An Analysis of Bid-Price Controls for Network Revenue Management

Kalyan Talluri • Garrett van Ryzin
Universitat Pompeu Fabra, 27 Ramon Trias Fargas, 08005 Barcelona, Spain
412 Uris Hall, Columbia University, New York, New York 10027

Bid-prices are becoming an increasingly popular method for controlling the sale of inventory in revenue management applications. In this form of control, threshold—or ''bid''—prices are set for the resources or units of inventory (seats on flight legs, hotel rooms on specific dates, etc.) and a product (a seat in a fare class on an itinerary or room for a sequence of dates) is sold only if the offered fare exceeds the sum of the threshold prices of all the resources needed to supply the product. This approach is appealing on intuitive and practical grounds, but the theory underlying it is not well developed. Moreover, the extent to which bid-price controls represent optimal or near optimal policies is not well understood. Using a general model of the demand process, we show that bid-price control is not optimal in general and analyze why bid-price schemes can fail to produce correct accept/deny decisions. However, we prove that when leg capacities and sales volumes are large, bid-price controls are asymptotically optimal, provided the right bid prices are used. We also provide analytical upper bounds on the optimal revenue. In addition, we analyze properties of the asymptotically optimal bid prices. For example, we show they are constant over time, even when demand is nonstationary, and that they may not be unique.
(*Bid Prices*; *Optimality*; *Yield Management*; *Revenue Management*; *Airlines*; *Dynamic Programming*; *Heuristics*; *Asymptotic Analysis*)

Introduction and Overview

Bid-price control is a revenue management method in which threshold values (called *bid prices*) are set for each leg of a network and an itinerary (path on the network) is sold only if its fare exceeds the sum of the bid prices along the path. (Throughout the paper, the term *bid-price control* refers to such an *additive*, leg-based bid-price scheme. See Definition 1.) The technique, which originated with the work of Simpson (1989) at MIT and was later studied by Williamson (1992) in her Ph.D. thesis, is being adopted by a number of airlines and hotels. Indeed, it is fast becoming the method of choice for origin-destination (OD) revenue management.

Yet the theory underlying bid-price controls is scant. The early development by Simpson (1989) and Williamson (1992), while quite innovative and intuitively appealing, is based on different mathematical programming formulations of the OD control problem. However, the deterministic mathematical programming models used in these analyses are clearly oversimplified models of the true OD revenue management problem. Even the more sophisticated probabilistic mathematical programming formulations, such as those proposed by Glover et al. (1982), Williamson (1992) and Wollmer (1986) assume one-time, static allocations of capacity.

In this paper, we propose a general model of the OD control problem and analyze it via dynamic programming. The model incorporates demand uncertainty and makes no assumptions about the network structure or the sequence of arrivals (timing of high and low fare class arrivals). It allows for random fares within a fare class, which is of significant practical importance in some applications (see §1). Using this general model,

0025-1909/98/4411/1577$05.00

we formulate a dynamic program to analyze both the structure of the optimal control and the performance of bid price controls.

Related Literature

The study of revenue management problems (or yield management) in the airlines dates back to the work of Rothstein (1971) on an overbooking model and to Littlewood (1972) on a model of space allocation for a stochastic two-fare, single-leg (a network with one leg) problem. Belobaba (1987a, 1987b) proposed and tested a multiple-fare-class extension of Littlewood's rule which he termed the *expected marginal seat revenue* (EMSR) heuristic. Extensions and refinements of the multiple-fare-class problem include recent papers by Brumelle and McGill (1993), Curry (1989), Robinson (1991) and Wollmer (1992) (all of which with minor differences in models give the optimal nested seat allocations when fare classes book sequentially). A recent review of research on revenue management as well as a taxonomy of perishable asset revenue management (PARM) problems is given by Weatherford and Bodily (1992). See also Barnhart and Talluri (1996) for a recent survey on airline operations that covers the practice of revenue management.

