Mathematical Programming Models for Revenue Management under Customer Choice

Lijian Chen
Department of Industrial Engineering
J.B. Speed School of Engineering
University of Louisville
Louisville, KY 40292
lijian.chen@louisville.edu

Tito Homem-de-Mello
Department of Industrial Engineering and Management Sciences
Northwestern University
Evanston, IL 60208
tito@northwestern.edu

April 15, 2009

Abstract

We study a network airline revenue management problem with discrete customer choice behavior. We discuss a choice model based on the concept of *preference orders*, in which customers can be grouped according to a list of options in decreasing order of preference. If a customer's preferred option is not available, the customer moves to the next choice on the list with some probability. If that option is not available, the customer moves to the third choice on the list with some probability, and so forth until either the customer has no other choice but to leave or his/her request is accepted. Using this choice model as an input, we propose some mathematical programs to determine seat allocations. We also propose a post-optimization heuristic to refine the allocation suggested by the optimization model. Simulation results are presented to illustrate the effectiveness of our method, including comparisons with other models.

1 Introduction

Revenue management involves the application of quantitative techniques to improve profits by controlling the prices and availabilities of various products that are produced with scarce resources. Perhaps the best known revenue management application occurs in the airline industry, where the products are tickets (for itineraries) and the resources are seats on flights. In view of many successful applications of revenue management in different areas, this topic has received considerable attention in the past few years both from practitioners and academics. The recent book by Talluri and van Ryzin [2004b] provides a comprehensive introduction to this field, see also references therein. A common way to model the airline booking process is as a sequential decision problem over a fixed time period, in which one decides whether each request for a ticket should be accepted or rejected. A typical assumption is that one can separate demand for individual *itinerary-fare pairs*; that is, each request is for a particular type of ticket on a particular itinerary, and yields a pre-specified fare. These combinations are called classes. Typically, a class is determined by particular constraints associated with the ticket rather than the physical seat. For example, a certain class may require a 14-day advance purchase, or a Saturday night stay, etc.

The existence of different classes reflects different customer behaviors. The classical example is that of customers traveling for leisure and those traveling on business. The former group typically books in advance and is more price-sensitive, whereas the latter behaves in the opposite way. Airline companies attempt to sell as many seats as possible to high-fare paying customers and at the same time avoid the potential loss resulting from unsold seats. In most cases, rejecting an early (and lower-fare) request saves the seat for a later (and higher-fare) booking, but at the same time that creates the risk of flying with empty seats. On the other hand, accepting early requests raises the percentage of occupation but creates the risk of rejecting a future high-fare request because of the constraints on capacity.

In practice, customers are not completely segmented into classes. A customer may be able to afford an expensive ticket for a particular class on a particular itinerary, but if a lower-priced ticket is available on the same flight, he or she will likely take it. Generally, customers have their own preferences among all the options.

Both academics and practitioners realize that a customer choice model for the network environment is an important step forward. Cooper et al. [2006] show that models that ignore customer choice may lead to policies that, when used repeatedly, drive revenues *down*, a phenomenon they call the "spiral-down effect." There are a few papers in the literature dealing with the customer choice behavior in airline setting. Belobaba [1989] models a heuristics called the expected marginal seat revenue (EMSR) to allocate seats to the different classes. The idea assumes that every cus-

tomer has a non-homogenous probability distribution on choosing multiple tickets. Although this work is based on the independent-demands assumption across classes, a "buying up-down" concept is introduced for modeling customer choice behavior with discounted seats. Belobaba and Weatherford [1996] and Bodily and Weatherford [1995] build their heuristics for the modified EMSR model for solving single-leg problems with customer choice. Brumelle et al. [1990] provide optimal conditions for the single-leg customer buy-up model. Talluri and van Ryzin [2004a] provide a thorough analysis of a single-leg model under a general discrete choice model of demand.

Despite the advances provided by the above works, those models have a major drawback — namely, that they deal with single-leg models. The operations in major airlines is largely based on the network environment which is expected to be a large and multiple resources problem; however, the extension of single-leg models to a network environment is far from clear, and can lead to problems that are too big to solve.

Some recent work has addressed the issue of customer choice in a network environment. Zhang and Cooper [2005] discuss a model for customer choice among different departure times between the same city pair. By assuming that customers only make a choice among alternatives within the same booking classes, they build a dynamic programming model for the problem and develop bounds and approximations for it. van Ryzin and Vulcano [2008] introduce a way to model customer behavior whereby customers choose substitutes if their preferred choices are not offered. This is accomplished by associating a preference list to each customer. Such a list contains the alternatives that are acceptable by that customer, ranked in the order of preference. They combine that demand model with a virtual nesting control approach, and solve the resulting revenue maximization problem using stochastic approximation techniques. Because the problem is non-convex, the proposed method yields local optimizers. Gallego and Hu [2006] approach the issue of customer choice from the viewpoint of dynamic pricing, and study a game-theoretical model for the problem.

The model that is closest to ours is the one described in Gallego et al. [2004] and Liu and van Ryzin [2008]. The authors model customer choice by means of *offer sets* with associated probabilities. This works as follows: Let J be the set of itinerary-fare classes available and let $\mathcal{S}$ be the collection of all subsets of J. The idea of this approach is to associate to each set $S \in \mathcal{S}$ the conditional probability that a class $j \in S$ is chosen by an arriving customer given that only the classes in S are offered. Thus, such a model captures the substitutability of products by consumers.

To optimize the system, Gallego et al. [2004] propose a linear programming model where the decision variables represent the amount of time each set $S \in \mathcal{S}$ is made available. Although the model is novel and yields good results, it suffers from a drawback — namely, the fact that the number of decision variables grow exponentially fast with the number of classes, which is easy to see since $|\mathcal{S}| = 2^{|J|}$. To overcome that problem, Liu and van Ryzin [2008] propose a column

generation algorithm together with a decomposition heuristic. The algorithm assumes that the choice probabilities are generated from a multinomial logit model. In such a model, customers are divided into segments (e.g., budget-sensitive and time-sensitive customers). Each segment corresponds to a vector of "preference weights" for each itinerary-fare class, and the probabilities are calculated from the ratios. For example, if the weights are $v_A = 3,\ v_B = 2,\ v_C = 1$, then $P(A|A) = 3/3 = 1,\ P(A|AB) = 3/(3+2) = 3/5,\ P(A|AC) = 3/(3+1) = 3/4\ P(A|ABC) = 3/(3+2+1) = 1/2$ and so on, where $P(A|AB)$ denotes the probability that a customer chooses A given that A and B are offered, and similarly for the other sets.

A crucial assumption for the efficiency of the algorithm in Liu and van Ryzin [2008] is that the segments described above are *disjoint*, i.e., no itinerary-fare class belongs to more than one segment. Such an assumption can be difficult to verify in practical problems, where many segments may exist. Miranda Bront et al. [2007] study the general case, where segments need not be disjoint. They show that the problem is NP-Hard, and propose a greedy heuristic to overcome the complexity of an exact algorithm. Zhang and Adelman [2007] study an approach to those models based on approximate dynamic programming techniques.

A central concept to our work is that of a *customer preference order* to describe the gross customer behavior regarding the order of classes for which they try to purchase tickets. If the customer's first choice is not available, she either tries her second choice or decides not to purchase anything. If her second choice is not available either, again she moves to her next choice or decides not to purchase anything, and so on. We model each customer's decision made at each step — i.e., between trying the next choice or leaving the system — as a Bernoulli random variable with known probability. Our basic assumption is that customers can be *grouped* by similar purchasing preferences.

The notion of preference orders is related to the idea of preference lists described in van Ryzin and Vulcano [2008], where each arriving customer comes with a list of options to be followed in case that customer's preferred choices are not available. However, in that case the order in which the customers arrive matter. For example, consider a preference order with two classes $A \to B$ and transition probability p_{AB}, and suppose that at most one customer can be booked from each class. Suppose also that exactly two customers request a ticket. In our model, we book both classes A and B with probability p_{AB}, and book class A only with probability $1 - p_{AB}$. In the model in van Ryzin and Vulcano [2008], we would have one customer with preference list A only (call it customer 1) and another with preference list A, B (call it customer 2). In that case, we book both classes A and B as long customer 1 arrives before customer 2. So, as long as that event happens with probability p_{AB} the two models are equivalent, but it is unclear whether this argument can be extended to a more general setting. We conjecture that the equivalence holds when the arrival

process is a homogeneous Poisson process, but since our model does not require that assumption we do not dwell on this point.

Our concept of preference orders is also very similar to that of segments described in Liu and van Ryzin [2008] and Miranda Bront et al. [2007]. The transition probabilities within a preference order can be viewed as a particular case of the offer set probabilities. To illustrate that point, consider a preference order with three classes $A \to B \to C$ and transition probabilities p_{AB}, p_{BC}, and let us express this in terms of offer set probabilities. We have

$$\begin{aligned} &P(A|A) = 1, \quad P(A|AB) = 1, \quad P(A|AC) = 1, \quad P(A|ABC) = 1, \\ &P(B|B) = p_{AB}, \; P(B|BC) = p_{AB}, \; P(C|C) = p_{AB}\, p_{BC}. \end{aligned}$$

It is clear however from the above example that our choice model cannot be expressed as a multinomial logit. This is a major difference between our work and that of Liu and van Ryzin [2008] and Miranda Bront et al. [2007], where the multinomial logit model is an important component to allow for the derivation of column generation algorithms. Moreover, we do not assume that the preference orders are disjoint as in Liu and van Ryzin [2008]. Another major difference is the optimization model — we assume that the demands for each preference order are independent with known distributions, and propose both deterministic as well as stochastic programming models to determine the number of seats to be allocated to each possible choice within each preference order. Since we do not make the assumption that preference orders are disjoint, such a solution yields multiple values for the same class. To overcome that issue we introduce heuristic algorithms — called *backup heuristics* — that process the solution from the model to determine a single allocation for each class.

