

# A Randomized Linear Program for the Network Revenue Management Problem with Customer Choice Behavior

(Research Paper)

Sumit Kunnumkal  
(Corresponding Author)

Indian School of Business, Gachibowli, Hyderabad, 500032, India  
sumit\_kunnumkal@isb.edu

Tel: 91-40-23187168 Fax: 91-40-23187200

Huseyin Topaloglu

School of Operations Research and Information Engineering,  
Cornell University, Ithaca, New York 14853, USA  
topaloglu@orie.cornell.edu

June 1, 2010

## Abstract

In this paper, we propose a randomized linear program to obtain bid prices for the network revenue management problem with customer choice behavior. Our development proceeds in two stages. We first formulate a mixed integer program that can be viewed as a deterministic approximation to the network revenue management problem. We then use the optimal solution to the mixed integer program to formulate a linear program that can be easily randomized. We establish that our solution method approximates the optimal total expected revenue arbitrarily closely in an asymptotic regime where the leg capacities and the number of time periods in the decision horizon increase linearly with the same rate. Numerical experiments indicate that our approach may be a viable alternative to the standard deterministic linear program that appears in the existing literature.

*Keywords.* network revenue management, customer choice behavior, linear programming, simulation, bid prices.

# A Randomized Linear Program for the Network Revenue Management Problem with Customer Choice Behavior

(Research Paper)

June 1, 2010

## **Abstract**

In this paper, we propose a randomized linear program to obtain bid prices for the network revenue management problem with customer choice behavior. Our development proceeds in two stages. We first formulate a mixed integer program that can be viewed as a deterministic approximation to the network revenue management problem. We then use the optimal solution to the mixed integer program to formulate a linear program that can be easily randomized. We establish that our solution method approximates the optimal total expected revenue arbitrarily closely in an asymptotic regime where the leg capacities and the number of time periods in the decision horizon increase linearly with the same rate. Numerical experiments indicate that our approach may be a viable alternative to the standard deterministic linear program that appears in the existing literature.

*Keywords.* network revenue management, customer choice behavior, linear programming, simulation, bid prices.

One of the fundamental issues in network revenue management is capacity control, which involves deciding which itineraries to make available for sale at each time point during the booking period. A widely used heuristic to make the capacity control decisions is a bid price control policy. A bid price control policy associates a bid price with each flight leg, which can be interpreted as the opportunity cost of a unit of capacity on that flight leg. In this case, an itinerary request is accepted only if the revenue generated by it exceeds the sum of the bid prices of the flight legs that are in the requested itinerary; see Talluri and van Ryzin (2004).

The literature on bid price controls can be broadly classified into two categories depending on the underlying demand model. The first category consists of methods that work under the so called independent demand assumption, which essentially states that each customer arrives into the system with the intention of purchasing a fixed itinerary. If this itinerary is available for purchase, then the customer purchases it. Otherwise, the customer leaves the system without purchasing anything. The classical method to obtain bid prices in this setting involves solving a linear program. This linear program is formulated under the assumption that itineraries can be sold in fractional quantities and the demand for an itinerary always takes on its expected value. In this linear program, there is one constraint for each flight leg, which ensures that the total capacity consumed by the accepted itinerary requests do not exceed the capacity of that flight leg. The optimal values of the dual variables associated with the flight leg capacity constraints can be viewed as an approximation to the opportunity cost of a unit of capacity on the flight legs. Therefore, these dual variables are used as bid prices. A clear deficiency of the linear program is that it is a deterministic approximation to a problem that actually takes place under uncertainty. Talluri and van Ryzin (1999) propose a randomized version of the linear program that compensates for this deficiency by working with samples of the demand random variables rather than their expected values. The latter linear program is solved for many samples of the demand random variables and the sample average of the optimal dual variables associated with the flight leg capacity constraints are used as the bid prices.

The second category of bid price control methods allow for the fact that there may be many itineraries that are acceptable to a customer and the purchasing decision of a customer may change depending on the set of itineraries that are made available for purchase. Many of the ideas from the independent demand setting can be transferred to the setting where customers choose among the available itineraries. For example, Liu and van Ryzin (2008) propose a linear program that can be viewed as an extension of the aforementioned deterministic linear program to handle customer choice behavior.

In this paper, we propose a randomized linear program to obtain bid prices for network revenue management problems with customer choice behavior. We start off with a mixed integer programming formulation, which is a somewhat nonstandard deterministic approximation to the network revenue management problem. This mixed integer program has a decision variable for each itinerary and each time period which indicates whether the itinerary is offered at that time period. The optimal solution to this mixed integer program therefore yields the set of itineraries to be offered at each time period. We next formulate a linear program that resembles the one used in the so called independent demand setting, but the key difference in our formulation is that the expected demand for an itinerary is not

exogenously given. Rather, it is given by the expected demand for the itinerary when we offer the itineraries in the optimal solution to the mixed integer program. The linear program that we propose can be easily randomized. We solve it for many different demand samples and use the average of the dual variables associated with the flight leg capacity constraints as our bid prices.

The approach that we propose has a number of appealing features. First, it can be easily implemented using commercially available optimization software. As a matter of fact, if the customer choices are governed by the multinomial logit model, then our solution method requires solving linear programs and linear mixed integer programs, which most commercial software packages are capable of. This is likely to boost the practical appeal of our approach. Moreover, our numerical study suggests that our approach is computationally efficient and the solution times are competitive with existing methods. Second, our approach allows randomization. As a result, the hope is that the bid prices generated by our approach better capture the uncertainty of the underlying network revenue management problem and perform better than the bid prices generated by the standard linear program. Our computational experiments indicate that this indeed seems to be the case. Third, we show that our solution method approximates the optimal total expected revenue arbitrarily closely in an asymptotic regime where the leg capacities and the number of time periods in the decision horizon increase linearly with the same rate. This asymptotic result provides a theoretical underpinning for our method and may explain to a certain degree its good performance in our computational experiments.

The randomized linear program that we propose generates bid prices that do not inherently depend on the remaining capacities on the flight legs. There has been some recent work on capacity dependent bid prices; see Zhang and Adelman (2009), Meissner and Strauss (2008) and Kunnumkal and Topaloglu (2009). Capacity dependent bid prices capture the intuitive notion that the opportunity cost of a unit of capacity should increase as the remaining capacity on the flight leg decreases and it is widely acknowledged that they can provide superior performance. However, the superior performance of capacity dependent bid price policies usually comes with substantial overhead, both in terms of computation and implementation, and there is still undeniable interest in the type of bid prices obtained by the randomized linear program; see Chaneton and Vulcano (2009). Moreover, starting with a set of bid prices that do not depend on the remaining capacities on the flight legs, we can obtain capacity dependent bid prices by using the dynamic programming decomposition technique of Liu and van Ryzin (2008) or the approaches described in Zhang (2009) or Meissner and Strauss (2010). Naturally, the performance of the resulting capacity dependent bid price policy depends on the starting values and so it becomes important to have a good set of initial bid prices. Our computational experiments indicate that the revenue gains carry over to the capacity dependent bid price policy that uses the bid prices generated by the randomized linear program as the starting value.

Our work is related to previous research in the network revenue management literature. The idea of a bid price control and generating bid prices through a deterministic linear program originated with the work of Simpson (1989) and Williamson (1992). Talluri and van Ryzin (1998) provide a theoretical basis for the deterministic linear program. In particular, they show that the deterministic linear program yields an upper bound on the optimal total expected revenue. They also establish an asymptotic

optimality result for this upper bound and the performance of the control policies obtained through the deterministic linear program. The deterministic linear program uses only the expected values of the random variables and discards all other distributional information. Talluri and van Ryzin (1999) propose a randomized linear program that tries to overcome this deficiency by using samples of the random variables. Computational experiments in Talluri and van Ryzin (1999) and Topaloglu (2009) indicate that the bid prices obtained by the randomized linear program can perform significantly better than the bid prices obtained by the deterministic linear program. The deterministic and randomized linear programs mentioned in this paragraph exclusively assume that customers come into the system with the intention of purchasing a fixed itinerary.