Dynamic programming has been applied to analyze single-leg problems in prior work. For example, Lee and Hersh (1993) use discrete time dynamic programming to develop optimal rules for the single-leg problem when demand in each fare class is modeled as a stochastic process. In terms of modeling approaches, their work is closest to ours. Diamond and Stone (1991) and Feng and Gallego (1995) develop optimal threshold rules when demands are modeled as continuous time stochastic process. Other dynamic programming models for the single-leg problem are analyzed by Chatwin (1992) and Janakiram et al. (1994).

In a network setting, various mathematical programming approaches have been proposed. Glover et al. (1982) address a deterministic network flow model for the allocation of seats between passenger itineraries and fare classes. Wang (1983) provides an algorithm for the sequential allocations of seats on a plane to different origin-destination city pairs and fare classes within a flight segment when demands are random. Dror et al. (1988) present a rolling horizon network flow formulation for the seat inventory control problem assuming deterministic demands. In two internal McDonell Douglas reports, Wollmer (1986a) and (1986b) proposes a mathematical programming formulation incorporating random demands, where the objective is to maximize the total expected network revenue.

Yet to date, few dynamic programming models of network revenue management have been analyzed. An exception is the work of Gallego and van Ryzin (1994), who address a network revenue management problem using a continuous time, dynamic pricing model. Their bounding techniques and asymptotic analysis are quite similar to ours. The fundamental difference between this work and ours is that Gallego and van Ryzin assume prices are set for each itinerary (so the number of control variables equals the number of itineraries), and customers either accept or reject the offered prices. In contrast, we focus on bid-price controls, in which values are set for the legs (so the number of controls equals the number of legs) and a booking has a random revenue that can only be accepted or rejected. Our situation, therefore, models the case where fares are set exogenously, and the main goal of our work is to analyze the effectiveness of bid prices as a mechanism for making these accept/deny decisions.

As mentioned, Simpson (1989) and Williamson (1992) introduced the idea of bid-price controls and they proposed many of the main approximation approaches in the area. Williamson (1992) in particular used extensive simulation studies to analyze a variety of approaches to network revenue management. Simpson and Williamson's work had a significant impact on the practice of revenue management. However, they do not provide a rigorous analysis of the structure of the optimal network policy nor do they provide a theoretical foundation for the bid-price approach. Our work puts this important practical development on a sound theoretical footing.

Organization of the Paper

In §1 we formally define our network model and define bid-price control. In §2, we analyze the structure of the optimal policy and show that, in general, it is not a bid-price control. Two counter-examples are given in §3 to illustrate why a bid-price structure can be suboptimal. In §4, we develop an upper bound on the optimal revenue. This bound is used to show that a bid-price policy is asymptotically optimal as the sales volumes and capacities in the network grow. The analysis also provides a constructive method for computing a set of asymptotically optimal bid prices. In §5, we briefly analyze some of the more common methods for computing bid prices and discuss their strengths and weaknesses. Conclusions and a discussion of future research is presented in §6.

1. Formulation

We use superscripts to denote components of a vector and subscripts to denote time. We generally try to follow the convention that k denotes current time (time-to-go) while t denotes an arbitrary time. Time is counted backwards, so time t represents a point t periods from the end of the horizon. We do not distinguish between row and column vectors, since the proper interpretation is usually clear from the context. Finally, if A is a matrix, then the j-th column of A is denoted A^j and the ith row is denoted A_i. Finally, we let $x^+ = \max\{x, 0\}$, $\mathbf{1}\{E\}$ denote the indicator function of the event E and (a.s.) denotes *almost surely*.

Our model is formulated as follows: An airline network has m arcs or *legs* which can be used to provide n origin-destination *itineraries*. We let a_{ij} be the number of seats on leg i used by itinerary j ($a_{ij} = 0$ if leg i is not part of itinerary j). Define the matrix $A = [a_{ij}]$. Thus, the j-th column of A, A^j, is a multiple of the incidence vector for itinerary j (or the incidence vector itself if itinerary j requires only one seat). We use the notation $i \in A^j$ to indicate that leg i is used by itinerary j and $j \in A_i$ to mean that itinerary j uses leg i.