In summary, the major contribution of this paper is the introduction of mathematical programming models for the discrete customer choice problem. Although the problem can be difficult to solve due to the large number of scenarios, we propose an approximating model that can be solved considerably faster. Our numerical experiments suggest that the approximation yields good results. We also propose a heuristic to implement the solution given by the model.

The remainder of the paper is organized as follows. We introduce the construction of preference orders in Section 2. In Section 3 we propose an optimization problem for the preference order model as well as an approximation to that problem that leads to a much smaller scenario tree. We show simulation results that suggest that the approximation can provide results that are close to those obtained with the original process. In Section 4 we describe our mathematical programming models and discuss solution techniques. In Section 5 we compare our models with the traditional origin-destination seat allocation models that assume *independent* demands for each class. In Section 6 we propose our *backup* heuristic to properly implement seat allocations resulting from solving the

mathematical programs. Numerical results are presented in Section 7 to illustrate the ideas set forth in the paper.

2 Model Description

We consider an airline company that has L combinations of origin and destinations (OD), which we index by $i = 1, \ldots, L$. An *OD combination* is the set of all the directed itineraries connecting the origin and the destination. For each specific OD combination, say the ith one, the company has n_i itineraries. For each itinerary there are k_i booking classes. Therefore, there are $n_i k_i$ booking classes available to complete this trip. To illustrate, consider for example an OD combination, say from city A to city B, that can be completed through either one non-stop flight or three different one-stop flights. Then for this OD combination, there are four possible itineraries, that is $n_1 = 4$. For all those flights, we have only 3 booking classes available, say first class, business, and coach class. Thus, the customers traveling from city A to city B have $n_1 k_1 = 4 \times 3 = 12$ possible choices to complete the trip.

An individual customer has personal preference in purchasing a ticket for the ith OD combination. The customer will (perhaps involuntarily) arrange some of the $n_i k_i$ available choices in an order from the most to the least preferred tickets. The list constructed by the customer is called a *customer preference order*. Theoretically, for each OD combination i there are $\sum_{\ell=1}^{n_i k_i} \ell!$ possible preference orders constructed. This is obviously beyond computational capability even for small models — indeed, in our small example above, the number of possible preference orders has order of magnitude 10^9. It is reasonable to think, however, that the actual number of preference orders in the model is much smaller because only a few selected preference orders may be sufficient to describe the possible customer behaviors. For instance, in a small problem we may have three preference orders, corresponding respectively to budget-sensitive customers, time-sensitive customers, and time-sensitive customers with a budget cap. In practical problems more preference orders would be identified, constructed, and integrated into the airline company operation, but the total number could still be relatively small.

During the booking process, a customer who has her first options rejected might give up booking because either the price or the time schedule becomes unattractive to her. As a result, some customers might leave the booking for other alternatives, such as those offered by a competitor. Assuming that customers behave independently of each other and have the same probability of moving on to the next option in the preference order, it follows that the number of customers who choose to stay can be modeled by random variables with binomial distribution.

We illustrate the leave-or-stay process with the following example. Consider a preference order

with three booking classes, 1, 2, and 3 arranged in the order of decreasing customer interest. The transition probability from class 1 in the preference order to class 2 is p_1, and from class 2 in this order to class 3 is p_2. Denote the total demand for this customer group by ξ_1 and the allocations for classes 1, 2, and 3 respectively by x_1, x_2, and x_3. Then, $\min(\xi_1, x_1)$ customers get their first-choice tickets, so $\xi_1 - \min(\xi_1, x_1) = [\xi_1 - x_1]_+$ customers face the choice between staying or leaving (here and thereafter, $[a]_+$ denotes $\max(a, 0)$). Thus, the number of customers who actually request class 2 is $\xi_2 := \mathcal{B}([\xi_1 - x_1]_+, p_1)$ (where $\mathcal{B}(n, p)$ denote the binomial distribution with parameters n and p). Likewise, $\xi_3 := \mathcal{B}([\xi_2 - x_2]_+, p_2)$ is the number of remaining customers who choose to stay and request a ticket for class 3. The process is illustrated in Figure 1.

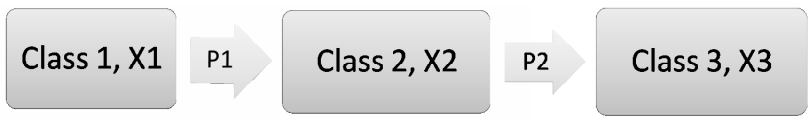

Figure 1: A buy-up example

In the general model, we consider a network with m legs, and denote the m-dimension leg capacity vector by c. There are $n_i k_i$ itinerary-fare classes for the ith OD combination. The number of preference orders for ith OD combination is S_i; for the sth preference order of that OD combination, we denote the actual number of choices by R_{is}. For example, for the preference order displayed in Figure 1 we have $R_{is} = 3$.

The decision variables are $\{x_{isr}\}$, indicating the seat allocation for rth choice inside sth preference order for the ith OD combination. Note that the seat allocation is assigned not only by booking classes but also by preference orders. When preference orders are not considered — this is the case with the standard linear program model for network revenue management, see e.g. Talluri and van Ryzin [2004b] — there are $\sum_{i=1}^{L} n_i k_i$ decision variables. In our model, the total number of decision variables is $\sum_{i=1}^{L} \sum_{s=1}^{S_i} R_{is}$. That is, we refine booking class allocations by constructing preference orders. We also have an $m \times (\sum_{i=1}^{L} \sum_{s=1}^{S_i} R_{is})$ matrix A which reflects the complete network leg structure by seat allocations. Such a matrix is constructed in a similar fashion to the matrix in standard OD models, i.e., $A_{jk} = 1$ if the itinerary corresponding to the triple (OD pair, preference order, choice within the preference order) indexed by k uses leg j.

Denote the demand for the sth preference order of the ith OD combination by ξ_{is1} and the demand thereafter (i.e., for the subsequent choices within that preference order) by $\xi_{isr} = \mathcal{B}([\xi_{is(r-1)} - x_{is(r-1)}]_+, p_{is(r-1)})$, $r = 1, \ldots, R_{is}$. Here, $p_{is(r-1)}$ is the "buy-up" probability. By construction, we have $p_{isr} > 0$, $r = 1, \ldots, R_{is} - 1$, and $p_{isR_{is}} = 0$. In practice, the buy-up probabilities must be estimated using some quantitative method; for example, Andersson [1998] and Algers and Beser

[2001] describe some techniques used in a project at Scandinavian Airlines Systems (SAS). Our work focuses on modeling, solving, and implementing an *optimization* model, so we will not discuss the probability estimation approaches; rather, we treat those probabilities as model inputs.

3 The Optimization Model

Using the notation defined in Section 2, we now exhibit the optimization model analyzed in this paper. The goal is to determine the seat allocations x_{isr} for each triple (OD pair, preference order, choice within the preference order) that maximize the expected accrued revenue. Below, f_{isr} indicates the fares corresponding to each of those triples.

$$
\begin{aligned}
&\max_x \mathbb{E}\left[\sum_{i=1}^{L}\sum_{s=1}^{S_i}\sum_{r=1}^{R_{is}} f_{isr}\min(\xi_{isr}, x_{isr})\right] \\
\text{Subject to: } & Ax \le c && \text{(P)} \\
& x_{isr} \in \mathbb{Z}^+ \\
& \xi_{is(r+1)} = \mathcal{B}([\xi_{isr} - x_{isr}]_+, p_{isr}), && (3.1) \\
& \qquad\qquad i = 1, \ldots, L,\ s = 1, \ldots, S_i,\ r = 1, \ldots, R_{is} - 1.
\end{aligned}
$$

Note that the above model resembles the classical simple-recourse stochastic program used in traditional revenue management models, sometimes called the probabilistic nonlinear programming model (PNLP); see, for instance, Talluri and van Ryzin [2004b] and references therein for the basic ideas, and de Boer et al. [2002] and Chen and Homem-de-Mello [2008] for further analysis. A major difference, however, is the introduction of the recursive equation (3.1) to define the demand corresponding to each triple (OD pair, preference order, choice within the preference order) — in the classical PNLP model, the demand for each itinerary-fare combination is assumed to be exogenous, which does not allow for the modeling of customer choice.

