}(x_j = k-1) \lambda P_j(S).$$

The probability that the latter event occurs is equal to

$$p_{t+1}^{y-1}(x_j = k) \lambda \sum_{i \in S \setminus \{j\}} P(S, i).$$

Combining all events shows that $\hat{p}_t^y(x_j = k)$ is equal to the stated formula for all k . \square

Proof of Lemma 3.5.

$\bar{x}_j(y, t)$ is the expected number of reservations for product j in state (y, t) , which is equal to

$$\bar{x}_j(y, t) = \sum_{k=0}^y p_t^y(x_j = k)k.$$

$\gamma(y, t)$ is the probability that a product is cancelled in state (y, t) . This is a combination of the cancellation probabilities of individual products. The cancellation probability of product j is equal to

$$\sum_{k=0}^y p_t^y(x_j = k)\gamma_j^t(k).$$

Therefore $\gamma(y, t)$ is equal to

$$\gamma(y, t) = \sum_{j=1}^n \sum_{k=0}^y p_t^y(x_j = k)\gamma_j^t(k).$$

Moreover, under the linear and equal cancellation rates Assumption 2.3 we have that

$$\gamma(y, t) = \sum_{j=1}^n \gamma_j^t \sum_{k=0}^y p_t^y(x_j = k)k = \sum_{j=1}^n \gamma_j^t \bar{x}_j(y, t).$$

$c(y, t)$ are the costs of a cancellation in state (y, t) . If $x_j = k$ then the expected costs are $r_j(t)$. The expected costs for product j are thus

$$\sum_{k=0}^y p_t^y(x_j = k)\gamma_j^t(k)r_j(t).$$

The expected costs for all products are thus

$$\sum_{j=1}^n \sum_{k=0}^y p_t^y(x_j = k)\gamma_j^t(k)r_j(t).$$

If we condition to the event that a cancellation occurs in state (y, t) we see that the expected costs for a cancellation are equal to

$$\begin{aligned} c(y, t) &= \frac{\sum_{j=1}^n \sum_{k=0}^y p_t^y(x_j = k)\gamma_j^t(k)r_j(t)}{\sum_{j=1}^n \sum_{k=0}^y p_t^y(x_j = k)\gamma_j^t(k)} \\ &= \frac{\sum_{j=1}^n \sum_{k=0}^y p_t^y(x_j = k)\gamma_j^t(k)r_j(t)}{\gamma(y, t)}. \end{aligned}$$

Moreover, under the linear and equal cancellation rates Assumption 2.3 we have that

$$c(y, t) = \frac{\sum_{j=1}^n r_j(t)\gamma_j^t \bar{x}_j(y, t)}{\gamma(y, t)}.$$

Finally, $p(y, t)$ is the probability that we are in state (y, t) at time t . First, we could come from state $(y + 1, t + 1)$ and we have a cancellation. The probability that this occurs is equal to

$$p(y + 1, t + 1) \sum_{j=1}^n \sum_{k=0}^{y+1} p_t^{y+1}(x_j = k)\gamma_j^t(k) = p(y + 1, t + 1)\gamma(y + 1, t + 1)$$

Second, we could come from state $(y, t + 1)$ and there was no event. This occurs with probability

$$\left(1 - \lambda \sum_{j \in S} P_j(S) - \gamma(y, t + 1)\right) p(y, t + 1).$$

Third, we could come from state $(y - 1, t + 1)$ and we have a reservation. This occurs with probability

$$\lambda \sum_{j \in S} P_j(S) p(y - 1, t + 1).$$

Combined this gives

$$\begin{aligned} p(y, t) &= p(y + 1, t + 1) \gamma(y + 1, t + 1) + \lambda \sum_{j \in S} P_j(S) p(y - 1, t + 1) \\ &\quad + \left(1 - \lambda \sum_{j \in S} P_j(S) - \gamma(y, t + 1) \right) p(y, t + 1). \end{aligned}$$

575

□

AppendixB. Estimating Parameters

AppendixB.1. Multinomial Logit Model

The *multinomial logit model* (MNL) is a method to describe a discrete probability distribution, given a set of independent variables. It is commonly used to predict the outcome of statistical classification problems, such as political elections (see for example Dow and Endersby [9]).