While it is simplest to imagine that each origin-destination itinerary requires only one unit of capacity from each of the legs it traverses, we do not impose this restriction. Indeed, the model can accommodate group requests (e.g. a family booking four seats together). To model this situation, we simply introduce one column in the matrix A for each possible group size, with the nonzero elements of the column equal to the group size and the itinerary revenue set equal to the total revenue of the group (e.g. for a group of size four, add a column j with $a_{ij} = 4$ for i on the path and with revenue R_t^j that is four times as large). The interpretation of the probability model in this case is that it reflects the likelihood of having requests for particular group sizes. (See Young and van Slyke (1994) for an exact analysis of monotonicity properties for a single-leg problem with group requests.)

The state of the network is described by a vector $x = (x^1, \ldots, x^m)$ of leg capacities. If itinerary j is sold, the state of the network changes to $x - A^j$. To keep our analysis simple, we will assume there are no cancellations or no-shows and, consequently, overbooking is not needed. Alternately, capacity may include so-called overbooking *pads*. (An overbooking *pad* is the number of seats an airline makes available for sale beyond the actual cabin capacity. These pads are sometimes set prior to performing fare class allocations.) If overbooking pads are computed independently, one can consider leg capacities in our model to be the sum of actual physical capacity and the overbooking pad.

In our formulation time is discrete, and k represents the number of periods left before departure. Within each time period, t, we assume that at most one request for an itinerary can arrive; that is, the discretization of time is sufficiently fine so that the probability of more than one request is negligible.

The way we model a request at time t is somewhat nonstandard but useful for analytical purposes. All booking events in time t are modeled as the realization of a *single* random vector $R_t = (R_t^1, \ldots, R_t^n)$. If $R_t^j > 0$, this indicates a request for itinerary j occurred and that its associated revenue is R_t^j; if $R_t^j = 0$, this indicates no request for j occurred. A realization $R_t = 0$ (all components equal to zero) indicates that no request from *any* itinerary occurred at time t. For example, if we have $n = 3$ itineraries, then a value $R_t = (0, 0, 0)$ indicates no requests arrived, a value $R_t = (120, 0, 0)$ indicates a request for itinerary 1 with revenue of \$120, a value of $R_t = (140, 0, 0)$ indicates a request for itinerary 1 with revenue of \$140, a value $R_t = (0, 70, 0)$ indicates a request for itinerary 2 with revenue of \$70, etc.

Note by our assumption that at most one arrival occurs in each time period, at most one component of R_t can be positive (as indicated in the example above). More formally, if we let $E_n = \{e_0, e_1, \ldots, e_n\}$ where e_j is the j-th unit n-vector and e_0 is the zero n-vector and define the set $\mathcal{S} = \{R : R = \alpha e, e \in E_n, \alpha \geq 0\}$, then $R_t \in \mathcal{S}$. The sequence $\{R_t;\ t \geq 1\}$ is assumed to be independent with known joint distributions $F_t(r)$ whose support is on $\mathcal{S}$. We further require that the marginal distributions, $F_t^j(r) = P(R_t^j \le r)$ be continuous on $(0, +\infty)$ and that all R_k^j have finite means. For the asymptotic results, we require that the revenues have bounded support (i.e. $P(R_t^j \le C) = 1$ for some constant C). Note that $F_t^j(r)$ is *not* the revenue distribution directly. Rather, $P(R_t^j > 0) = 1 - F_t^j(0)$ is the probability of getting a request from itinerary j in period t, and $P(R_t^j \le r \mid R_t^j > 0) = F_t^j(r)/(1 - F_t^j(0))$ is the distribution of the actual revenue from itinerary j at time t.