3.1 Model Simplification

The presence of equation (3.1) in problem (P) makes it a nonlinear nonconvex stochastic program. Moreover, (P) is likely to have a large number of scenarios. For instance, consider a small setting with just one OD combination for a one-way trip, $L = 1$. There are only $n_1 = 2$ itineraries. For either itinerary, there are $k_1 = 2$ classes, such as high-fare and low-fare classes. Then, there are $n_1 k_1 = 2 \times 2 = 4$ choices for all customers. We also assume there are only two preference orders, one for budget-sensitive customers and the other one for time-sensitive customers. The four booking

classes are itinerary 1 high-fare class (Class 1), itinerary 1 low-fare class (Class 2), itinerary 2 high-fare class (Class 3), and itinerary 2 low-fare class (Class 4). Suppose itinerary 1's departure time is more attractive than that of the itinerary 2. As a result, the fare level for the comparable class in itinerary 2 is lower than that of itinerary 1. Therefore, the preference order for time-sensitive travelers would be Class 2, Class 1, Class 4, and finally Class 3. For budget-sensitive travelers, the preference order should be Class 4, Class 2, Class 3, and finally Class 1.

To give an idea of how large the number of scenarios can be even for such a small problem, consider the case where the exogenous demand $\xi_1 = (\xi_{11}, \xi_{12})^T$ is deterministic and equal to $(20, 5)^T$, and suppose the seat allocation is $x = (x_{111}, x_{112}, x_{113}, x_{114}, x_{121}, x_{122}, x_{123}, x_{124})^T = (8, 3, 1, 1, 1, 1, 3, 2)^T$. Even though the exogenous demand is deterministic, the size of the scenario tree grows because of the leave-or-stay process. For example, for the budget-sensitive customers' preference order, since $\xi_{11} = 20$ and $x_{111} = 8$ there are 12 customers who would have to make a choice between trying the next option (Class 2) or leaving. Thus, the corresponding binomial random variable has 13 scenarios for the allocation x_{112}. Since $x_{112} = 3$, up to 10 customers may try the third option (Class 3), but the actual number depends on that 13-scenario random variable as well as on the binomial process describing the move from the second to the third option. The number of scenarios for the budget-sensitive preference order can be calculated as 178. Similarly, for the time-sensitive preference order, the number of scenarios is 11. Therefore, for this simple problem, which is shown in Figure 2, the total number of scenarios is $178 \times 11 = 1958$.

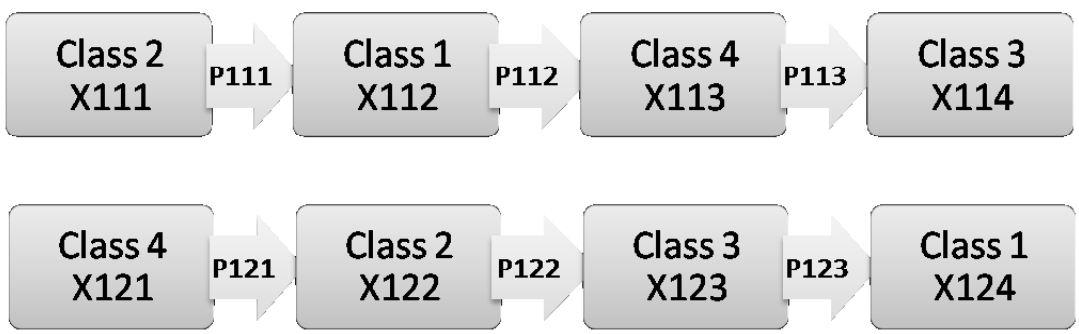

Figure 2: Example for the huge scenario tree

It is clear from the above discussion that it can be difficult to solve (P). In order to circumvent that problem, we shall consider an approximation to (P) whereby the binomial random variables are replaced with their expectations and the integrality constraint is relaxed. Note that the resulting problem is still stochastic, since the incoming demands (ξ_{is1}) are random. Thus, we obtain:

$$\max_x \mathbb{E}\left[\sum_{i=1}^{L}\sum_{s=1}^{S_i}\sum_{r=1}^{R_{is}} f_{isr}\min(\xi_{isr}, x_{isr})\right]$$

$$\text{Subject to: } Ax \le c \tag{PA}$$

$$x_{isr} \ge 0$$

$$\xi_{is(r+1)} = [\xi_{isr} - x_{isr}]_+ \, p_{isr}, \tag{3.2}$$

$$i = 1, \dots, L, \;\; s = 1, \dots, S_i, \;\; r = 1, \dots, R_{is} - 1.$$

3.2 Examples

We show now some numerical examples to compare problems (P) and (PA).

Example 1. *Consider an OD combination with three choices available at prices of \$200, \$400, and \$500. To describe discrete customer choice, we have two preference orders with deterministic demands information. The capacity of the plane is 15 seats. In this case, problem* (P) *can be solved*

Preference Orders	Demands	Transition Probability	Fare Levels (\$)
1	20	0.3, 0.7	200, 400, 500
2	8	0.8, 0.6	500, 400, 200

Table 1: Parameters for Example 1.

by enumerating the entire scenario tree. One obtains the same optimal solution for both models, $(1\ 6\ 0\ 8\ 0\ 0)^T$ *with the objective value equal to \$6,227 for problem* (P) *and \$6,600 for problem* (PA).

In the above example problem (PA) has a higher objective value than (P) at the same solution. The converse may occur as well. For instance, consider the case of one preference order in which customers have three choices, with buy-up probabilities all equal to 0.5. Suppose the exogenous demand is deterministic with $\xi = 4$. The fares are $f = (\$100, \$150, \$400)$. Then, for the seat allocation $x = (2, 1, 2)$, it is easy to calculate the objective value of (P) as $2f_1 + 0.5 \times f_2 + 0.125 \times f_2 + 0.125 \times [f_2 + f_3] = \$362.5$, whereas the objective value of (PA) is $2f_1 + f_2 = \$350$.

Another type of comparison can be conducted in terms of comparing *critical points* for a given seat allocation x. We need the following definition 1.

Definition 1. In a given preference order with R choices, there exists a unique q^* such that

$$q^* \;=\; \max\{n : \xi_n \ge x_n, \; \xi_n > 0, \; 1 \le n \le R\}$$

($q^* = 0$ if $\xi_1 < x_1$), so we have $\xi_{q^*+1} < x_{q^*+1}, \xi_{(q^*+2)}, \ldots, \xi_R = 0$. We call q^* the *critical point* for this seat allocation x.

In words, the critical point is the place within the preference order up to which the residual demand is sufficient to cover the seat allocation. Notice that q^* is random since it depends on the demand. Evaluating the critical points is important because the allocations prior to the critical point account for most of the revenue. Moreover, the revenue prior to the critical point is deterministic, since it is simply the allocation times the fare. The example below suggests that the critical points for problems (P) and (PA) can be close. Notice that, since demand for different preference orders is assumed to be independent, it suffices to study a single preference order.

Example 2. *We consider a preference order with ten choices. To study how the critical point q^* behaves as demand increases or increases, consider three settings in which the exogenous demand is Poisson distributed with means respectively equal to 120 (low), 150 (medium) and 180 (high). The remaining parameters are listed in Table 2 below.*

Choices	Seat Allocation x	Fare levels \$	Transition Probability
1	35	400	0.9
2	25	450	0.95
3	25	700	0.8
4	23	800	0.8
5	11	1000	0.5
6	8	800	0.8
7	8	400	0.7
8	3	250	0.7
9	5	300	0.5
10	7	400	0

Table 2: Seat allocation and transition probabilities for Example 2

Table 3 shows some statistics on the critical point, obtained by simulating both systems 10,000 times.

We can see that the critical points are indeed close in all cases. Moreover, although variances for problem (P) are higher than their counterparts in problem (PA) — which is not surprising since (P) has an extra level of uncertainty — the expected values of total revenue are very close. This suggests that (PA) can actually be a good approximation for (P).

Model	Average C.P.	Variance on C.P.	Revenue	Variance on Revenue
Problem (P)-low	3.36	0.29	58,139	4.183×10^7
Problem (PA)-low	3.31	0.25	58,047	3.537×10^7
Problem (P)-medium	4.67	0.41	71,860	2.291×10^7
Problem (PA)-medium	4.65	0.32	72,072	1.773×10^7
Problem (P)-high	5.70	0.41	78,332	6.229×10^6
Problem (PA)-high	5.66	0.26	78,536	3.474×10^6

Table 3: Simulation results for Example 2 for the three demand levels (low, medium, high).

4 Solving Problem (PA)

We discuss now how to solve problem (PA), which is a stochastic programming model for which the random inputs are the random variables ξ_{is1}, $i = 1, \ldots, L$, $s = 1, \ldots, S_i$. We shall first consider an approximation where ξ_{is1} is replaced by its *expected value*. The proposition below shows the resulting problem. Here and thereafter the notation fx indicates $\sum_{i=1}^{L} \sum_{s=1}^{S_i} \sum_{r=1}^{R_{is}} f_{isr} x_{isr}$.