Assume that customers can choose among n products and the no-purchase alternative. Each product j provides a utility U_j for each customer, i.e., the value of product j to the customer. The no-purchase alternative provides a utility U_0 . Each customer chooses alternative $j \in N \cup \{0\}$ such that his utility is maximised. The MNL assumes that the utilities are random variables of the form

$$U_j = u_j + \xi_j,$$

where u_j is the expected value of U_j and ξ_j is a random draw from the Gumbel distribution² with location equal to Euler's constant and scale equal to one, such that the expected value is zero. Under these assumptions it can be shown (see Ben-Akiva and Lerman [5]) that the choice probabilities are equal to

$$P_j(S) = \frac{e^{u_j}}{\sum_{i \in S} e^{u_i} + e^{u_0}}.$$

The mean utility u_j of a product or the no-purchase option depends on $k \in \mathbb{N}$ different attributes. The values of the attributes are stored in the k -dimensional vector Z_j . The corresponding weights of each attribute are stored in the k -dimensional vector β . Hence the mean utility u_j for product j is given by

$$u_j = \beta^\top Z_j.$$

AppendixB.2. Complete Knowledge

Suppose that we have complete knowledge on arrivals, purchases and no-purchases, offer sets, and cancellations. Then we can estimate the parameters in the following way. Define $a_\lambda(t)$, $\{a_j(t)\}_{j \in N}$, and $a(t)$ by

$$\begin{aligned} a_\lambda(t) &= \begin{cases} 1 & \text{arrival at time } t, \\ 0 & \text{otherwise.} \end{cases} & a(t) &= \begin{cases} 1 & \text{no arrival or cancellation at time } t, \\ 0 & \text{otherwise.} \end{cases} \\ a_j(t) &= \begin{cases} 1 & \text{product } j \text{ cancelled at time } t, \\ 0 & \text{otherwise.} \end{cases} & &= (1 - a_\lambda(t)) \prod_{j \in N} (1 - a_j(t)). \end{aligned}$$

²The cumulative distribution function of the *Gumbel distribution* is given by

$$F(x|\mu, \beta) = e^{-e^{-(x-\mu)/\beta}},$$

where μ is called the *location* and β is called the *scale* parameter. The expected value is equal to $\mu + \gamma\beta$, where $\gamma \approx 0.5772$ is Euler's constant.

The parameters defined above are indicators to specify which event happened in time period t . Note that exactly one indicator is equal to 1 at time t and that this is in agreement with the assumption that only one event happens in each time period. In order to ensure a unique solution we set $\alpha = 0$ for convenience. Let $j(t)$ be the product that is purchased at time t under the condition that a customer arrived and $j(t) = 0$ if we observed an arrival but no purchase. Then $\lambda P_{tj(t)}(0, \beta, S_t, Z_t)$ is the probability that $j(t)$ is purchased at time t , or the probability that a customer arrived and purchased nothing if $j(t) = 0$. If product j was cancelled at time t then the probability that this occurs is $\gamma_j(x_j)$. If no event took place this occurs with probability $1 - \lambda - \sum_{j \in N} \gamma_j(x_j)$. Let D be the set of time periods. Then the likelihood function L is given by

$$L(\beta, \gamma, \lambda | a, x, S_t, Z_t) = \prod_{t \in D} [\lambda P_{tj(t)}(0, \beta, S_t, Z_t)]^{a_\lambda(t)} \prod_{j \in N} \gamma_j(x_j)^{a_j(t)} \cdot \left[1 - \lambda - \sum_{j \in N} \gamma_j(x_j) \right]^{a(t)}$$

The log-likelihood function LL is then

$$\begin{aligned} LL(\beta, \gamma, \lambda | a, x, S_t, Z_t) = \log(L) = & \sum_{t \in D} \left[a_\lambda(t) \log(P_{tj(t)}(\alpha, \beta, S_t, Z_t)) + a_\lambda(t) \log(\lambda) \right. \\ & \left. + \sum_{j \in N} a_j(t) \log(\gamma_j(x_j)) + a(t) \log \left(1 - \lambda - \sum_{j \in N} \gamma_j(x_j) \right) \right]. \end{aligned}$$

Note that LL is separable in β and (γ, λ) . Maximising the separated log-likelihood functions with respect to β and (γ, λ) gives estimates $\hat{\beta}$ and $(\hat{\gamma}, \hat{\lambda})$, respectively.