The time dependence of the revenue distribution models a variety of nonstationarities in the arrival process. For example, due to purchase timing restrictions, the mix of available fare products changes with time. Purchase patterns in various customer segments (e.g. the ratio of leisure/business purchases) change as the flight departure date approaches as well. No assumption is made on the particular order of arrival in our model.

Allowing uncertainty in revenues is important for several reasons. First, airlines often offer a variety of fares in each fare class for each itinerary and also pay varying commissions on these fares. Under such conditions, there is a potential to generate more revenue by discriminating among the various net revenues *within* a particular fare class, and the value function should reflect this potential. Second, the practice of negotiating fares in some industries (advertising, broadcasting, hotels) contributes to uncertainty in fares. Including fare variance provides more modeling flexibility in these emerging revenue management applications. Finally, modeling fare variance provides flexibility in constructing forecasts. Specifically, one can decrease the relative forecast error by aggregating fare classes, at the expense of increasing the variance in the fares within a fare class. This ability to vary the level of aggregation in the forecast data in this way has the potential of leading to a better overall forecasting-optimization scheme.

Given the time-to-go, k, the current seat inventory x and the current request R_k, we are faced with a decision: Do we or do we not accept the current request?

Let an n-vector u_k denote this decision, where $u_k^j = 1$ if we accept a request for itinerary j at time k, and $u_k^j = 0$ otherwise. In general, the decision to accept, u_k^j, is a function of the capacity vector x and the fare r^j offered for itinerary j, i.e. $u_k^j = u_k^j(x, r^j)$ and hence $u_k = u_k(x, r)$, where $r = (r^1, \ldots, r^n) \in \mathcal{S}$. Since we can accept at most one request in any period, $u_k \in E_n$, where E_n is the collection of unit n-vectors as defined above. Since we assume cancellations and no-shows do not occur and that legs cannot be oversold, if the current seat inventory is x, then u_k is restricted to the set $\mathcal{U}(x) = \{e \in E_n : Ae \le x\}$.

We can now define precisely what a bid-price control scheme is in the context of this model.

DEFINITION 1. *A control $u_k(x, r)$ is said to be a <u>bid-price control</u> if there exist real-valued functions $\mu_k(x) = (\mu_k^1(x), \ldots, \mu_k^m(x))$, $k = 1, 2, \cdots$ (called bid prices) such that*

$$u_k^j(x, r^j) = \begin{cases} 1 & r^j \ge \sum_{i \in A^j} \sum_{h=0}^{a_{ij}-1} \mu_k^i(x - h),\ A^j \le x, \\ 0 & \textit{otherwise}. \end{cases} \tag{1}$$

That is, a bid-price control specifies a set of bid prices for each leg at each point in time and for each capacity, such that we accept a request for a particular itinerary if and only if there is available capacity and the fare exceeds the sum of the bid prices for all the units of capacity used by the itinerary. We next examine whether this bid-price structure is optimal.

2. Structure of the Optimal Control

In order to formulate a dynamic program to determine optimal decisions $u_k^*(x, r)$, let $J_k(x)$ denote the maximum expected revenue (cost-to-go) for a given seat inventory x at time k. Then $J_k(x)$ must satisfy the Bellman equations (see Bertsekas (1995), p. 18)

$$\begin{aligned} J_k(x) = \max_{u_k(\cdot) \in \mathcal{U}(x)} & E[R_k u_k(x, R_k) \\ & + J_{k-1}(x - A u_k(x, R_k))] \end{aligned} \tag{2}$$

with the boundary condition

$$J_0(x) = 0, \quad \forall x. \tag{3}$$

This leads to our first proposition, which establishes the existence of an optimal policy and characterizes the form of the optimal control.