Proposition 1. *The model resulting from replacing ξ_{is1} with its expected value in* (PA) *is* equivalent *to the following deterministic linear program:*

$$\begin{aligned} &\max_x \ fx \\ \textit{Subject to: } & Ax \leq c \qquad \text{(PA-Lin)} \\ & x_{is1} + \frac{x_{is2}}{p_{is1}} + \ldots + \frac{x_{isR_{is}}}{p_{is1} \cdots p_{is(R_{is}-1)}} \leq \mathbb{E}[\xi_{is1}], \quad i = 1, \ldots, L, \ s = 1, \ldots, S_i \qquad (4.1) \\ & x \geq 0. \end{aligned}$$

Proof. Clearly, the resulting term $f_{is1} \min(\mathbb{E}[\xi_{is1}], x_{is1})$ in the objective function can be replaced with $f_{is1} x_{is1}$, together with the constraint $x_{is1} \leq \mathbb{E}[\xi_{is1}]$. Then, the term ξ_{is2} in (3.2) can be written as $\xi_{is2} = [\mathbb{E}[\xi_{is1}] - x_{is1}]_+ \, p_{is1} = (\mathbb{E}[\xi_{is1}] - x_{is1}) \, p_{is1}$. It follows that the term $f_{is2} \min(\xi_{is2}, x_{is2})$ in the objective function can be replaced with $f_{is2} x_{is2}$, together with the constraint $x_{is2} \leq (\mathbb{E}[\xi_{is1}] - x_{is1}) \, p_{is1}$, which in turn is equivalent to

$$x_{is1} + \frac{x_{is2}}{p_{is1}} \leq \mathbb{E}[\xi_{is1}]. \qquad (4.2)$$

Since the constraint $x_{is1} \leq \mathbb{E}[\xi_{is1}]$ is a relaxation of (4.2), it can be removed. By repeating this argument for the remaining values of r we obtain problem (PA-Lin). □

Alternatively, we consider a two-stage stochastic programming approximating formulation for (PA). The proposition below shows the resulting problem.

Proposition 2. *The following two-stage stochastic linear program with integer recourse provides the* same optimal solutions *as model* (PA)*:*

$$\begin{aligned} \max_x \; & fx - \mathbb{E}_{\xi_1}[Q(x,\xi_1)] \\ \textit{Subject to: } & Ax \le c \\ & x \ge 0, \end{aligned} \tag{PA-SP}$$

where

$$Q(x,\xi_1) = \min_{y,w,h\ge 0} \; fy \tag{4.3}$$

$$\begin{aligned} \textit{Subject to:} \quad x_{is1} - y_{is1} &\le \xi_{is1}, \\ x_{is1} + \frac{w_{is1} - h_{is1}}{p_{is1}} &= \xi_{is1}, \\ x_{isr} - y_{isr} - w_{is(r-1)} &\le 0, \quad r = 2,\dots,R_{is} \\ x_{isr} + \frac{w_{isr} - h_{isr}}{p_{isr}} - w_{is(r-1)} &= 0, \quad r = 2,\dots,R_{is}-1 \\ w_{isr} &\le \xi_{is1} b_{isr}, \quad r = 1,\dots,R_{is}-1 \\ h_{isr} &\le C(1-b_{isr}), \quad r = 1,\dots,R_{is}-1 \\ b_{isr} &\in \{0,1\}, \quad r = 1,\dots,R_{is}-1. \end{aligned}$$

In the above, all constraints are repeated for $i = 1,\dots,L$, $s = 1,\dots,S_i$, *and* C *is defined as* $C := \max_\ell c_\ell$.

Proof. Let z_{isr} denote the number of sold seats for the rth choice inside sth preference order for the ith OD combination, $i = 1,\dots,L$, so the terms $f_{isr}\min(\xi_{isr}, x_{isr})$ in the objective function of (PA) are replaced with $f_{isr}z_{isr}$, together with the constraints $z_{isr} \le x_{isr}$ and $z_{isr} \le \xi_{isr}$.

Fix now i and s. Note that the term ξ_{is2} in (3.2) can be written as $\xi_{is2} = \max\{(\xi_{is1} - x_{is1})\,p_{is1},\ 0\}$. We can write the latter term as w_{is1}, together with the constraints

$$w_{is1} - h_{is1} = (\xi_{is1} - x_{is1})\,p_{is1} \tag{4.4}$$

$$w_{is1} \le \xi_{is1} b_{is1} \tag{4.5}$$

$$h_{is1} \le C(1 - b_{is1}) \tag{4.6}$$

$$w_{is1} \ge 0,\ h_{is1} \ge 0,\ b_{is1} \in \{0,1\}. \tag{4.7}$$

Constraints (4.4)-(4.7), which involve binary variables, enforce that w_{is1} and h_{is1} be respectively the positive and negative parts of the term $(\xi_{is1} - x_{is1})\,p_{is1}$.

It follows from the above developments that the inequality $z_{is2} \leq \xi_{is2}$ generates $z_{is2} \leq w_{is1}$, in addition to (4.4)-(4.7). Similarly, the term ξ_{is3} in (3.2) can be written as $\xi_{is3} = \max\{(\xi_{is2} - x_{is2})\, p_{is2},\ 0\}$. Call the latter term w_{is2}. Since we must have $\xi_{is2} = w_{is1}$, we can write the constraints

$$
\begin{aligned}
w_{is2} - h_{is2} &= (w_{is1} - x_{is2})\, p_{is2} && (4.8)\\
w_{is2} &\leq \xi_{is1} b_{is2} && (4.9)\\
h_{is2} &\leq C(1 - b_{is2}) && (4.10)\\
w_{is2} &\geq 0,\ h_{is2} \geq 0,\ b_{is2} \in \{0,1\}. && (4.11)
\end{aligned}
$$

Hence, the inequality $z_{is3} \leq \xi_{is3}$ generates $z_{is3} \leq w_{is2}$, in addition to (4.8)-(4.11). By repeating this argument for successive values of r, we obtain the problem

$$
\begin{aligned}
&\max_x\ \mathbb{E}[fz]\\
\text{Subject to: } & Ax \leq c\\
& z_{is1} \leq \xi_{is1},\\
& z_{isr} \leq w_{is(r-1)}, \quad r = 2, \ldots, R_{is}\\
& w_{isr} \leq \xi_{is1} b_{isr}, \quad r = 1, \ldots, R_{is} - 1\\
& h_{isr} \leq C(1 - b_{isr}), \quad r = 1, \ldots, R_{is} - 1\\
& x_{is1} + \frac{w_{is1} - h_{is1}}{p_{is1}} = \xi_{is1},\\
& x_{isr} + \frac{w_{isr} - h_{isr}}{p_{isr}} = w_{is(r-1)}, \quad r = 2, \ldots, R_{is} - 1\\
& z \leq x,\\
& x \geq 0,\ z \geq 0,\ w \geq 0,\ b_{isr} \in \{0,1\}
\end{aligned}
$$

(where all constraints are repeated for $i = 1, \ldots, L,\ s = 1, \ldots, S_i$), which in turn is equivalent to (PA-SP) — this can be seen by defining $y_{isr} := x_{isr} - z_{isr}$. It is clear from the above derivation that the feasibility set of this problem is the same as that of (PA). □

It is easy to see from the above developments that, when the distribution of ξ_{is1} is atomic with point mass at $\xi_{is1} = \hat{\xi}_{is1}$, the optimal solution of (4.3) has $y = 0$, $w_{is1} = p_{is1}(\xi_{is1} - x_{is1})$, and $w_{isr} = p_{isr}(w_{is(r-1)} - x_{isr})$, provided that x is chosen in such a way that

$$
\begin{aligned}
x_{is1} &\leq \xi_{is1}\\
x_{is2} &\leq w_{is1} = p_{is1}(\xi_{is1} - x_{is1})\\
x_{is3} &\leq w_{is2} = p_{is2}[p_{is1}(\xi_{is1} - x_{is1}) - x_{is2}]\\
&\vdots
\end{aligned}
$$

(there is no benefit in choosing x otherwise, i.e., there might be other solutions that do not satisfy the above property but they cannot be better than that one). This in turn implies that $(w_{isr} - h_{isr})/p_{isr} \geq 0$ for all i, s, r, i.e., $h_{isr} = 0$ and $b_{isr} = 1$. In particular, when $\hat{\xi}_{is1} = \mathbb{E}[\xi_{is1}]$ we have that the formulation in (PA-SP) is the same as (PA-Lin). Since Q is convex in ξ_1, it follows from Jensen's inequality that $Q(x, \mathbb{E}[\xi_1]) \leq \mathbb{E}[Q(x, \xi_1)]$ for all x and hence the optimal value of (PA-Lin) provides an *upper bound* on the optimal value of (PA-SP). On the other hand, any optimal solution x^{Lin} to (PA-Lin) is feasible for (PA-SP); therefore, the objective value of (PA-SP) at x^{Lin} is a *lower bound* on the optimal value of (PA-SP). Note that such conclusions are analogous to those derived in case of the traditional non-choice-based mathematical programming formulations for network revenue management; see, for instance, Chen and Homem-de-Mello [2008].

While model (PA-SP) provides in principle a more accurate solution — since it is equivalent to model (PA) — the advantage of formulation (PA-Lin) is that it can be solved more easily. (PA-Lin) can be solved with any linear programming solver; (PA-SP), on the other hand, can be solved as a binary linear program provided the total number of scenarios in the problem is small. In case the number of scenarios is large one can resort to sampling techniques for stochastic programs as we do in the experiments below (see for instance Shapiro 2003 for a review of such methods), but even then the size of the problem may become an obstacle. One could combine sampling techniques with a specialized algorithm for stochastic programs with integer recourse, but such an investigation is out of the scope of this paper.