Incorporating customer choice behavior into the network revenue management problem is an active area of research and there is large body of work on the subject. Here, we only review the literature on methods to generate bid prices. Liu and van Ryzin (2008) propose a deterministic linear program for the network revenue management problem with customer choice behavior. They also show how the traditional dynamic programming decomposition method can be extended to the customer choice setting. Zhang and Adelman (2009), Meissner and Strauss (2008) and Zhang (2009) use the linear programming representation of the dynamic programming formulation of the revenue management problem to come with up different approximations to the value functions that appear in the dynamic programming formulation of the problem. Kunnumkal and Topaloglu (2008) and Kunnumkal and Topaloglu (2009) use Lagrangian relaxation ideas to obtain value function approximations. In both cases, bid price policies are obtained by using the value function approximations to estimate the opportunity costs of capacity on the flight legs. Chaneton and Vulcano (2009) use stochastic approximation to obtain bid prices. Talluri (2010) proposes a deterministic concave program for the network revenue management problem with customer choice behavior and describes a way to randomize it. However, the approach is quite different from what we propose in this paper, as his approach requires solving continuous concave optimization problems, whereas we work with linear integer programs.

In this paper, we make the following research contributions. 1) We propose a new method to generate bid prices for the network revenue management problem with customer choice behavior. Our method is based on a deterministic approximation to the network revenue management problem that is non-standard and allows randomization. As a result, the bid prices that we obtain may better capture the underlying uncertainty in the problem. 2) We show that the our solution method approximates the optimal total expected revenue arbitrarily closely in an asymptotic regime where the leg capacities and the number of time periods in the decision horizon increase linearly with the same rate. 3) Computational experiments indicate that the bid price policies obtained by our method can perform significantly better than the bid price policies obtained by the standard deterministic linear program.

The rest of the paper is organized as follows. Section 1 formulates the network revenue management problem with customer choice behavior as a dynamic program. Section 2 describes the linear program proposed by Liu and van Ryzin (2008). Section 3 builds on this linear program and develops the randomized method that we propose to obtain bid prices. Section 4 establishes the asymptotic optimality result described in the previous paragraph. Section 5 presents our computational experiments.

## 1 PROBLEM FORMULATION

We have an airline network consisting of a set of flight legs that we can use to serve the customers that arrive over time with the intention of purchasing itineraries. We use  $\mathcal{L}$  to denote the set of flight legs in the airline network. The initial capacity on flight leg  $i$  is  $c_i$ . We use  $\mathcal{J}$  to denote the set of itineraries that can be offered to the customers. An itinerary  $j$  has a revenue associated with it, which we denote by  $r_j$ . If we accept a request for itinerary  $j$ , then we consume capacity on one or more flight legs. We use  $a_{ij}$  to denote the number of units of capacity consumed by itinerary  $j$  on flight leg  $i$ . Naturally, we have  $a_{ij} = 0$  if itinerary  $j$  does not include flight leg  $i$ . We discretize the planning horizon into a finite number of time periods  $\mathcal{T} = \{1, \dots, \tau\}$  and assume that the discretization is fine enough so that there is at most one customer arrival at each time period. The probability of a customer arrival at time period  $t$  is  $\lambda$ . The fact that the arrival probability is constant over time is only for ease of exposition and it is straightforward to allow the arrival probability to depend on the time period  $t$ .

At each time period, we have to decide which itineraries to make available for sale taking into account the state of the remaining leg capacities. That is, we are constrained by the fact that an itinerary can be offered only if there is sufficient capacity on the flight legs that it uses. Using  $x_{it}$  to denote the remaining capacity on flight leg  $i$  at time period  $t$ ,  $x_t = \{x_{it} : i \in \mathcal{L}\}$  captures the state of the remaining leg capacities. We use  $u_t = \{u_{jt} : j \in \mathcal{J}\} \in \{0, 1\}^{|\mathcal{J}|}$  to denote the set of itineraries that we make available at time period  $t$ , with the interpretation that  $u_{jt}$  takes value 1 if itinerary  $j$  is offered at time period  $t$  and takes value 0 otherwise. The decision variables are constrained to lie in the set

$$\mathcal{U}(x_t) = \{u_t \in \{0, 1\}^{|\mathcal{J}|} : a_{ij}u_{jt} \leq x_{it} \quad \forall i \in \mathcal{L}, j \in \mathcal{J}\},$$

which ensures that if we want to offer itinerary  $j$  and itinerary  $j$  consumes capacity on flight leg  $i$ , then we have to have capacity available on flight leg  $i$ . Given that  $u_t$  is the set of itineraries that are offered, an arriving customer purchases itinerary  $j$  with probability  $P_j(u_t)$ . Naturally, we have  $P_j(u_t) = 0$  whenever  $u_{jt} = 0$ . We allow the possibility that the customer may not purchase any of the offered itineraries and let  $P_\phi(u_t) = 1 - \sum_{j \in \mathcal{J}} P_j(u_t)$  denote this probability. As per our notation, the probability  $P_j(\cdot)$  does not depend on the time period. Again, this is only for ease of exposition and it is straightforward to allow this probability to depend on the time period.

The decision problem is to determine the set of itineraries to offer to the customers at each time period so as to maximize the expected total revenue over the planning horizon. Under the assumption that the customer arrivals in the different time periods and the purchasing decisions of the different customers are independent of each other, we can obtain the value functions  $\{V_t(\cdot) : t \in \mathcal{T}\}$  through the optimality equation

$$\begin{aligned} V_t(x_t) &= \max_{u_t \in \mathcal{U}(x_t)} \left\{ \sum_{j \in \mathcal{J}} \lambda P_j(u_t) \left[ r_j + V_{t+1}(x_t - \sum_{i \in \mathcal{L}} a_{ij} e_i) \right] + \left[ 1 - \lambda + \lambda P_\phi(u_t) \right] V_{t+1}(x_t) \right\} \\ &= \max_{u_t \in \mathcal{U}(x_t)} \left\{ \sum_{j \in \mathcal{J}} \lambda P_j(u_t) \left[ r_j + V_{t+1}(x_t - \sum_{i \in \mathcal{L}} a_{ij} e_i) - V_{t+1}(x_t) \right] \right\} + V_{t+1}(x_t), \end{aligned} \quad (1)$$

where  $e_i$  is the  $|\mathcal{L}|$ -dimensional unit vector with a one in the element corresponding to  $i \in \mathcal{L}$  and the second equality follows from the fact that  $P_\phi(u_t) = 1 - \sum_{j \in \mathcal{J}} P_j(u_t)$ . The boundary condition

for the optimality equation above is  $V_{\tau+1}(\cdot) = 0$ . Throughout the rest of the paper, we assume that  $\lambda = 1$  for notational brevity. We note that this is equivalent to letting  $\tilde{P}_j(u_t) = \lambda P_j(u_t)$  and  $\tilde{P}_\phi(u_t) = 1 - \lambda + \lambda P_\phi(u_t)$  and working with the probabilities  $\{\tilde{P}_j(u_t) : j \in \mathcal{J}\}$ .

If we let  $\mathcal{C}_i = \{0, \dots, c_i\}$ , then the state space of the above dynamic program is  $\prod_{i \in \mathcal{L}} \mathcal{C}_i$ . Therefore, the size of the state space increases exponentially with the number of flight legs in the airline network and it becomes computationally quite difficult to solve the optimality equation for practical problem instances. In the next two sections, we describe approximate methods that can be used to decide which itineraries to offer to the customers at each time period in the booking horizon.