PROPOSITION 1. *If R_k^j has finite first moments for all k and j, then $J_k(x)$ is finite for all finite x, and an optimal control u_k^* exists of the form*

$$u_k^{j*}(x, r^j) = \begin{cases} 1 & r^j \geq J_{k-1}(x) - J_{k-1}(x - A^j) \text{ and } A^j \leq x, \\ 0 & \text{otherwise}. \end{cases} \tag{4}$$

PROOF. Note that $u_k^*(x, r)$ defined by (4) maximizes

$$ru + J_{k-1}(x - Au)$$

subject to the constraint $u \in \mathcal{U}(x)$. Therefore,

$$E[R_k u_k^*(x, R_k) + J_{k-1}(x - Au_k^*(x, R_k))] \geq \max_{u_k(\cdot) \in \mathcal{U}(x)} E[R_k u_k(x, R_k) + J_{k-1}(x - Au_k(x, R_k))].$$

Thus, u_k^* satisfies the Bellman equation provided we can show that the expectation on the left-hand side exists. To do so, we use induction. First, assume that $J_{k-1}(x)$ is finite for all finite x. Then applying u_k^* we have that

$$E[R_k u_k^*(x, R_k) + J_{k-1}(x - Au_k^*(x, R_k))] = J_{k-1}(x) + \sum_{j: A^j \leq x} E(R_k^j + J_{k-1}(x - A^j) - J_{k-1}(x))^+.$$

By the induction assumption, $J_{k-1}(x - A^j) + J_{k-1}(x)$ is finite. Therefore, since $E(R - c)^+ \leq ER + |c|$, the right-hand-side above is finite if the revenues R_k^j have finite first moments. The finiteness of J_{k-1} then follows using induction on k and the fact that $J_0(x) = 0$ for all x, and that under u_k^*,

$$J_k(x) = J_{k-1}(x) + \sum_{j: A^j \leq x} E(R_k^j + J_{k-1}(x - A^j) - J_{k-1}(x))^+. \quad \square \tag{5}$$

3. Nonoptimality of Bid-Price Controls

Proposition 1 says that an optimal policy for accepting requests is of the form: accept fare r^j for itinerary j if and only if we have sufficient remaining capacity and

$$r^j \geq J_{k-1}(x) - J_{k-1}(x - A^j).$$

This reflects the rather intuitive notion that we accept a fare of r for a given itinerary only when it exceeds the *opportunity cost* of the reduction in leg capacities. It is precisely this intuition and its analogy to the role of dual prices in deterministic optimization that motivated the early development of bid-price control schemes (Simpson 1989, Williamson 1992). However, in general this form of control is not a bid price control, a fact which we illustrate next via two counter examples.

3.1. A Counter Example to Bid-Price Optimality

In this first example, we have a simple network with two legs. There is one unit of capacity on each leg and two time periods remaining in the horizon. The itinerary data are shown in Table 1. In period 2, there are two local itineraries. (A *local* itinerary on a leg is a nonstop itinerary consisting of that leg, while a *through* itinerary on a leg is a multi-leg itinerary involving that leg.) each with a fare of \$250 and probability of arrival 0.3, and one through itinerary with a fare of \$500 and probability of arrival 0.4; in the last period, there is only a through fare with the same \$500 revenue and a probability of arrival of 0.8. Recall, that arrivals in each period are mutually exclusive (i.e., only one itinerary per period arrives).

In this example, we report the data in the form of an arrival probability and a fare for each itinerary. This can easily be translated into a single distribution of arriving revenues $R_t = (R_t^1, \ldots, R_t^n)$ where R_t^j is the revenue associated with a request for itinerary j at time k and $R_t^j = 0$ indicates no request for itinerary j occurred. Note also that the assumption of a deterministic fare violates the continuity assumption on the distribution of R_t. It is not hard to show, however, that we can get essentially the same counter example by replacing the deterministic fare with a random fare that has a continuous distribution arbitrarily close to the (degenerate) deterministic distribution.