Alternately, we propose a simpler two-stage stochastic program, called (PA-SPs), that still generalizes (PA-Lin) but is much easier to solve. The first stage is the same as in (PA-SP), but the second stage is defined as

$$Q(x, \xi_1) = \min_{y \geq 0} \widetilde{f} y \tag{PA-SPs}$$

Subject to:

$$x_{is1} + \frac{x_{is2}}{p_{is1}} + \ldots + \frac{x_{isR_{is}}}{p_{is1} \cdots p_{is(R_{is}-1)}} - y_{is} \leq \xi_{is1}, \quad i = 1, \ldots, L, \; s = 1, \ldots, S_i. \tag{4.12}$$

In the above, each element of the vector $\widetilde{f}$ is defined as

$$\widetilde{f}_{is} \;:=\; \max \left\{ f_{is1}, \; f_{is2} p_{is1}, \; f_{is3} p_{is1} p_{is2}, \ldots, f_{isR_{is}} \prod_{r=1}^{R_{is}-1} p_{isr} \right\}. \tag{4.13}$$

Again, it is easy to see that when the distribution of ξ_{is1} is atomic with point mass at $\xi_{is1} = \hat{\xi}_{is1}$, the optimal solution of the second stage in (PA-SPs) has $y = 0$. This happens because in order for (4.12) to be satisfied, each additional unit of y_{is} must be compensated by $\prod_{k=1}^{r-1} p_{isk}$ units of x_{isr}

for some r. The resulting gain in the objective function of the first stage is therefore $f_{isr} \prod_{k=1}^{r-1} p_{isk}$ which, by definition (4.13) of the coefficients of y, is offset by the cost of the second stage. Thus, (PA-SPs) generalizes (PA-Lin), so the same relationship (in terms of bounds) between (PA-Lin) and (PA-SP) holds between (PA-Lin) and (PA-SPs).

5 Relationship with the Independent Demand Model

An important question that arises about the preference order model concerns the value of such a model compared to traditional independent demand models studied in the literature. In the independent demand case, each booking class has a random demand ξ. The customer in each class requests a reservation once and, if that class is not available, the customer leaves the booking process without looking at other available booking classes. In contrast, the preference order model assumes the customer is willing to consider alternatives. Figures 3 and 4 illustrate both processes for the ith OD combination with $n := n_i k_i$ booking classes.

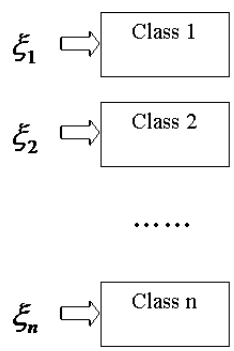

Figure 3: Independent demand model

It is clear that the traditional independent demand is a particular case of the preference order model, obtained by setting the number of preference orders S_i to $n_i k_i$ and the number of choices within each preference order to one. It is interesting to compare the corresponding mathematical programs. The standard linear program for the independent demand model (see, e.g., Talluri and

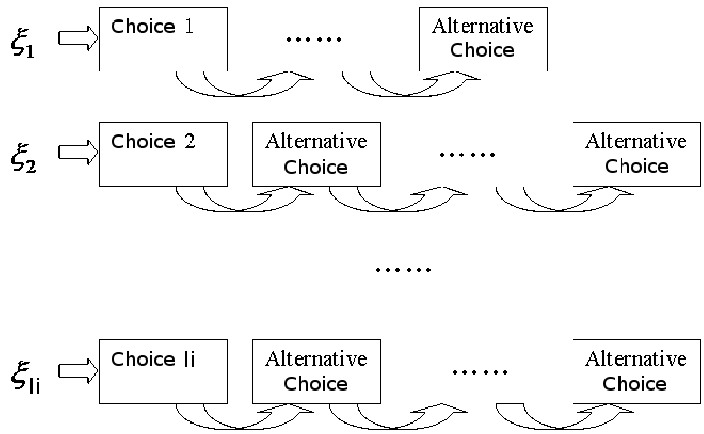

Figure 4: Independent preference order model

van Ryzin 2004a) is

$$\begin{aligned} &\max\ fx \\ &\text{subject to} \\ &Ax \le c, \\ &x \le \mathbb{E}[\xi] \\ &x \ge 0. \end{aligned} \tag{DLP}$$

In (DLP), the total number of booking classes is $\sum_{i=1}^{L} n_i k_i$. Clearly, (DLP) is identical to model (PA-Lin) with $S_i = n_i k_i$ and $R_{is} = 1$, $s = 1, \ldots, S_i$.

Consider now the traditional stochastic program for the independent demand model, often referred to as the probabilistic nonlinear program (PNLP):

$$\begin{aligned} &\max\ f\,\mathbb{E}[\min(\xi, x)] \\ &\text{subject to} \\ &Ax \le c, \\ &x \ge 0. \end{aligned} \tag{PNLP}$$

Let us look at models (PA-SP) and (PA-SPs) with $S_i = n_i k_i$ and $R_{is} = 1$, $s = 1, \ldots, S_i$. In that case the second stage becomes simply $\min\{fy : y \ge 0,\ y \ge x - \xi\}$, which is more briefly written as $f[x - \xi]_+$. Since $x - [x - \xi]_+ = \min(x, \xi)$, it follows that (PNLP) coincides with both (PA-SP) and (PA-SPs) in that case.

The above discussion demonstrates that models (PA-Lin), (PA-SP) and (PA-SPs) generalize the traditional (DLP) and (PNLP) models to the case of consumer choice. This suggests that many

of the results obtained for those well-studied models (such as asymptotic optimality or the use of dual variables as bid prices, for example) may be applicable to these models as well. Such a study is, however, out of the scope of this paper.

6 Implementing the Policy

The seat allocation provided by models (PA-Lin), (PA-SP) and (PA-SPs), if implemented directly, may not lead to a good booking policy. Broadly speaking, the reason is that, by construction, several triples (OD combination, preference order, choice within the preference order) correspond to the same booking class but are located in different preference orders. Below we discuss this issue in detail and propose a heuristic procedure as a remedy.

6.1 Drawbacks of applying the model recommendation directly

Models (PA-Lin), (PA-SP) and (PA-SPs) yield seat allocations for each choice within each preference order. The natural way to translate such allocations into a booking policy is to accept requests for a pair (OD, preference order) until the allocation corresponding to the first choice class is used up, after which the next class within that preference order becomes available. When the last class within that preference order is used up, the customer is rejected.

To see why such a policy is not practical, consider the following example. Assume an airline company has two itineraries, A and B. Itinerary A is a connected trip while itinerary B is a direct trip. There are two classes for both itineraries, such as coach class and business class. The fares for coach and business class are \$80, \$300 respectively in itinerary A and \$100, \$500 for itinerary B. Suppose we have three preference orders, corresponding to budget-sensitive, time-sensitive, or budget-time hybrid type customers. The fares for budget-sensitive customers arranged by preference are \$80, \$100, \$300, \$500. Likewise, the fares are \$100, \$500, \$80, \$300 for time-sensitive customers and \$100, \$80, \$500, \$300 for hybrid-type customers.

One drawback happens when booking requests for high-fare classes are rejected while there are still some seats for the same booking class left available in another preference order. For instance, suppose there is a booking request from a customer who only wants itinerary B due to time constraints. Suppose also that no more \$100 seats are available. The customer will then attempt to purchase a \$500 seat. If that class happens to be used up, the customer will leave and purchase his trip elsewhere. Meanwhile, there are quite a few \$500 seats assigned to the other preference orders. Thus, if we follow the seat allocation policy without flexibility, that valued customer will be rejected.

The above discussion suggests that it is a good idea to "merge" allocations corresponding to the same class. Such merging, however, must be done in a careful way. To illustrate that problem, suppose that a time-sensitive customer is booking the \$100 class and that a hybrid customer is also booking the same class. But their next classes' rates are noticeably different from each other, \$500 and \$80. If the hybrid customers could use \$100 seats from the allocations for both time-sensitive and hybrid customers, the shift for hybrid customers to buy \$80 ticket would be delayed and the shift for time sensitive customers to buy \$500 ticket would tend to happen earlier. By the same token, it is advantageous not to let time-sensitive customers use \$100 seats allocated to hybrid customers, to avoid delaying the shift for the former to purchase the \$500 ticket.

Intuitively, those delays or accelerations would benefit the booking process by lifting the total revenue. We describe next a heuristic procedure called *backup heuristic* to improve the seat allocation policy.

6.2 Backup heuristic on the seat allocation

We first introduce additional notation. For each rth choice open for booking, we have three values: x_{isr} is the seat allocation, f_{isr} is the fare level, and

$$\text{UBUR}_{isr} := \max\{p_{isr} f_{is(r+1)},\ p_{isr} p_{is(r+1)} f_{is(r+2)},\ \ldots,\ p_{isr} \cdots p_{is(R_{is}-1)} f_{isR_{is}}\} \tag{6.1}$$

is the *unit buying up revenue* which indicates the potential unit revenue if customers are asked to *buy-up*. The term "*buy-up*" describes the activity to forward booking requests to the next available but less favorite choice in the given preference order. When $\text{UBUR}_{isr} > f_{isr}$, asking customers to buy-up tends to generate more revenue. In this case, we should try to let buy-up activity happen earlier. Otherwise, when $\text{UBUR}_{isr} \leq f_{isr}$, buy-up activity tends to hurt the total revenue, so we should try to postpone buy-up call. Figure 5 illustrates the UBUR concept for the example described in Section 6.1. There are three values in each block. The first bucket is the model recommended allocation, x. The second bucket denotes the fare level while the third bucket is the calculated UBUR.