## 2 CHOICE BASED DETERMINISTIC LINEAR PROGRAM

Liu and van Ryzin (2008) propose a deterministic linear program for the network revenue management problem with customer choice behavior. This linear program replaces random quantities by their expected values and assumes that itineraries can be sold in fractional amounts. With a slight abuse of notation, we let  $P_j(\mathcal{S})$  be the probability that a customer chooses itinerary  $j$  given that  $\mathcal{S}$  is the set of itineraries that are offered. The deterministic linear program is based on the assumption that if  $\mathcal{S}$  is the set of itineraries that are offered at a time period, then the number of sales for itinerary  $j$  at that time period is  $P_j(\mathcal{S})$ . In this case, the revenue generated by offering set  $\mathcal{S}$  is  $\sum_{j \in \mathcal{J}} r_j P_j(\mathcal{S})$  and the capacity consumed on flight leg  $i$  is  $\sum_{j \in \mathcal{J}} a_{ij} P_j(\mathcal{S})$ . The deterministic linear program determines the frequency with which each set should be offered in order to maximize the total revenue subject to the flight leg capacity constraints. Therefore, we have a decision variable  $h_t(\mathcal{S})$ , which denotes the frequency with which we offer set  $\mathcal{S}$  at time period  $t$  and the linear program that we solve is

$$\max \sum_{t \in \mathcal{T}} \sum_{\mathcal{S} \subset \mathcal{J}} \sum_{j \in \mathcal{J}} r_j P_j(\mathcal{S}) h_t(\mathcal{S}) \tag{2}$$

$$\text{subject to } \sum_{t \in \mathcal{T}} \sum_{\mathcal{S} \subset \mathcal{J}} \sum_{j \in \mathcal{J}} a_{ij} P_j(\mathcal{S}) h_t(\mathcal{S}) \leq c_i \quad \text{for all } i \in \mathcal{L} \tag{3}$$

$$\sum_{\mathcal{S} \subset \mathcal{J}} h_t(\mathcal{S}) = 1 \quad \text{for all } t \in \mathcal{T} \tag{4}$$

$$h_t(\mathcal{S}) \geq 0 \quad \text{for all } \mathcal{S} \subset \mathcal{J}, t \in \mathcal{T}. \tag{5}$$

In the above problem, the first set of constraints ensure that the total capacity consumed by the itineraries does not exceed the capacity of each flight leg. The second set of constraints ensure that the total frequency with which we offer the sets at each time period is equal to one. Since the empty set is a subset of  $\mathcal{J}$ , the second set of constraints allow not offering any itinerary with a certain frequency. We note that the number of decision variables in problem (2)-(5) increases exponentially with the number of itineraries. However, since the number of constraints is manageable, problem (2)-(5) can be solved by using column generation; see Liu and van Ryzin (2008).

Problem (2)-(5) is useful for three main reasons. First, the optimal objective value of problem (2)-(5) provides an upper bound on the optimal total expected revenue. That is, letting  $Z_{LP}$  denote the optimal objective value of problem (2)-(5) and  $c = \{c_i : i \in \mathcal{L}\}$ , we have  $V_1(c) \leq Z_{LP}$ ; see Liu

and van Ryzin (2008). The upper bound property is useful since it provides a benchmark to assess the performance of various heuristic control policies.

Second, we can use the solution to problem (2)-(5) to come up with a bid price control policy for the network revenue management problem. In particular, letting  $\hat{\mu} = \{\hat{\mu}_i : i \in \mathcal{L}\}$  be the optimal values of the dual variables associated with constraints (3), we can use  $\hat{\mu}_i$  as the bid price of flight leg  $i$ , with the intuition that  $\hat{\mu}_i$  approximates the opportunity cost of a unit of capacity on flight leg  $i$ . Therefore, we can approximate the value function  $V_t(x_t)$  by a linear function of the form  $\sum_{i \in \mathcal{L}} \hat{\mu}_i x_{it}$ , plug this approximation into the right side of optimality equation (1) and solve the problem

$$\max_{u_t \in \mathcal{U}(x_t)} \left\{ \sum_{j \in \mathcal{J}} P_j(u_t) \left[ r_j - \sum_{i \in \mathcal{L}} a_{ij} \hat{\mu}_i \right] \right\} + \sum_{i \in \mathcal{L}} \hat{\mu}_i x_{it} \quad (6)$$

to decide which set of itineraries to offer at time period  $t$ . This idea is described in Zhang and Adelman (2009) and it can be viewed as an extension of the traditional bid pricing approach to the setting with customer choice behavior.

Third, we can use the solution to problem (2)-(5) to decompose the network revenue management problem into a number of single flight leg problems. This is the dynamic programming decomposition method described in Liu and van Ryzin (2008). Given an initial estimate of the opportunity costs of capacity,  $\hat{\mu}$ , the portion of the revenue of itinerary  $j$  that is allocated to flight leg  $i$  is taken to be  $r_{ij}(\hat{\mu}) = r_j - \sum_{k \in \mathcal{L} \setminus \{i\}} a_{kj} \hat{\mu}_k$ . We solve the dynamic program

$$\vartheta_{it}(x_{it} | \hat{\mu}) = \max_{u_t \in \mathcal{U}_i(x_{it})} \left\{ \sum_{j \in \mathcal{J}} P_j(u_t) \left[ r_{ij}(\hat{\mu}) + \vartheta_{i,t+1}(x_{it} - a_{ij} | \hat{\mu}) - \vartheta_{i,t+1}(x_{it} | \hat{\mu}) \right] \right\} + \vartheta_{i,t+1}(x_{it} | \hat{\mu}), \quad (7)$$

for flight leg  $i$ , where  $\mathcal{U}_i(x_{it}) = \{u_t \in \{0, 1\}^{|\mathcal{J}|} : a_{ij} u_{jt} \leq x_{it} \forall j \in \mathcal{J}\}$  and  $\vartheta_{i,\tau+1}(\cdot | \hat{\mu}) = 0$ . Noting that  $\vartheta_{it}(\cdot | \hat{\mu})$  approximates the value of capacity on flight leg  $i$  at time period  $t$ , we can approximate the value function  $V_t(x_t)$  by  $\sum_{i \in \mathcal{L}} \vartheta_{it}(x_{it} | \hat{\mu})$ , plug this approximation into the right side of optimality equation (1) and solve the problem

$$\max_{u_t \in \mathcal{U}(x_t)} \left\{ \sum_{j \in \mathcal{J}} P_j(u_t) \left[ r_j - \sum_{i \in \mathcal{L}} \left\{ \vartheta_{i,t+1}(x_{it} | \hat{\mu}) - \vartheta_{i,t+1}(x_{it} - a_{ij} | \hat{\mu}) \right\} \right] \right\} + \sum_{i \in \mathcal{L}} \vartheta_{i,t+1}(x_{it} | \hat{\mu}) \quad (8)$$

to decide which set of itineraries to offer at time period  $t$ . Note that we can interpret  $\vartheta_{i,t+1}(x_{it} | \hat{\mu}) - \vartheta_{i,t+1}(x_{it} - a_{ij} | \hat{\mu})$  as the bid price of flight leg  $i$  and the bid price depends on the remaining capacity on the flight leg.

### 3 RANDOMIZED LINEAR PROGRAM

In this section, we describe our approach to generate bid prices for the network revenue management problem with customer choice behavior. We ultimately use a randomized linear program to generate bid prices, but our development proceeds in two stages. First, we solve a mixed integer program that can be viewed as a deterministic approximation to the optimality equation in (1). Next, we use the



solution to this mixed integer program to formulate and solve a randomized linear program. The mixed integer program that we solve has decision variables  $\{z_{jt} : j \in \mathcal{J}, t \in \mathcal{T}\}$  and  $\{y_j : j \in \mathcal{J}\}$ . In particular,  $z_{jt}$  is a binary decision variable that indicates whether itinerary  $j$  is offered at time period  $t$ . Note that the total expected demand for itinerary  $j$  as a function of  $\{z_{jt} : j \in \mathcal{J}, t \in \mathcal{T}\}$  is given by  $\sum_{t \in \mathcal{T}} P_j(z_t)$ , where we let  $z_t = \{z_{jt} : j \in \mathcal{J}\}$ . In this case, letting  $y_j$  denote the portion of the total demand for itinerary  $j$  that is accepted, we solve the mixed integer program

$$\max \sum_{j \in \mathcal{J}} r_j y_j \tag{9}$$

$$\text{subject to } \sum_{j \in \mathcal{J}} a_{ij} y_j \leq c_i \quad \text{for all } i \in \mathcal{L} \tag{10}$$

$$y_j \leq \sum_{t \in \mathcal{T}} P_j(z_t) \quad \text{for all } j \in \mathcal{J} \tag{11}$$

$$y_j \geq 0, z_{jt} \in \{0, 1\} \quad \text{for all } j \in \mathcal{J}, t \in \mathcal{T}. \tag{12}$$

Constraints (10) ensure that the total capacity consumed by the accepted itinerary requests does not exceed the capacity of each flight leg, whereas constraints (11) ensure that we do not accept more itinerary requests than the expected demand for the itinerary.