It is not hard to see by inspection what an optimal policy is for this example. Accepting either of the local itineraries in period 2 yields \$250 in revenue and prevents us from accepting a through itinerary in period 1.

Table 1 Problem Data for Bid-Price Counter Example

Time (t)	Itin. (A^j)	Fare	Prob.
2	(1 1)	\$500	0.4
	(1 0)	\$250	0.3
	(0 1)	\$250	0.3
1	(1 1)	\$500	0.8
	No arrival		0.2

However, if we do not accept a local itinerary in period 2 and leave both legs available for the through demand in period 1, the expected revenue is \$400. So it is optimal to *reject* both local itineraries in period 2.

On the other hand, we clearly want to accept the through itinerary in period 2. Together, this implies that the bid prices, μ_1 and μ_2, in period 2 must satisfy $\mu_1 > 250$, $\mu_2 > 250$ and $\mu_1 + \mu_2 \le 500$, which is, of course, impossible. Therefore, no bid-price policy can produce an optimal decision in period 2. Indeed, it is not hard to show that the best a bid-price policy can do in this example is to reject *all* demand in period 2 and accept only the through fare (if it arrives) in period 1, yielding a \$400 expected revenue. The optimal policy, in contrast, generates an expected revenue of \$440—fully 10% more expected revenue than the best possible bid-price policy.

Finally, it is possible to construct other counter examples in which bid-price sub-optimality occurs with arbitrarily large remaining leg capacities. Thus, the problem of bid price sub-optimality is not confined to the end of the horizon, but can in fact occur at any point in the booking process.

3.2. Bid-Price Optimality and the Structure of the Value Function

A structural insight into why bid prices are not optimal in general is obtained by considering the implication of bid-price optimality for the value function $J_k(x)$. In some cases, it implies a certain linearity of the value function, as the next proposition demonstrates.

PROPOSITION 2. *Suppose the elements of A are only zero or one (i.e. no multiple requests) and that A has the identity matrix as a sub matrix. Further, suppose the marginal distributions $F_t^j(x)$ are strictly increasing on $(0, +\infty)$ for all t and j. Then a bid-price control scheme is optimal only if J_k satisfies*

$$J_k(x) - J_k(x - A^j) = \sum_{i \in A^j} (J_k(x) - J_k(x - e_i))$$

for all k, j and $x \ge A^j$.

PROOF. Without loss of generality, let the first m columns of A be the identity matrix. If a bid-price control scheme is optimal, then by considering the first m itineraries we must have that

$$\mu_k^i(x) = J_k(x) - J_k(x - e_i), \quad i = 1, \ldots, m,$$

by Proposition 1 and the definition of bid-price control. Now suppose for some k, j and $x \ge A^j$, that $J_k(x - A^j) > J_k(x) - \Sigma_{i \in A^j} (J_k(x) - J_k(x - e_i))$. Then by Proposition 1 the threshold for accepting fares for itinerary j in state x in period $k + 1$ is

$$J_k(x) - J_k(x - A^j) < \sum_{i \in A^j} (J_k(x) - J_k(x - e_i)) = \sum_{i \in A^j} \mu_k^i(x).$$

Using this inequality together with the fact that the distributions $F_k^j(x)$ are strictly increasing, one can show that the threshold as determined by the bid prices, $\Sigma_{i \in A^j} \mu_k^i(x)$, violates the optimality conditions (2); hence, bid prices cannot be optimal. A similar contradiction is obtained when $J_k(x - A^j) < J_k(x) - \Sigma_{i \in A^j} (J_k(x) - J_k(x - e_i))$. □

Proposition 2 shows that bid prices are only optimal in this case when the opportunity cost of the itinerary, $J_k(x) - J_k(x - A^j)$, is equal to the sum of the opportunity costs of selling each leg i separately, $\Sigma_{i \in A^j} (J_k(x) - J_k(x - e_i))$. This sort of linearity in the value function cannot be expected to hold in general.