Appendix B.3. Incomplete Knowledge: Expectation Maximisation Algorithm

Suppose we can not observe the event that a customer arrives but does not make a purchase. Then we do not know the values of $a_\lambda(t)$ and $a(t)$. One way to overcome this problem is to apply the well known expectation maximisation algorithm (see Dempster et alii, 1977 [8]). Define \mathcal{P} as the set of time periods where a purchase is observed and \mathcal{C}_j the set of time periods where a cancellation for product j is observed. Then we can rewrite LL to

$$\begin{aligned} LL(\beta, \gamma, \lambda | x, S_t, Z_t) = & \sum_{t \in \mathcal{P}} [\log(\lambda) + \log(P_{tj(t)}(0, \beta, S_t, Z_t))] + \sum_{j \in N} \sum_{t \in \mathcal{C}_j} \log(\gamma_j(x_j)) \\ & + \sum_{t \notin \mathcal{P}, t \notin \mathcal{C}_j} \log \left(1 - \lambda - \sum_{j \in N} \gamma_j(x_j) \right) \\ & + a_\lambda(t) \left[\log(\lambda) + \log(P_{t0}(0, \beta, S_t, Z_t)) - \log \left(1 - \lambda - \sum_{j \in N} \gamma_j(x_j) \right) \right]. \end{aligned}$$

585 We know the values for $a_\lambda(t)$ for $t \in \mathcal{P}$ and $t \in \mathcal{C}_j$ for all j . However, we do not know the values for $a_\lambda(t)$ for $t \notin \mathcal{P}$ and $t \notin \mathcal{C}_j$ for all j . If we have estimates $\hat{\beta}$, $\hat{\gamma}$, and $\hat{\lambda}$ we can use Bayes theorem to calculate $\hat{a}_\lambda(t)$ (similar to Talluri and van Ryzin [27]):

$$\begin{aligned} \hat{a}_\lambda(t) &= \mathbb{E}[a_\lambda(t) | t \notin \mathcal{P}, t \notin \mathcal{C}_j, \hat{\beta}, \hat{\lambda}, \hat{\gamma}] \\ &= \mathbb{P}[a_\lambda(t) = 1 | t \notin \mathcal{P}, t \notin \mathcal{C}_j, \hat{\beta}, \hat{\lambda}, \hat{\gamma}] \\ &= \frac{\mathbb{P}[t \notin \mathcal{P}, t \notin \mathcal{C}_j | a_\lambda(t) = 1, \hat{\lambda}, \hat{\beta}, \hat{\gamma}] \cdot \mathbb{P}[a_\lambda(t) = 1 | \hat{\lambda}, \hat{\beta}, \hat{\gamma}]}{P[t \notin \mathcal{P}, t \notin \mathcal{C}_j | \hat{\lambda}, \hat{\beta}, \hat{\gamma}]} \\ &= \frac{P_{t0}(\hat{0}, \hat{\beta}, S_t, Z_t) \cdot \hat{\lambda}}{1 - \hat{\lambda} P_{t*}(\hat{0}, \hat{\beta}, S_t, Z) - \sum_{j \in N} \hat{\gamma}_j(x_j)}. \end{aligned} \tag{B.1}$$

With \hat{a}_λ we can approximate $a(t)$ by

$$\begin{aligned} \hat{a}(t) &= (1 - \hat{a}_\lambda(t)) \prod_{j \in N} (1 - a_j(t)) \\ &= \begin{cases} 0 & \text{if a product was cancelled in time period } t, \\ 1 - \hat{a}_\lambda(t) & \text{otherwise.} \end{cases} \end{aligned}$$

Under $\hat{a}_\lambda(t)$ the expected log-likelihood function is

$$\begin{aligned} \mathbb{E}[LL|\hat{\beta}, \hat{\lambda}, \hat{\gamma}] &= \sum_{t \in \mathcal{P}} [\log(\lambda) + \log(P_{tj(t)}(\beta, Z, S_t))] + \sum_{j \in N} \sum_{t \in \mathcal{C}_j} \log(\gamma_j(x_j)) \\ &+ \sum_{t \notin \mathcal{P}, t \notin \mathcal{C}_j} \log\left(1 - \lambda - \sum_{j \in N} \gamma_j(x_j)\right) \\ &+ \hat{a}_\lambda(t) \left[\log(\lambda) + \log(P_{t0}(\beta, Z, S_t)) - \log\left(1 - \lambda - \sum_{j \in N} \gamma_j(x_j)\right) \right]. \end{aligned} \quad (\text{B.2})$$

Maximising this function with respect to (β, λ, γ) gives new estimates $(\hat{\beta}, \hat{\lambda}, \hat{\gamma})$. These values in turn can be used to find a new estimate $\hat{a}_\lambda(t)$, etc. The algorithm is summarised in Algorithm [AppendixB.1](#).

Algorithm AppendixB.1. Expectation-Maximisation

1. Start with initial parameters $(\hat{\lambda}, \hat{\beta}, \hat{\gamma})$.
2. *Expectation* Estimate $\hat{a}_\lambda(t)$ using Equation (B.1).
3. *Maximisation* Estimate $(\hat{\lambda}, \hat{\beta}, \hat{\gamma})$ using $\hat{a}_\lambda(t)$ and Equation (B.2).
4. Go to step 2 unless we hit a stopping criterion.