Intuitively, we would like to use the values of the dual variables corresponding to constraints (10) as estimates of the opportunity costs of capacity. However, dual variables are not well defined for problem (9)-(12) as it is a mixed integer program. This motivates the second step of our solution method. Let  $\hat{y} = \{\hat{y}_j : j \in \mathcal{J}\}$ ,  $\hat{z}_t = \{\hat{z}_{jt} : j \in \mathcal{J}\}$  for all  $t \in \mathcal{T}$  denote an optimal solution to problem (9)-(12). We let  $D(\hat{z}) = \{D_{jt}(\hat{z}_t) : j \in \mathcal{J}, t \in \mathcal{T}\}$  be random variables such that  $D_{jt}(\hat{z}_t) \in \{0, 1\}$  and  $\mathbb{P}\{D_{jt}(\hat{z}_t) = 1\} = P_j(\hat{z}_t)$ . Note that  $D_{jt}(\hat{z}_t)$  can be viewed as the demand for itinerary  $j$  when we offer the set  $\hat{z}_t$ . We now solve the linear program

$$\max \sum_{j \in \mathcal{J}} r_j y_j \tag{13}$$

$$\text{subject to } \sum_{j \in \mathcal{J}} a_{ij} y_j \leq c_i \quad \text{for all } i \in \mathcal{L} \tag{14}$$

$$y_j \leq \sum_{t \in \mathcal{T}} D_{jt}(\hat{z}_t) \quad \text{for all } j \in \mathcal{J} \tag{15}$$

$$y_j \geq 0 \quad \text{for all } j \in \mathcal{J}. \tag{16}$$

The decision variables in this problem are  $\{y_j : j \in \mathcal{J}\}$ , where  $y_j$  is the total number of requests for itinerary  $j$  that are accepted. Constraints (14) are analogous to constraints (10). Noting that  $\mathbb{E}\{D_{jt}(\hat{z}_t)\} = P_j(\hat{z}_t)$ , we can interpret constraints (15) as being analogous to constraints (11), but with the expected value of the random variable replaced by its sample. Since the right side of constraints (15) is random, the optimal solution to problem (13)-(16) is also random. Letting  $\hat{\pi}(D(\hat{z})) = \{\hat{\pi}_i(D(\hat{z})) : i \in \mathcal{L}\}$  be the optimal values of the dual variables corresponding to constraints (14), we use  $\mathbb{E}\{\hat{\pi}(D(\hat{z}))\} = \{\mathbb{E}\{\hat{\pi}_i(D(\hat{z}))\} : i \in \mathcal{L}\}$  as estimates of the opportunity costs of capacity.

To summarize our solution method, we first solve problem (9)-(12) to obtain  $\hat{z}_t = \{\hat{z}_{jt} : j \in \mathcal{J}\}$  for all  $t \in \mathcal{T}$ . We then solve the linear program (13)-(16) and use the expected values of the optimal

dual variables corresponding to constraints (14),  $\mathbb{E}\{\hat{\pi}(D(\hat{z}))\}$ , as estimates of the opportunity costs of capacity. The exact computation of the expectation  $\mathbb{E}\{\hat{\pi}(D(\hat{z}))\}$  can be difficult in practice, but we resolve this difficulty by estimating the expectation through Monte Carlo samples. This amounts to solving problem (13)-(16) for  $N$  independent samples of  $D(\hat{z})$ , say  $D^n(\hat{z})$  for  $n = 1, \dots, N$ . Letting  $\hat{\pi}(D^n(\hat{z})) = \{\hat{\pi}_i(D^n(\hat{z})) : i \in \mathcal{L}\}$  be the optimal values of the dual variables corresponding to constraints (14) when we use demand sample  $D^n(\hat{z})$ , we can estimate  $\mathbb{E}\{\hat{\pi}(D(\hat{z}))\}$  by  $\bar{\pi} = \frac{1}{N} \sum_{n=1}^N \hat{\pi}(D^n(\hat{z}))$ . We can directly use  $\bar{\pi}$  to decide on the set of itineraries to offer at time period  $t$  by solving the problem

$$\max_{u_t \in \mathcal{U}(x_t)} \left\{ \sum_{j \in \mathcal{J}} P_j(u_t) \left[ r_j - \sum_{i \in \mathcal{L}} a_{ij} \bar{\pi}_i \right] \right\} + \sum_{i \in \mathcal{L}} \bar{\pi}_i x_{it}. \quad (17)$$

Note that this is equivalent to approximating  $V_t(x_t)$  by the linear function  $\sum_{i \in \mathcal{L}} \bar{\pi}_i x_{it}$ . Alternatively, we can use  $\bar{\pi}$  as the initial estimate of the opportunity costs of capacity in the dynamic programming decomposition method described in Section 2 and use the resulting value function approximations  $\{\vartheta_{it}(\cdot | \bar{\pi}) : i \in \mathcal{L}, t \in \mathcal{T}\}$  to decide on the set of itineraries to offer at time period  $t$  by solving problem (8).

In closing this section, we note that solving problem (9)-(12) for an arbitrary choice model can be difficult as  $P_j(z_t)$  can depend on  $z_t$  in an arbitrary fashion. However, for the multinomial logit choice model, Theorem 2 in Zhang and Adelman (2009) shows that we can solve problem (9)-(12) as a linear mixed integer program.

#### 4 ASYMPTOTIC OPTIMALITY

In this section, we show that the percent gap between the optimal objective value of problem (9)-(12) and the optimal total expected revenue diminishes to zero in a regime where the leg capacities and the number of time periods in the decision horizon increase linearly with the same rate. This result implies that the optimal objective value of problem (9)-(12) approximates the optimal total expected revenue more precisely as the scale of the problem increases, where scale is measured in terms on the leg capacities and the length of the decision horizon. We first show that the optimal objective value of problem (2)-(5) is an upper bound on the optimal objective value of problem (9)-(12). We then show that the percent gap between the optimal objective values of problems (9)-(12) and (2)-(5) diminishes to zero in the asymptotic regime. Proposition 2 in Liu and van Ryzin (2008) establishes that the percent gap between the optimal total expected revenue and the optimal objective value of problem (2)-(5) diminishes to zero in the same asymptotic regime. This then implies that the percent gap between the optimal objective value of problem (9)-(12) and the optimal total expected revenue also diminishes to zero asymptotically. Our results require the following assumption on the customer choice model.

**Assumption 1** *Let  $z, \tilde{z} \in \{0, 1\}^{|\mathcal{J}|}$  be such that  $\tilde{z}_l = 0$  for some  $l \in \mathcal{J}$  and  $\tilde{z}_j = z_j$  for all  $j \in \mathcal{J} \setminus \{l\}$ . Then we have  $P_j(\tilde{z}) \geq P_j(z)$  for all  $j \in \mathcal{J} \setminus \{l\}$ .*

Assumption 1 simply states that the probability of a customer choosing an itinerary  $j$  does not decrease if an itinerary  $l \neq j$  is removed from the set of itineraries offered. This assumption is satisfied,

for example, by the multinomial logit choice model, which is a commonly used choice model in the network revenue management literature. We begin with the following lemma.