We define now the *backup* operation.

Definition 2. When $\text{UBUR}_{isr} \leq f_{isr}$ for rth choice in a given preference order, we use the same booking class in other allocations to accept upcoming booking requests until all seats in this class are reserved. This is the operation called *backup*.

Figure 6 shows the backup decisions for the UBUR values in Figure 5.

The backup operation, as described, is static, in the sense that the UBUR values do not change with booking requests. It is intuitive, however, that a better procedure would be obtained if UBUR

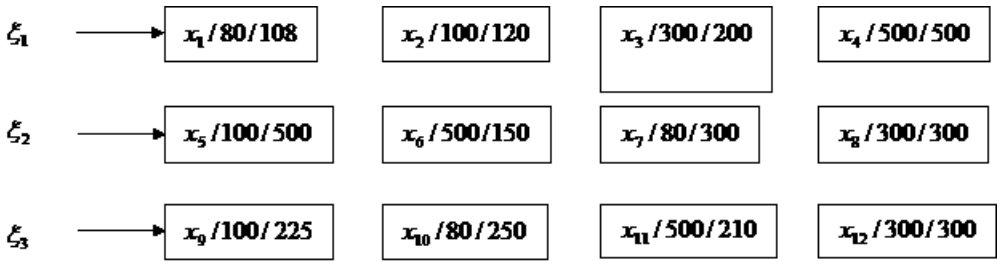

Figure 5: Example to illustrate UBUR values

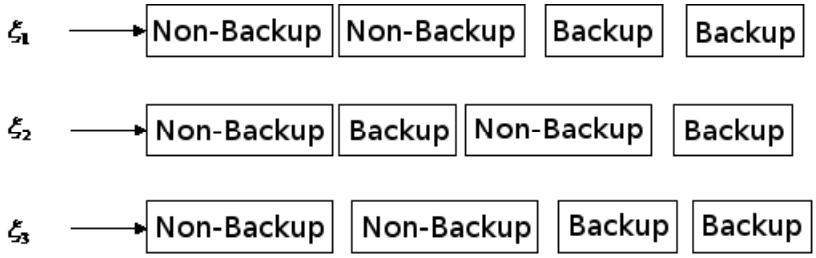

Figure 6: Example for the backup policy

values changed upon the current seat availability. In that case, any materialized reservation would trigger a change in UBUR values. In what follows, we consider two heuristics: the *static backup* method which is based on initial UBUR values, and the *dynamic backup* method which updates UBUR values upon changes on current seat availabilities. The dynamic backup heuristic algorithm is described next.

- Step 1: Pull out seat allocation from the models (PA-Lin), (PA-SP) or (PA-SPs).
- Step 2: Initialize UBURs using (6.1).
- Step 3: Calculate the backup policy by UBUR and analyze the next booking request:

  3.1 If requested booking is Non-Backup (fare level is lower than UBUR), then check the seat allocation availability in step 3.2. Otherwise, go to step 3.3.

  3.2 If no seat available for this request, upon customer's buying-up decision go to step 3.1 with the booking request set to the next choice in that preference order (if there is no next choice, the request is rejected). Otherwise, honor the request, update the seat availability, and go to step 4.

  3.3 Check seat availabilities from all allocations for the same booking class in the other preference orders. If there is no seat for the requested booking class, upon customer's

buying-up decision go to step 3.1 with the booking request set to the next choice in that preference order (if there is no next choice, the request is rejected). Otherwise, honor the request, update the seat availability, and go to step 4.

- Step 4: Check if any class is no longer available (i.e., if all allocations corresponding to that booking class have been reduced to zero); if so, eliminate the corresponding "blocks" from the model. Go to step 5.
- Step 5: If there are no more requests or no more seats available, stop the process; Otherwise, go back to step 3.

7 Numerical Experiments

We present now some numerical experiments to illustrate the performance of the methods discussed in the previous sections and compare it with some of the existing models. We start with a very small single-leg example, for which most models yield an analytical solution, and then move on to larger networks.

7.1 A single-leg example

Consider a single-leg flight with two classes, say, A and B, with $f_A > f_B$. The arrival rate of customers who plan to purchase the high-fare ticket (e.g., due to time constraints) is $\lambda_A = 0.3$. Of those customers, 80% are willing to purchase the low-fare ticket if the high-fare ticket is not available. Likewise, the arrival rate of customers who plan to purchase the low-fare ticket is $\lambda_B = 0.7$, and 50% of those customers are willing to purchase the high-fare ticket if the low-fare ticket is not available. The arrival processes are assumed to be Poisson, the capacity of the plane is $c = 100$, and the time horizon is $T = 100$.

How would one approach this problem with an independent demand model? One possibility, of course, is to simply ignore the second-choice of each type of customer, but such an approach is obviously sub-optimal. A more realistic approach is to use a heuristic procedure to accommodate the choice probabilities. For example, one could set $\lambda'_A := \lambda_A + 0.5\lambda_B = 0.65$ as the "inflated" rate of high-fare customers that takes into account the ones that were denied a low-fare ticket. Likewise, one sets $\lambda'_B := \lambda_B + 0.8\lambda_A = 0.94$. Using the inflated rates, one can then apply a standard model

like the LP model or even the Littlewood rule. The LP model solves

$$\max \; f_A x_A + f_B x_B$$

subject to

$$x_A + x_B \leq 100,$$
$$0 \leq x_A \leq 65$$
$$0 \leq x_B \leq 94$$

Clearly, the optimal solution is $x_A = 65$, $x_B = 35$. If no adjustment is made, that is, if the original rates λ_A and λ_B are used, the solution to the LP model is $x_A = 30$, $x_B = 70$. The Littlewood rule uses the f_B/f_A quantile of the demand distribution F_A for the high-fare class, i.e., it sets $x_A := F_A^{-1}(1 - f_B/f_A)$ and $x_B := c - x_A$. The distribution F_A may correspond to the original rates or the inflated rates.

Let us consider the preference order models (PA-Lin) and (PA-SP) for this problem. In that case we have two preference orders (A→B and B→A) with rates $\lambda_1 = \lambda_A = 0.3$ and $\lambda_2 = \lambda_B = 0.7$, and four variables x_{11}, x_{12}, x_{21} and x_{22}. The (PA-Lin) model is

$$\max \; f_A x_{11} + f_B x_{12} + f_B x_{21} + f_A x_{22}$$

subject to

$$x_{11} + x_{12} + x_{21} + x_{22} \leq 100,$$
$$x_{11} + \frac{1}{0.8} x_{12} \leq 30$$
$$x_{21} + \frac{1}{0.5} x_{22} \leq 70$$
$$x \geq 0.$$

It easy to see that, if $f_A \geq 2f_B$, the optimal solution is $x_{11} = 30$, $x_{12} = x_{21} = 0$, and $x_{22} = 35$. Since x_{11} and x_{22} correspond to the same class (A), the model sets $x_A = 65$, $x_B = 0$ (for simplicity, in this example we do not consider the more elaborate backup heuristic described in Section 6.2). It is interesting to compare that allocation with the solution of the adjusted independent model: both set $x_A = 65$, but (PA-Lin) shuts down the low-fare class to "save" low-fare customers who are willing to pay more. If $f_A < 2f_B$ the optimal solution is $x_{11} = 30$, $x_{12} = x_{22} = 0$, and $x_{21} = 70$.

Model (PA-SP) can be solved numerically. As mentioned before, sampling techniques can be used to reduce the size of the model. For this experiment we took a small sample size of 15. Since the solutions vary according to the sample, we conducted multiple replications. The actual revenues corresponding to the obtained solutions in all instances were close to each other (more precisely, within 0.5%), so in the results below we report the average allocation and average revenue. In

some cases all replications yielded the same allocation — for example, for $f_A = 100$ and $f_B = 50$, all solutions satisfied $x_{11} + x_{22} = 100$ and $x_{12} = x_{21} = 0$. That is, (PA-SP) allocates all 100 seats to the high-fare customers. Clearly, such a solution can be no worse than the solution given by (PA-Lin).

Model (PA-SPs) can also be solved numerically. In fact, for this small example it can be solved almost exactly by truncating the Poisson distribution of the arrival process in such a way that the probability of the truncated region of each variable is 0.99. That resulted in 1,040 scenarios, but since (PA-SPs) can be formulated as a linear program it can be solved very fast even for that number of scenarios.