References

- [1] N. Aydin, Ş.İ. Birbil, J.B.D. Frenk, and N. Noyan. Single-leg airline revenue management with overbooking. *Transportation Science*, 47(4):560–583, 2012.
- [2] P.P. Belobaba. *Air Travel Demand and Airline Seat Inventory Management*. PhD Thesis, Massachusetts Institute of Technology, Cambridge, MA, 1987.
- [3] P.P. Belobaba. Airline yield management: An overview of seat inventory control. *Transportation Science*, 21:63–73, 1987.
- [4] P.P. Belobaba. Application of a probabilistic decision model to airline seat inventory control. *Operations Research*, 37:183–197, 1989.
- [5] M. Ben-Akiva and S.R. Lerman. *Discrete Choice Analysis*. MIT Press, Cambridge, MA, 1985.
- [6] D. Bertsimas and I. Popescu. Revenue management in a dynamic network environment. *Transportation science*, 37(3):257–277, 2003.
- [7] Kym Clarke. Case study: Implementing a revenue management system at malaysia airlines. *Journal of Revenue and Pricing Management*, 3(1):41–48, 2004.
- [8] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society Series B (Methodological)*, 39(1):1–38, 1977.
- [9] J.K. Dow and J.W. Endersby. Multinomial probit and multinomial logit: a comparison of choice models for voting research. *Electoral Studies*, 23:107–122, 2004.
- [10] A. Erdelyi and H. Topaloglu. A dynamic programming decomposition method for making overbooking decisions over an airline network. *INFORMS Journal on Computing*, 22(3):443–456, 2010.
- [11] M.K. Geraghty and E. Johnson. Revenue management saves national car rental. *Interfaces*, 27(1):107–127, 1997.
- [12] F. Guerriero, G. Miglionico, and F. Olivito. Strategic and operational decisions in restaurant revenue management. *European Journal of Operational Research*, 237(3):1119–1132, 2014.

- [13] D. Hopman. Revenue management in small-scale hotels. *Working Paper*, 2013.
- [14] D. C. Ilescu, L. A. Garrow, and R. A. Parker.
- [15] S.E. Kimes. Revenue management on the links: applying yield management to the golf-course industry. *The Cornell Hotel and Restaurant Administration Quarterly*, 41(1):120–127, 2000.
- [16] S. Kunnumkal, K. Talluri, and H. Topaloglu. A randomized linear programming method for network revenue management with product-specific no-shows. *Transportation Science*, 46(1):90–108, 2012.
- [17] C.B.G. Langeveld. *Economie van het theater: vestiging, prijsvorming en economies of scale in een wereld met bezieling*. Erasmus University Rotterdam, 2006.
- [18] C.J. Lautenbacher, S. Stidham, and J. Subramanian. Airline yield management with overbooking, cancellations, and no-shows. *Transportation Science*, 33(2):147–167, 1999.
- [19] K. Littlewood. Forecasting and control of passenger bookings. *AGIFORS 12-th Annual Symposium Proceedings*, pages 95–128, 1972.
- [20] Q. Liu and G. van Ryzin. On the choice-based linear programming model for network revenue management. *Manufacturing & Service Operations Management*, 10(2):288–310, 2008.
- [21] B. Maddag, L. Moussawi-Haidar, M. El-Taha, and H. Rida. Dynamic cruise ship revenue management. *European Journal of Operational Research*, 207(1):445–555, 2010.
- [22] J.I. McGill and G.J. van Ryzin. Revenue management: Research overview and prospects. *Transportation Science*, 33(2):233–256, 1999.
- [23] J. Meissner and A. Strauss. Network revenue management with inventory-sensitive bid prices and customer choice. *European Journal of Operational Research*, 216(2):459–468, 2012.
- [24] J.P. Newman, L. Garrow, M. Ferguson, T.L. Jacobs, and H. Purnomo. Estimation of choice-based models using sales data from a single firm. *Working Paper*, 16(2):184–197, 2014.
- [25] M.L. Puterman. *Markov Decision Processes*. Wiley, 1994.
- [26] A. Strauss and K. Talluri. Tractable consideration set structures and new inequalities for choice network revenue management. *Working Paper*, 2012.
- [27] K. Talluri and G. van Ryzin. Revenue management under a general choice model of consumer behavior. *Management Science*, 50(1):15–33, 2004.
- [28] G. van Ryzin and G. Vulcano. An expectation-maximization method to estimate a rank-based choice model of demand. *Technical Note*, 2013.
- [29] B. Vinod. Unlocking the value of revenue management in the hotel industry. *Journal of Revenue and Pricing Management*, 3(2):178–190, 2004.