**Lemma 1** *Let  $Z_{MIP}$  and  $Z_{LP}$ , respectively, denote the optimal objective values of problems (9)-(12) and (2)-(5). If Assumption 1 holds, then we have that  $Z_{MIP} \leq Z_{LP}$ .*

**Proof:** Let  $\hat{\mu} = \{\hat{\mu}_i : i \in \mathcal{L}\}$  be the optimal values of the dual variables associated with constraints (3) in problem (2)-(5). Dualizing constraints (3) and rearranging terms, we have by strong duality that

$$\begin{aligned} Z_{LP} = \max \quad & \sum_{t \in \mathcal{T}} \sum_{\mathcal{S} \subset \mathcal{J}} \sum_{j \in \mathcal{J}} [r_j - \sum_{i \in \mathcal{L}} a_{ij} \hat{\mu}_i] P_j(\mathcal{S}) h_t(\mathcal{S}) + \sum_{i \in \mathcal{L}} \hat{\mu}_i c_i \\ \text{subject to} \quad & (4), (5). \end{aligned}$$

Letting  $\hat{\mathcal{S}} = \operatorname{argmax}_{\mathcal{S} \subset \mathcal{J}} \sum_{j \in \mathcal{J}} [r_j - \sum_{i \in \mathcal{L}} a_{ij} \hat{\mu}_i] P_j(\mathcal{S})$ , it is easy to see that an optimal solution to the above maximization problem has  $h_t(\hat{\mathcal{S}}) = 1$  and  $h_t(\mathcal{S}) = 0$  for all  $\mathcal{S} \in 2^{\mathcal{J}} \setminus \{\hat{\mathcal{S}}\}$ , where  $2^{\mathcal{J}}$  denotes the set of all possible subsets of  $\mathcal{J}$ . Therefore, we can restrict the variables  $h_t(\mathcal{S})$  to be binary and still get the same optimal objective value. That is, we have

$$\begin{aligned} Z_{LP} = \max \quad & \sum_{t \in \mathcal{T}} \sum_{\mathcal{S} \subset \mathcal{J}} \sum_{j \in \mathcal{J}} [r_j - \sum_{i \in \mathcal{L}} a_{ij} \hat{\mu}_i] P_j(\mathcal{S}) h_t(\mathcal{S}) + \sum_{i \in \mathcal{L}} \hat{\mu}_i c_i \\ \text{subject to} \quad & \sum_{\mathcal{S} \subset \mathcal{J}} h_t(\mathcal{S}) = 1 \quad \text{for all } t \in \mathcal{T} \\ & h_t(\mathcal{S}) \in \{0, 1\} \quad \text{for all } \mathcal{S} \subset \mathcal{J}, t \in \mathcal{T}. \end{aligned}$$

Noting that the above maximization problem can be equivalently written in terms of the binary decision variables  $\{z_{jt} : j \in \mathcal{J}, t \in \mathcal{T}\}$ , we have

$$\begin{aligned} Z_{LP} = \max \quad & \sum_{t \in \mathcal{T}} \sum_{j \in \mathcal{J}} [r_j - \sum_{i \in \mathcal{L}} a_{ij} \hat{\mu}_i] P_j(z_t) + \sum_{i \in \mathcal{L}} \hat{\mu}_i c_i \\ \text{subject to} \quad & z_{jt} \in \{0, 1\} \quad \text{for all } j \in \mathcal{J}, t \in \mathcal{T}. \end{aligned}$$

By Assumption 1, it follows that an optimal solution to the above maximization problem has  $z_{jt} = 0$  for any itinerary  $j$  with  $r_j - \sum_{i \in \mathcal{L}} a_{ij} \hat{\mu}_i < 0$ . Therefore, we have

$$Z_{LP} = \max \quad \sum_{j \in \mathcal{J}} [r_j - \sum_{i \in \mathcal{L}} a_{ij} \hat{\mu}_i] y_j + \sum_{i \in \mathcal{L}} \hat{\mu}_i c_i \quad (18)$$

$$\text{subject to} \quad 0 \leq y_j \leq \sum_{t \in \mathcal{T}} P_j(z_t) \quad \text{for all } j \in \mathcal{J} \quad (19)$$

$$z_{jt} \in \{0, 1\} \quad \text{for all } j \in \mathcal{J}, t \in \mathcal{T}. \quad (20)$$

Finally, letting  $\{\hat{y}_j : j \in \mathcal{J}\}, \{\hat{z}_{jt} : j \in \mathcal{J}, t \in \mathcal{T}\}$  be an optimal solution to problem (9)-(12), we have

$$Z_{LP} \geq \sum_{j \in \mathcal{J}} [r_j - \sum_{i \in \mathcal{L}} a_{ij} \hat{\mu}_i] \hat{y}_j + \sum_{i \in \mathcal{L}} \hat{\mu}_i c_i = \sum_{j \in \mathcal{J}} r_j \hat{y}_j + \sum_{i \in \mathcal{L}} \hat{\mu}_i [c_i - \sum_{j \in \mathcal{J}} a_{ij} \hat{y}_j] \geq Z_{MIP},$$

where the first inequality holds since  $\{\hat{y}_j : j \in \mathcal{J}\}, \{\hat{z}_{jt} : j \in \mathcal{J}, t \in \mathcal{T}\}$  is a feasible but not necessarily optimal solution to problem (18)-(20) and the last inequality holds since  $\hat{\mu}_i \geq 0$  for all  $i \in \mathcal{L}$  and  $\{\hat{y}_j : j \in \mathcal{J}\}$  satisfies constraints (10).  $\square$

Liu and van Ryzin (2008) show that  $Z_{LP}$  provides an upper bound on the optimal total expected revenue. An analogous result does not hold for problem (9)-(12) and it is possible to come up with examples where we have  $Z_{MIP} < V_1(c)$ . However, we show below that the percent gap between the optimal objective value of problem (9)-(12) and the optimal total expected revenue diminishes to zero in a regime where the leg capacities and the number of time periods in the decision horizon increase linearly with the same rate. For this purpose, we consider a family of network revenue management problems  $\{\mathcal{P}^\theta : \theta \in \mathbb{Z}_+\}$  parameterized by the scaling parameter  $\theta$ . Problem  $\mathcal{P}^\theta$  takes place over the decision horizon  $\mathcal{T}^\theta = \{1, \dots, \theta\tau\}$  and the initial leg capacity on flight leg  $i$  in this problem is  $\theta c_i$ . All other parameters of problem  $\mathcal{P}^\theta$  are the same as those described in Section 1. This is a standard way of scaling the problem in the revenue management literature to obtain asymptotic results; see Talluri and van Ryzin (1998).

Let  $Z_{MIP}^\theta$  and  $Z_{LP}^\theta$  respectively denote the optimal objective values of problems (9)-(12) and (2)-(5) when these problems are solved with decision horizon  $\mathcal{T}^\theta$  and leg capacities  $\{\theta c_i : i \in \mathcal{L}\}$ . Note that we have  $Z_{MIP}^1 = Z_{MIP}$  and  $Z_{LP}^1 = Z_{LP}$ . We show below that  $\lim_{\theta \rightarrow \infty} Z_{MIP}^\theta / Z_{LP}^\theta = 1$ . Lemma 1 implies that  $\limsup_{\theta \rightarrow \infty} Z_{MIP}^\theta / Z_{LP}^\theta \leq 1$ . We next establish that  $\liminf_{\theta \rightarrow \infty} Z_{MIP}^\theta / Z_{LP}^\theta \geq 1$ .

**Lemma 2** *We have  $\liminf_{\theta \rightarrow \infty} Z_{MIP}^\theta / Z_{LP}^\theta \geq 1$ .*

**Proof:** Let  $\{\hat{h}_t(\mathcal{S}) : t \in \mathcal{T}, \mathcal{S} \subset \mathcal{J}\}$  be an optimal solution to problem (2)-(5). We construct a feasible solution to problem (9)-(12) with decision horizon  $\mathcal{T}^\theta$  and leg capacities  $\{\theta c_i : i \in \mathcal{L}\}$  in the following manner. We offer a nonempty set  $\mathcal{S}$  a total of  $\lfloor \theta \hat{h}_t(\mathcal{S}) \rfloor$  times over the time periods  $\{t, \tau + t, \dots, (\theta - 1)\tau + t\}$  for  $t \in \mathcal{T}$ , where the  $\lfloor \cdot \rfloor$  operator rounds a number down to its nearest integer. That is, letting  $\mathcal{S}_{(n-1)\tau+t}^\theta$  denote the set of itineraries offered at time period  $(n-1)\tau + t$  for  $n \in \{1, 2, \dots, \theta\}$ ,  $t \in \mathcal{T}$ , we have  $\sum_{n=1}^\theta \mathbf{1}(\mathcal{S}_{(n-1)\tau+t}^\theta = \mathcal{S}) = \lfloor \theta \hat{h}_t(\mathcal{S}) \rfloor$  for all  $t \in \mathcal{T}$  for a nonempty set  $\mathcal{S} \subset \mathcal{J}$ . In this solution, we offer the empty set in the remaining  $\theta - \sum_{\mathcal{S} \subset \mathcal{J}, \mathcal{S} \neq \emptyset} \lfloor \theta \hat{h}_t(\mathcal{S}) \rfloor$  periods. That is, we have  $\sum_{n=1}^\theta \mathbf{1}(\mathcal{S}_{(n-1)\tau+t}^\theta = \emptyset) = \theta - \sum_{\mathcal{S} \subset \mathcal{J}, \mathcal{S} \neq \emptyset} \lfloor \theta \hat{h}_t(\mathcal{S}) \rfloor$ . Letting  $z_{j,(n-1)\tau+t}^\theta = \mathbf{1}(j \in \mathcal{S}_{(n-1)\tau+t}^\theta)$  and  $z_t^\theta = \{z_{jt}^\theta : j \in \mathcal{J}\}$ , we obtain