For comparison, let us look at the CDLP model described in Liu and van Ryzin [2008]. For this problem the resulting LP is very small, since there are only three offer sets (A only, B only, A+B). Calling the resulting variables t_A, t_B and t_{AB}, respectively, we have

$$\max\ 0.65 f_A t_A + 0.94 f_B t_B + (0.3 f_A + 0.7 f_B) t_{AB}$$

subject to

$$\begin{aligned} 0.65 t_A + 0.94 t_B + t_{AB} &\leq 100 \\ t_A + t_B + t_{AB} &\leq 100 \\ t_A,\ t_B,\ t_{AB} &\geq 0. \end{aligned}$$

Clearly, the solution to the above problem is to allocate everything to the offer set that yields the maximizer of $\{0.65 f_A,\ 0.94 f_B,\ (0.3 f_A + 0.7 f_B)\}$. In the examples below the maximizer was always the third element, which means that the policy recommended by the CDLP model was not to impose any control, i.e., leave both classes A and B open throughout the booking horizon and reject a request only if the overall capacity has been filled. This corresponds to setting $t_{AB} = 100$, $t_A = t_B = 0$.[1]

We tested the above models with different price ratios. More specifically, f_A was fixed at \$100, and f_B took values in $\{\$50, \$60, \$75\}$. Table 4 summarizes the simulated revenues for all models, together with 95% confidence intervals. All simulations were conducted with the same stream of random numbers to facilitate comparison.

The results suggest that models (PA-SP) and (PA-SPs) are quite robust to variation in the fares — although they do not always yield the highest revenue, both methods perform consistently well. The other models perform well in some cases but poorly on others. Model CDLP also performs

[1] It must be noted that in the case $f_A = \$100$, $f_B = \$50$ any solution satisfying $t_A + t_{AB} = 100$ is optimal. Different solutions, however, yield different revenues. For example, for the solution $t_A = 100$ (open class A only) the revenue is identical to that given by models (PA-SP) and (PA-SPs).

	$f_A = \$100,\ f_B = \$75$		$f_A = \$100,\ f_B = \$60$		$f_A = \$100,\ f_B = \$50$	
Model	Allocation (A,B)	Revenue 95% C.I.	Allocation (A,B)	Revenue 95% C.I.	Allocation (A,B)	Revenue 95% C.I.
LP	(30,70)	7917±39	(30,70)	6903±35	(30,70)	6226±33
LP-Adj.	(65,35)	7492±66	(65,35)	6967±66	(65,35)	6617±66
Littlewood	(59,41)	7612±63	(63,37)	6981±65	(65,35)	6617±66
(PA-Lin)	(30,70)	7917±39	(30,70)	6903±35	(65,0)	6224±37
(PA-SP)	(37,63)	7956±46	(47,53)	7043±53	(100,0)	6590±71
(PA-SPs)	(31,69)	7936±40	(44,56)	7059±53	(100,0)	6590±71
CDLP	both open	7963±41	both open	6956±39	both open	6284±38

Table 4: Results for single-leg example

well but, as pointed out earlier, in the case $f_A = \$100,\ f_B = \$50$ it may lead to a poor policy depending on which solution is chosen.

7.2 Numerical results for networks

We present now some numerical experiments for networks of flights. We use two examples, for which we apply both (PA-Lin) and (PA-SPs) models to determine the seat allocations. To test the efficacy of the backup heuristic, we compare the results obtained using the backup heuristic with its counterpart which applies allocation directly. As discussed earlier, since (PA-SPs) is a two stage stochastic programming problem with large number of scenarios, it can be solved by using the sample average approximation approach. It must be noticed that the computational times can change depending on the number of samples used to solve the stochastic program, which in turn is related to the desired accuracy of the estimates. Still, to give a sense of the complexity of the problem we can easily calculate the number of variables. For a sample of size N, the total number of variables for the linear programming formulation of (PA-SPs) is (no. of OD pairs) $\times$ (no. of preference orders for each pair) $\times$ (N+no. of choices within each preference order). The total number of variables for (PA-Lin) is just (no. of OD pairs) $\times$ (no. of preference orders for each pair) $\times$ (no. of choices within each preference order).

For both examples, presented in Sections 7.2.1 and 7.2.2, we sampled $10,000$ scenarios to represent the problem. Based on the discussion above, we can see that model (PA-SPs) has $40,024$ variables in the first example, and $100,060$ variables in the second one. The computational time to solve the second example was seven times higher than the time needed to solve the

first one.The optimization problems were solved using the simplex solver from the Coin-OR library, on a Debian/Lenny/AMD64-Dual-Xeon workstation platform. The computational time to solve (PA-Lin) was negligible in either case.

We also compare our preference order model with/without backup heuristics with the CDLP model (see Section 1 for a discussion of that model). In order to make a fair comparison with those counterparts, we use the same preference order models for the parallel flight example described in Zhang and Cooper [2005], Miranda Bront et al. [2007], and Liu and van Ryzin [2008]. The second example's network structure is borrowed from Miranda Bront et al. [2007] with the same purpose. In the parallel flight example, we have $2^6 - 1 = 63$ decision variables for the CDLP model. In the hub network, we have 22 booking classes and therefore, we have $2^{22} - 1 = 4194303$ decision variables. It is clear that the time to solve the second example is far longer than the time required for the parallel flight model.

We introduce some new notation.

1. $\mathbf{NB}^{\mathrm{LP}}$ means "no backup" by the optimal solution of (PA-Lin) model. We take the optimal solution from (PA-Lin) model and implement it directly without any post-optimization heuristic.

2. $\mathbf{NB}^{\mathrm{SP}}$ means "no backup" by implementing the optimal solution of (PA-SPs) model directly.

3. $\mathbf{DB}^{\mathrm{LP}}$ is "dynamic backup" by the optimal solution of (PA-Lin) model.

4. $\mathbf{DB}^{\mathrm{SP}}$ is "dynamic backup" by the optimal solution of (PA-SPs) model.

5. **PI** is the perfect information model. We use this result to measure the performance of our preference order model and CDLP model.

7.2.1 Experiments for parallel flights

The first example is based on a network with three parallel flights, and two fare classes, business and coach, on each flight. This example is illustrated in Figure 7. We construct four preference orders listed in Table 5 and the available itineraries listed in Table 6. The initial capacity is $c = (30, 50, 40)$. The booking horizon consists of $T = 300$ time periods with an average of 150 arrivals. There are four customer preference orders, which are shown in Table 6.

In order to test our model under different loads, we simulated the system under the policies suggested by our methods on three demand scenarios: heavy demand, fair demand, and soft demand. We use λ be the parameter of demand loads. $\lambda = 0.75$ indicates the demand is at only 75% of regular demand level. $\lambda = 1.25$ indicates the network is overloaded by additional 25%. For each case, we have the capacity fixed.

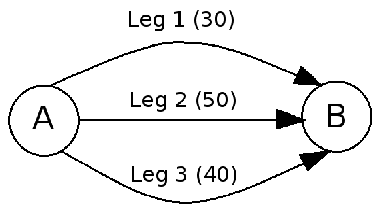

Figure 7: Example 1 of comparisons on multiple methods

Itinerary	Leg	Class	Fare
1, morning flight	1	coach	400
2, morning flight	1	business	800
3, afternoon flight	2	coach	500
4, afternoon flight	2	business	1,000
5, evening flight	3	coach	300
6, evening flight	3	business	600

Table 5: Itineraries setting for the parallel flights

Preference order	Ordered itineraries	Transition probability	Demand level
1	\$1000, \$800,\$600	0.5,0.2	30
2	\$300, \$400,\$500	0.5,0.2	45
3	\$400, \$800, \$500, \$1000, \$300,\$600	0.8,0.75,0.66,0.75,0.33	60
4	\$800, \$400, \$1000, \$500, \$600,\$300	0.8,0.75,0.66,0.75,0.33	15

Table 6: Preference orders' setting for the parallel flights

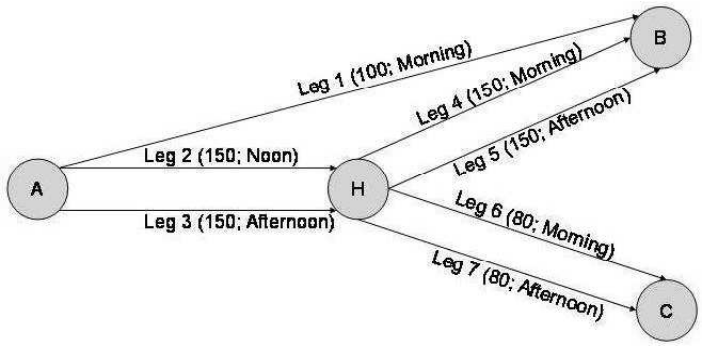

Figure 8: The hub network example of comparisons on multiple methods

It is important to clarify that the CDLP results we present here were obtained for our buy-up/down structure rather than for a multinomial logit choice model. Therefore, our numerical results on the CDLP model are close but do not match the results in Miranda Bront et al. [2007] and Liu and van Ryzin [2008].

The numerical results are presented in Table 7. We use $\pm$ to denote a 95% confidence interval from the simulation results.

Demand	$\mathbf{NB}^{\mathrm{LP}}$	$\mathbf{NB}^{\mathrm{SP}}$	$\mathbf{DB}^{\mathrm{LP}}$	$\mathbf{DB}^{\mathrm{SP}}$	**CDLP**	**PI**
$\lambda = 0.75$	65622 ± 122	65651 ± 123	65988 ± 131	66123 ± 131	65994 ± 139	66180 ± 119
$\lambda = 1$	78706 ± 83	79252 ± 96	80711 ± 86	81287 ± 91	79917 ± 108	82334 ± 94
$\lambda = 1.25$	83076 ± 34	84684 ± 44	83909 ± 13	84866 ± 15	83241 ± 52	87652 ± 21

Table 7: Numerical results for the parallel flights example

7.2.2 Experiment for a hub network

This example is based on a network considered in Liu and van Ryzin [2008] and Miranda Bront et al. [2007], which is illustrated in Figure 8. The network has a total of 22 booking classes. There is local traffic, and traffic over the hub which is connected through 10 legs. We list the parameters in Table 8 followed by the setting of preference orders in Table 9.