$$\sum_{t \in \mathcal{T}^\theta} P_j(z_t^\theta) = \sum_{t \in \mathcal{T}} \sum_{n=1}^\theta P_j(z_{(n-1)\tau+t}^\theta) = \sum_{t \in \mathcal{T}} \sum_{n=1}^\theta \sum_{\mathcal{S} \subset \mathcal{J}} P_j(\mathcal{S}) \mathbf{1}(\mathcal{S}_{(n-1)\tau+t}^\theta = \mathcal{S}) = \sum_{t \in \mathcal{T}} \sum_{\mathcal{S} \subset \mathcal{J}} \lfloor \theta \hat{h}_t(\mathcal{S}) \rfloor P_j(\mathcal{S})$$

for all  $j \in \mathcal{J}$ , where the last equality uses the fact that  $P_j(\emptyset) = 0$ . If we let  $y_j^\theta = \sum_{t \in \mathcal{T}^\theta} P_j(z_t^\theta)$ , then we immediately have that constraint (11) is satisfied. We also have the chain of inequalities  $\sum_{j \in \mathcal{J}} a_{ij} y_j^\theta = \sum_{j \in \mathcal{J}} \sum_{t \in \mathcal{T}} \sum_{\mathcal{S} \subset \mathcal{J}} a_{ij} P_j(\mathcal{S}) \lfloor \theta \hat{h}_t(\mathcal{S}) \rfloor \leq \theta (\sum_{j \in \mathcal{J}} \sum_{t \in \mathcal{T}} \sum_{\mathcal{S} \subset \mathcal{J}} a_{ij} P_j(\mathcal{S}) \hat{h}_t(\mathcal{S})) \leq \theta c_i$ , where the last inequality follows from (3). Therefore  $(y^\theta, z^\theta)$  is feasible to problem (9)-(12) with decision horizon  $\mathcal{T}^\theta$

and leg capacities  $\{\theta c_i : i \in \mathcal{L}\}$ . We have

$$\begin{aligned} Z_{MIP}^\theta &\geq \sum_{j \in \mathcal{J}} r_j y_j^\theta = \sum_{t \in \mathcal{T}} \sum_{\mathcal{S} \subset \mathcal{J}} \sum_{j \in \mathcal{J}} r_j P_j(\mathcal{S}) [\theta \hat{h}_t(\mathcal{S})] \geq \sum_{t \in \mathcal{T}} \sum_{\mathcal{S} \subset \mathcal{J}} \sum_{j \in \mathcal{J}} r_j P_j(\mathcal{S}) (\theta \hat{h}_t(\mathcal{S}) - 1) \\ &= \theta \left( \sum_{t \in \mathcal{T}} \sum_{\mathcal{S} \subset \mathcal{J}} \sum_{j \in \mathcal{J}} r_j P_j(\mathcal{S}) \hat{h}_t(\mathcal{S}) \right) - \sum_{t \in \mathcal{T}} \sum_{\mathcal{S} \subset \mathcal{J}} \sum_{j \in \mathcal{J}} r_j P_j(\mathcal{S}) = Z_{LP}^\theta - \sum_{t \in \mathcal{T}} \sum_{\mathcal{S} \subset \mathcal{J}} \sum_{j \in \mathcal{J}} r_j P_j(\mathcal{S}), \end{aligned}$$

where the last equality uses the fact that  $Z_{LP}^\theta = \theta Z_{LP}$ ; see Liu and van Ryzin (2008). Dividing both sides of the above expression by  $Z_{LP}^\theta$ , we have  $\liminf_{\theta \rightarrow \infty} Z_{MIP}^\theta / Z_{LP}^\theta \geq 1$ .  $\square$

Lemmas 1 and 2 together yield the main result of this section.

**Proposition 1** *If Assumption 1 holds, then we have  $\lim_{\theta \rightarrow \infty} Z_{MIP}^\theta / Z_{LP}^\theta = 1$ .*

Letting  $\{V_t^\theta(\cdot) : t \in \mathcal{T}^\theta\}$  be the value functions obtained by solving the optimality equation in (1) with decision horizon  $\mathcal{T}^\theta$  and leg capacities  $\{\theta c_i : i \in \mathcal{L}\}$ , Liu and van Ryzin (2008) show that  $\lim_{\theta \rightarrow \infty} Z_{LP}^\theta / V_1^\theta(\theta c) = 1$ . An immediate corollary to Proposition 1 is that the same holds for the objective value of problem (9)-(12).

**Corollary 1** *If Assumption 1 holds, then we have  $\lim_{\theta \rightarrow \infty} Z_{MIP}^\theta / V_1^\theta(\theta c) = 1$ .*

## 5 COMPUTATIONAL EXPERIMENTS

In this section, we numerically test the performance of the bid price policy obtained by the randomized linear program.

### 5.1 COMPUTATIONAL SETUP

In all of our test problems, we assume that the customers choose among the offered itineraries according to the multinomial logit model with disjoint consideration sets. For brevity, we refer to this choice model simply as the logit model. The logit model assumes that there are multiple customer segments and each customer segment is interested only in a subset of all the available itineraries. Letting  $\mathcal{J}_l$  be the set of itineraries that are of interest to customer segment  $l$ , we further assume that each itinerary is of interest to one and exactly one customer segment. That is, we have  $\mathcal{J}_{l_1} \cap \mathcal{J}_{l_2} = \emptyset$  whenever  $l_1 \neq l_2$ . At each time period, a customer from segment  $l$  arrives with probability  $\lambda_l$ . If the set of itineraries that we offer at time period  $t$  is given by  $u_t = \{u_{jt} : j \in \mathcal{J}\}$  and a customer from segment  $l$  arrives, then this customer purchases itinerary  $j$  with probability

$$\mathbf{1}(j \in \mathcal{J}_l) \frac{v_j u_{jt}}{\sum_{k \in \mathcal{J}_l} v_k u_{kt} + v_l^0},$$

where  $v_j$  is a positive preference weight associated with itinerary  $j \in \mathcal{J}$  and  $v_l^0$  is the preference weight for customer segment  $l$  associated with purchasing nothing. Roughly speaking, the preference weight  $v_j$  characterizes the appeal of itinerary  $j$  to a customer that belongs to the segment  $l$  that satisfies  $j \in \mathcal{J}_l$ ,

whereas the preference weight  $v_l^0$  characterizes the appeal of purchasing nothing to a customer that belongs to the segment  $l$ .

The logit model is crucial for the computational tractability of the optimization problems that we solve. As mentioned earlier, Zhang and Adelman (2009) show that we can solve problem (9)-(12) as a linear mixed integer program when the customer choices are governed by the logit model. The logit model also turns out to be crucial for the computational tractability of the choice based deterministic linear program. In particular, Liu and van Ryzin (2008) show that the column generation subproblem for problem (2)-(5) is tractable under the logit model.

We present computational experiments on two groups of test problems, both of which are based on the setup in Zhang and Adelman (2009). The first group involves parallel flight legs that operate between the same origin destination pair. The second group involves hub-and-spoke networks.

## 5.2 BENCHMARK SOLUTION METHODS

We compare the performance of the bid price policy obtained by the randomized linear program with that obtained by the choice-based deterministic linear program. We also compare the performance of the dynamic programming decomposition method that uses the dual solution of the randomized linear program with that which uses the dual solution of the choice-based deterministic linear program.

*Bid price policy obtained by the randomized linear program (RLP)* This is the solution method described in Section 3 with the exception that in our implementation, we divide the decision horizon into five equal segments and resolve problem (9)-(12) at the beginning of each segment. In particular, we resolve problem (9)-(12) at the beginning of a segment by replacing the right hand side of constraints (10) with the current remaining leg capacities and the set of time periods  $\mathcal{T}$  with the current set of remaining time periods. We use the new optimal solution  $\hat{z}$  and solve problem (13)-(16) for  $N$  demand samples to obtain a fresh set of optimal dual values  $\{\hat{\pi}(\hat{D}^n(\hat{z})) : n = 1, \dots, N\}$ . Finally, we solve problem (17) to decide which itineraries to offer at time period  $t$ . We use  $N = 100$  for RLP in all of our computational experiments. Increasing the number of samples further did not seem to produce any significant performance improvements.

*Bid price policy obtained by the choice based deterministic linear program (CDLP)* This is the solution method described in Section 2. As with RLP, CDLP divides the decision horizon into five equal segments and resolves problem (2)-(5) at the beginning of each segment by replacing the right hand side of constraints (3) with the current remaining leg capacities and the set of time periods  $\mathcal{T}$  with the current set of remaining time periods. This yields a fresh set of optimal dual values  $\hat{\mu}$ . After this, CDLP solves problem (6) to decide on the set of itineraries to be offered at time period  $t$ .

*Dynamic programming decomposition based on the randomized linear program (DPD-RLP)* This is the dynamic programming decomposition method that uses the dual solution obtained from the randomized linear program as the initial estimate of the opportunity costs of capacity. As with RLP, we divide the decision horizon into five equal segments and resolve problem (9)-(12) at the beginning of each segment. We use the new optimal solution  $\hat{z}$  and solve problem (13)-(16) for  $N$  demand samples to obtain a fresh

set of optimal dual values  $\{\hat{\pi}(\hat{D}^n(\hat{z})) : n = 1, \dots, N\}$ . We solve the optimality equation in (7) with  $\hat{\pi} = \frac{1}{N} \sum_{n=1}^N \hat{\pi}(D^n(\hat{z}))$  as the initial estimate of the opportunity costs of capacity and obtain the value function approximations  $\{\vartheta_{it}(\cdot | \hat{\pi}) : i \in \mathcal{L}, t \in \mathcal{T}\}$ . After this, we solve problem (8) to decide which itineraries to offer at time period  $t$ .

*Dynamic programming decomposition based on the choice based deterministic linear program (DPD-CDLP)* This is the dynamic programming decomposition method that uses the dual solution obtained from the choice based deterministic linear program as the initial estimate of the opportunity cost of capacity. As with CDLP, we divide the decision horizon into five equal segments and resolve problem (2)-(5) at the beginning of each segment. We solve the optimality equation in (7) with the fresh set of optimal dual values  $\hat{\mu}$  as the initial estimate of the opportunity costs of capacity and obtain the value function approximations  $\{\vartheta_{it}(\cdot | \hat{\mu}) : i \in \mathcal{L}, t \in \mathcal{T}\}$ . After this, we solve problem (8) to decide which itineraries to offer at time period  $t$ .

### 5.3 COMPUTATIONAL RESULTS ON PARALLEL FLIGHT LEGS

We consider  $n$  flight legs that operate between the same origin destination pair. There is one itinerary associated with each flight leg so that the number of itineraries is also  $n$ . There is one customer segment and an arriving customer purchases at most one of the offered itineraries according to the logit model. The preference weights associated with the itineraries are generated from the Poisson distribution with mean 100. The preference weight associated with purchasing nothing is set to  $\sum_{j \in \mathcal{J}} v_j / 2$ . The revenue associated with an itinerary is generated from the uniform distribution over  $[10, 100]$ . We obtain different test problems by using different seeds to initialize the random number generator. We label our test problems by  $(n, \tau, k) \in \{6, 8\} \times \{200, 400\} \times \{1, 2, 3\}$ , where  $n$  is the number of flight legs,  $\tau$  is the number of time periods in the decision horizon and  $k$  is the seed used for generating the preference weights and the revenues. This experimental setup parallels the one in Zhang and Adelman (2009).

Table 1 shows the computational results. The first column in this table gives the characteristics of the problem by using  $(n, \tau, k)$ . The second column gives the optimal objective value of problem (2)-(5), which is an upper bound on the optimal total expected revenue. Columns three to six respectively give the total expected revenues obtained by RLP, CDLP, DPD-RLP and DPD-CDLP. We estimate these total revenues by simulating the decisions made by the four solution methods under multiple customer arrival trajectories. The seventh column gives the percent gap between the total expected revenues obtained by RLP and CDLP, while the last column gives the percent gap between the total expected revenues obtained by DPD-RLP and DPD-CDLP. The results indicate that RLP performs significantly better than CDLP and the average performance gap is about 9%. DPD-RLP and DPD-CDLP perform significantly better than RLP and CDLP, respectively. This is not surprising since the dynamic programming decomposition method produces bid prices that are capacity dependent and it is widely acknowledged that capacity dependent bid prices can provide superior performance. Interestingly, DPD-RLP performs significantly better than DPD-CDLP and the average performance gap is around 6%. Comparing the revenues with the upper bound in the second column, we see that RLP and DPD-RLP generate essentially the optimal total expected revenues for the smaller test problems.

Problem ( $n, \tau, k$ )	Upper Bound	Total Exp. Rev				% Gap between	
		RLP	CDLP	DPD-RLP	DPD-CDLP	RLP and CDLP	DPD-RLP and DPD-CDLP
(6, 200, 1)	6,093	6,085	5,291	6,093	5,763	13.05	5.41
(6, 200, 2)	4,309	4,105	3,751	4,308	4,308	12.93	0.02
(6, 200, 3)	4,030	4,015	3,523	4,030	3,864	12.25	4.11
(6, 400, 1)	9,010	8,997	8,078	9,010	8,733	10.21	3.07
(6, 400, 2)	7,600	7,582	6,730	7,600	7,312	11.24	3.78
(6, 400, 3)	8,450	8,435	7,679	8,450	8,190	8.96	3.08
(8, 200, 1)	6,916	6,649	5,983	6,784	6,053	10.02	10.78
(8, 200, 2)	7,215	7,005	6,237	7,067	6,386	10.96	9.64
(8, 200, 3)	7,535	7,366	6,647	7,399	6,680	9.76	9.72
(8, 400, 1)	9,670	9,275	8,907	9,639	8,999	3.97	6.64
(8, 400, 2)	11,350	11,032	10,276	11,316	10,474	6.85	7.44
(8, 400, 3)	7,719	7,420	6,919	7,693	7,026	6.75	8.67

Table 1: Comparison of the total expected revenues for the test problems with parallel flight legs.

#### 5.4 COMPUTATIONAL RESULTS ON HUB-AND-SPOKE NETWORKS

We consider an airline network that serves  $n$  spokes through a single hub. Half of the spokes have two parallel flights to the hub and the other half have two parallel flights from the hub. Figure 1 shows the structure of the airline network. Therefore, there are  $n/2$  spoke-to-hub origin destination pairs,  $n^2/4$  spoke-to-spoke origin destination pairs and  $n/2$  hub-to-spoke origin destination pairs. There are two itineraries associated with a spoke-to-hub or hub-to-spoke origin destination pair, corresponding to the flight legs that connect the origin destination pair. On the other hand, there are four itineraries associated with a spoke-to-spoke origin destination pair, corresponding to the four possible combinations of flight legs connecting that pair. There is a customer type associated with each origin destination pair in the airline network and an arriving customer purchases at most one of the itineraries that connect the origin destination pair that it is interested in. The preference weights are generated in the same fashion as in the test problems with parallel flight legs. The revenues associated with spoke-to-hub and hub-to-spoke itineraries are respectively generated from the uniform distributions over  $[1, 10]$  and  $[10, 100]$ . The revenue associated with a spoke-to-spoke itinerary is 95% of the sum of the revenues associated with the corresponding spoke-to-hub and hub-to-spoke itineraries. We obtain different test problems by using different seeds to initialize the random number generator. We label our test problems by  $(n, \tau, k) \in \{4, 6\} \times \{200, 400\} \times \{1, 2, 3\}$ , where  $n$  is the number of spokes and the interpretations of  $\tau$  and  $k$  are the same as those for the test problems with parallel flight legs.