The numerical results for the hub network are listed in Table 10. We must remark the computational burden for the CDLP model grows exponentially. As we mentioned earlier, the column-generation approach developed in Liu and van Ryzin [2008] and Miranda Bront et al. [2007] cannot be used in this setting since our choice model is not derived from a multinomial logit. Furthermore, the complexity of a simulation model for CDLP also becomes higher.

Itinerary	Leg	Class	Fare	Itinerary	Leg	Class	Fare
1, morning flight	1	business	1000	12, evening flight	1	coach	500
2, morning flight	2	business	400	13, morning flight	2	coach	200
3, afternoon flight	3	business	400	14, morning flight	3	coach	200
4, afternoon flight	4	business	300	15, afternoon flight	4	coach	150
5, evening flight	5	business	300	16, afternoon flight	5	coach	150
6, evening flight	6	business	500	17, evening flight	6	coach	250
7, morning flight	7	business	500	18, evening flight	7	coach	250
8, morning flight	2,4	business	600	19, morning flight	2,4	coach	300
9, afternoon flight	3,5	business	600	20, morning flight	3,5	coach	300
10, afternoon flight	2,6	business	700	21, afternoon flight	2,6	coach	350
11, evening flight	3,7	business	700	22, afternoon flight	3,7	coach	350

Table 8: Itineraries setting for the hub network

PO	Ordered itineraries	Transition probability	Demand level
1	\$1000, \$600,\$600,\$500,\$300,\$300	0.8,1,0.75,0.66,1	80
2	\$300, \$300,\$500,\$600,\$600,\$1000	1,0.8,0.25,1,0.5	200
3	\$400, \$400, \$200, \$200	1,0.5,1	50
4	\$200, \$200, \$400, \$400	1,0.2,1	200
5	\$300, \$300, \$150, \$150	1,0.5,1	100
6	\$150, \$150, \$300, \$300	1,0.2,0.8	150
7	\$500, \$500, \$250, \$250	0.8,0.625,1	20
8	\$250, \$250, \$500, \$500	1,0.2,0.8	50
9	\$700, \$700, \$350, \$350	0.8,0.625,1	20
10	\$350, \$350, \$700, \$700	1,0.2,1	40

Table 9: Preference orders' setting for the hub network

Demand	$\mathbf{NB}^{\mathrm{LP}}$	$\mathbf{NB}^{\mathrm{SP}}$	$\mathbf{DB}^{\mathrm{LP}}$	$\mathbf{DB}^{\mathrm{SP}}$	**CDLP**	**PI**
$\lambda = 0.75$	210730 ± 165	195630 ± 213	211730 ± 177	211760 ± 168	211340 ± 213	227013 ± 145
$\lambda = 1$	246780 ± 79	255670 ± 163	248910 ± 73	256950 ± 159	246970 ± 215	268277 ± 147
$\lambda = 1.25$	254090 ± 19	274050 ± 61	254160 ± 20	274320 ± 67	255810 ± 98	304223 ± 130

Table 10: Numerical results for the hub network example

7.3 Numerical experiments summary

The simulation results presented above lead to the following conclusions:

- The seat allocation from model (PA-SPs) tends to generate more expected revenue than the counterpart from model (PA-Lin), especially when dynamic backup is applied. This is natural since (PA-SPs) incorporates more information about the demand distribution other than just the mean.

- The backup heuristic is one of the critical features of the preference order model. Applying a backup heuristic can significantly improve the booking performance. The simple example in Section 6.1 illustrates the issue, and the numerical results corroborate that intuition for larger scale models.

- The preference order model can significantly improve the booking performance by granting other choices to customers whose first choice was not available. While intuitive, such a result quantifies the benefit of using a more accurate model that incorporates more information.

- The CDLP model is a good alternative to our preference order model. However, as the number of booking classes grows, the number of choice sets — and hence the number of variables — grows exponentially, whereas the size our preference order model tends to grow much more slowly (the actual complexity depends on the preference order structure). On the other hand, CDLP uses a more general choice model based on offer sets, whereas ours is developed for the case of buy-up/down probabilities. Conceivably, the computational burden to run CDLP could be reduced by developing column generation algorithms for our preference order structure.

- We notice that our preference order model (with dynamic backup) outperforms the CDLP model in the experiments. We conjecture that this is due to the fact that the CDLP model does not incorporate randomness, which in some cases makes the model unable to distinguish between different policies (cf. footnote comment in Section 7.1).

8 Conclusions

Research on discrete customer choice models for revenue management is an important topic of ongoing activity in the academic community. It has been well recognized that discrete customer choice is an important part of booking process.

In this paper we have proposed some mathematical programming models for network optimization where customer choice is formulated through the concept of preference orders. The model gives

a seat allocation that maximizes the expected revenue, much like the traditional seat allocation models do in the independent demand case — indeed, as we have discussed, our approximating models constitute extensions of the traditional mathematical programming models to the case of consumer choice. Similarly, our models yield only heuristic policies; on the other hand, we do not encounter the curse of dimensionality and other modeling issues concerning the state space in a network.

Besides proposing a novel model for network optimization under customer choice, we have also discussed a way to implement the model recommendations. We have proposed our backup heuristic and also compared the preference order model with alternative models. Simulation results indicate that the preference order model provides robust solutions to the problem.

In reality, an individual customer chooses either to stay or to leave randomly. Thus, a binomial process such as the one used in model (P) would be more accurate to model the buy-up activity. However, incorporating binomial random variables into the model leads to an exploding scenario tree. To reduce the computational complexity, we have used deterministic staying rates as an approximation. A topic for future research would be to determine the bounds on the error of that approximation compared to the true model.

Acknowledgement

We thank two anonymous referees for their helpful comments and suggestions. This work has been supported in part by the National Science Foundation under grant DMI-0115385.

References

S. Algers and M. Beser. Modelling choice of flight and booking class — a study using stated preference and revealed preference data. *International Journal of Services Technology and Management*, 2(1/2):28–45, 2001.

S. Andersson. Passenger choice analysis for seat capacity control: A pilot project in Scandinavian Airlines. *International Transactions in Operational Research*, 5(6):471–486, 1998.

P. Belobaba. Application of a probabilistic decision model to airline seat inventory control. *Operations Research*, 37(2):183–197, 1989.

P. Belobaba and L. Weatherford. Comparing decision rules that incorporate customer diversion in perishable asset revenue management situations. *Decision Sciences*, 27(2):343–363, 1996.

S. E. Bodily and L. R. Weatherford. Perishable-asset revenue management: Generic and multi-price yield management with diversion. *Omega*, 23(2):173–185, 1995.

S. L. Brumelle, J. I. McGill, T. H. Oum, K. Sawaki, and M. W. Tretheway. Allocation of airline seats between stochastically dependent demands. *Transportation Science*, 24:183–192, 1990.

L. Chen and T. Homem-de-Mello. Re-solving Stochastic Programming Models for Airline Revenue Management. *Annals of Operations Research*, 2008. In press.

W. L. Cooper, T. Homem-de-Mello, and A. Kleywegt. Models of the spiral down effect in revenue management. *Operations Research*, 54(5):968–987, 2006.

S. V. de Boer, R. Freling, and N. Piersma. Mathematical programming for network revenue management revisited. *European Journal of Operational Research*, 137:72–92, 2002.

G. Gallego and M. Hu. Dynamic pricing of perishable assets under competition. Technical report, Columbia University, 2006.

G. Gallego, G. Iyengar, R. Phillips, and A. Dubey. Managing flexible products on a network. CORC Technical Report Tr-2004-01, IEOR Department, Columbia University, 2004.

Q. Liu and G. van Ryzin. On the choice-based linear programming model for network revenue management. *Manufacturing and Service Operations Management*, 10(2):288–310, 2008.

J. J. Miranda Bront, I. Mendez-Diaz, and G. Vulcano. A column generation algorithm for choice-based network revenue management. Technical report, University of Buenos Aires, Buenos Aires, Argentina, 2007.

A. Shapiro. Monte Carlo sampling methods. In A. Ruszczyński and A. Shapiro, editors, *Handbook of Stochastic Optimization*. Elsevier Science Publishers B.V., Amsterdam, Netherlands, 2003.

K. Talluri and G. van Ryzin. Revenue management under a general discrete choice model of consumer behavior. *Management Science*, 50(1):15–33, 2004a.

K. Talluri and G. van Ryzin. *The Theory and Practice of Revenue Management*. Kluwer Academic Publishers, 2004b.

G. van Ryzin and G. Vulcano. Computing virtual nesting controls for network revenue management under customer choice behavior. *Manufacturing and Service Operations Management*, 10(3):448–467, 2008.

D. Zhang and D. Adelman. An approximate dynamic programming approach to network revenue management with customer choice. Working paper, Desautels Faculty of Management, McGill University, 2007.

D. Zhang and W. L. Cooper. Revenue management for parallel flights with customer-choice behavior. *Operations Research*, 53:415–431, 2005.