Table 2 shows the computational results for the test problems with hub-and-spoke networks. The columns have the same interpretation as that for Table 1. The results display essentially the same trend that we observed for the test problems with parallel flight legs. RLP continues to perform better than CDLP and the average performance gap is about 2%. The performance gap between RLP and CDLP appears to decrease as the number of time periods in the planning horizon increases. However, even for the problems with the longest planning horizon, we observe performance gaps as large as 3%. Such performance gaps are considered quite significant in the network revenue management setting. On average, DPD-RLP performs about 2% better than DPD-CDLP. It is encouraging that the boost in



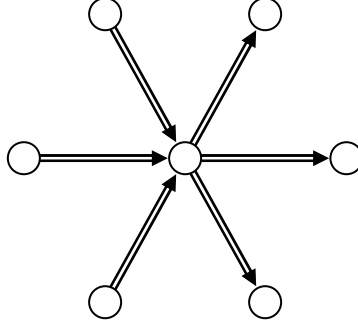


Figure 1: Structure of the hub-and-spoke network for the case where  $n = 6$ .

Problem ( $n, \tau, k$ )	Upper Bound	Total Exp. Rev				% Gap between	
		RLP	CDLP	DPD-RLP	DPD-CDLP	RLP and CDLP	DPD-RLP and DPD-CDLP
(4, 200, 1)	1,772	1,629	1,525	1,682	1,586	6.38	5.71
(4, 200, 2)	2,392	2,212	2,167	2,264	2,231	2.03	1.43
(4, 200, 3)	3,647	3,385	3,287	3,504	3,392	2.9	3.18
(4, 400, 1)	4,340	4,068	4,012	4,235	4,209	1.38	0.61
(4, 400, 2)	7,057	6,694	6,319	6,891	6,525	5.6	5.31
(4, 400, 3)	5,206	5,000	4,946	5,112	5,114	1.08	-0.04
(6, 200, 1)	3,773	3,376	3,352	3,518	3,464	0.71	1.53
(6, 200, 2)	3,032	2,752	2,685	2,847	2,800	2.43	1.65
(6, 200, 3)	3,311	3,020	2,294	3,111	3,013	3.18	3.16
(6, 400, 1)	7,152	6,589	6,550	6,960	6,824	0.59	1.96
(6, 400, 2)	5,454	5,079	4,986	5,281	5,161	1.83	2.26
(6, 400, 3)	5,382	5,084	4,913	5,270	5,154	3.86	2.20

Table 2: Comparison of the total expected revenues for the test problems with hub-and-spoke networks.

performance from RLP remains intact when we apply the dynamic programming decomposition method to it. This further illustrates the need for methods that provide good initial estimates of the opportunity cost of capacity.

Table 3 shows the CPU seconds for RLP and CDLP on a Pentium Core 2 Duo PC with 3 GHz CPU and 4 GB RAM. The CPU seconds for DPD-RLP and DPD-CDLP differ by the same amount and we do not report them separately. This is because once the randomized linear program and the choice based deterministic linear program have been solved, the additional work required is the same for both DPD-RLP and DPD-CDLP. We use CPLEX 11.2 to solve all the mixed integer programs and QSOPT to solve all the linear programs. The CPU seconds is primarily affected by the parameters  $n$  and  $\tau$ . Hence, we only provide the average CPU seconds over the test problems. The first column in Table 3 gives the problem characteristics. The second and third columns respectively show the CPU seconds required by RLP and CDLP. The CPU seconds for RLP correspond to the total time taken to solve problem (9)-(12) once and problem (13)-(16) for 100 samples. The CPU seconds for CDLP correspond to the time taken to solve problem (2)-(5). The CPU times are generally of the order of seconds and RLP takes at most a minute and a half to solve the largest test problem. Especially for the larger test problems, the CPU times for RLP and CDLP are of the same order of magnitude, signaling that RLP

Problem ( $n, \tau$ )	RLP CPU secs.	CDLP CPU secs.
Par. Fli. (6, 200)	1.2	0.1
Par. Fli. (6, 400)	3.2	0.4
Par. Fli. (8, 200)	3.8	0.3
Par. Fli. (8, 400)	7.6	1.1
Hub. Spk. (4, 200)	9.7	2.5
Hub. Spk. (4, 400)	24.7	12.1
Hub. Spk. (6, 200)	31.1	5.7
Hub. Spk. (6, 400)	84.8	20.6

Table 3: CPU seconds required for the different test problems.

is expected to be as attractive as CDLP for practical applications. However, we still caution the reader that a more extensive numerical study would be required in order to understand the applicability of RLP to real-world problem instances.

## 6 CONCLUSIONS

In this paper, we presented a new method to compute bid prices for the network revenue management problem with customer choice behavior. We first formulated and solved a mixed integer program that assumed that the demands for the itineraries took on their expected values. We used the optimal solution to this mixed integer program to formulate a linear program. The novel aspect of our linear programming formulation is that it naturally allows randomization and thus may better capture the probabilistic nature of the customer choices. We showed that our method approximates the optimal total expected revenue arbitrarily closely in a scaling of the network revenue management problem where the leg capacities and the number of time periods in the decision horizon increase linearly with the same rate. From a practical standpoint, our approach is appealing as it can be easily implemented by using commercially available software and it can provide substantial improvements over the standard choice based deterministic linear program.

## REFERENCES

- Chaneton, J. and Vulcano, G. (2009), Computing bid-prices for revenue management under customer choice behavior, Working paper, New York University, New York City, NY.
- Kunnumkal, S. and Topaloglu, H. (2008), ‘A refined deterministic linear program for the network revenue management problem with customer choice behavior’, *Naval Research Logistics Quarterly* **55**, 563–580.
- Kunnumkal, S. and Topaloglu, H. (2009), ‘A new dynamic programming decomposition method for the network revenue management problem with customer choice behavior’, *Production and Operations Management* . (to appear).
- Liu, Q. and van Ryzin, G. (2008), ‘On the choice-based linear programming model for network revenue management’, *Manufacturing & Service Operations Management* **10**(2).
- Meissner, J. and Strauss, A. K. (2008), Network revenue management with inventory sensitive bid prices and customer choice, Working paper, Department of Management Science, Lancaster University.
- Meissner, J. and Strauss, A. K. (2010), Improved bid prices for choice-based network revenue management, Working paper, Department of Management Science, Lancaster University.

- Simpson, R. W. (1989), Using network flow techniques to find shadow prices for market and seat inventory control, Mit flight transportation laboratory memorandum m89-1, cambridge, ma.
- Talluri, K. (2010), A randomized concave programming method for choice network revenue management, Working paper, Universitat Pompeu Fabra, Barcelona, Spain.
- Talluri, K. T. and van Ryzin, G. J. (2004), *The Theory and Practice of Revenue Management*, Kluwer Academic Publishers.
- Talluri, K. and van Ryzin, G. (1998), ‘An analysis of bid-price controls for network revenue management’, *Management Science* **44**(11), 1577–1593.
- Talluri, K. and van Ryzin, G. (1999), ‘A randomized linear programming method for computing network bid prices’, *Transportation Science* **33**(2), 207–216.
- Topaloglu, H. (2009), ‘On the asymptotic optimality of the randomized linear program for network revenue management’, *European Journal of Operational Research* **197**(3), 884–896.
- Williamson, E. L. (1992), Airline Network Seat Control, PhD thesis, MIT, Cambridge, MA.
- Zhang, D. (2009), An improved dynamic programming decomposition approach for network revenue management, Working paper, McGill University.
- Zhang, D. and Adelman, D. (2009), ‘An approximate dynamic programming approach to network revenue management with customer choice’, *Transportation Science* **43**, 381–394.