Indeed, the counter example above illustrates the two main reasons why it does not hold. Bid prices in this example fail in part because selling a seat is a ‘‘large’’ change in the capacity of a leg. Large relative changes in capacity on several legs simultaneously cannot, in general, be expected to have the same revenue effect as the sum of the individual changes. This is one reason why gradient-based reasoning falls short in explaining bid-price optimality. Indeed, this issue was first raised by Curry (1992), who used an analogy to the Taylor series expansion of the value function to argue that second-order ‘‘interaction’’ terms may be significant in determining optimal revenue thresholds.

The second reason bid prices may fail to capture the opportunity cost is that future revenues may depended in a highly nonlinear way on the remaining capacity. Specifically, in the counter example note that it is the *minimum* capacity on the two legs that determines future expected revenues. Hence, the opportunity cost of using a single leg exactly equals the opportunity cost of using both legs simultaneously, so the linearity required in Proposition 2 is destroyed. This phenomenon is very similar to degeneracy in mathematical programming, and it can occur in the optimal value function or in various approximation to the optimal value function, as shown in §5.5.

4. An Asymptotic Analysis of Bid-Price Controls

We next analyze the degree of suboptimality of a bid-price control scheme. Our first step is to consider an upper bound based on a relaxation of the original problem. The upper bound provides dual prices that are then used in our second step to construct a bid-price policy. In contrast to the negative results of the previous section, we show that the control generated by this particular set of bid prices has good asymptotic properties if the number of seats sold on each leg is large.

4.1. An Upper Bound Problem

Let u_t represent a given control policy. We consider u_t somewhat more abstractly as simply a process which is adapted to history of requests from k to t. That is, if $\mathcal{F}_t = \sigma(\{R_k;\ T \geq k \geq t\})$, then u_t is a process that is $\mathcal{F}_t$-measurable. The problem can then be stated as finding such a process u_t that solves the problem

$$J_k(x) = \max E\left[\sum_{t=1}^{k} R_t u_t\right],$$

$$\sum_{t=1}^{k} A u_t \leq x \quad \text{(a.s.)},$$

$$u_t \in E_n. \tag{6}$$

Note that the process u_t is defined for a given starting state x and k, and therefore optimizing over u_t only provides a control policy for those states reachable from (x, k). Nevertheless, this formulation is sufficient for determining $J_k(x)$.

For any m-vector $\mu \geq 0$, consider a relaxed version of this problem

$$\bar{J}_k(x, \mu) = \max_{\{u_t \in E_n\}} E\left[\sum_{t=1}^{k} R_t u_t\right] + E\left[\mu\left(x - \sum_{t=1}^{k} A u_t\right)\right]$$

$$= \max_{\{u_t \in E_n\}} E\left[\sum_{t=1}^{k} (R_t - \mu A) u_t\right] + \mu x. \tag{7}$$

Here, u_t is a process which is adapted to $\mathcal{F}_t$ and satisfies $u_t \in E_n$, but need not satisfy $\sum_{t=1}^{k} A u_t \leq x$ (a.s.). That is, the policy might oversell a leg i, but at a cost of μ_i for each oversold seat.

The values of these two problems are related as follows:

LEMMA 1. *For any $\mu \geq 0$, $J_k(x) \leq \bar{J}_k(x, \mu)$.*

PROOF. For any control process $\{u_t : 1 \leq t \leq k\}$ which is an optimal policy for (6), we have that $\mu(x - \sum_{t=1}^{k} A u_t) \geq 0$ (a.s.), and hence because u_t is bounded, $E[\mu(x - \sum_{t=1}^{k} A u_t)] \geq 0$ as well. In addition, by definition $E[\sum_{t=1}^{k} R_t u_t] = J_k(x)$. Since such a policy is also feasible for (7), the above inequality follows. □

To create the best upper bound possible from this relation, we will minimize $\bar{J}_k(x, \mu)$ over $\mu \geq 0$. Fortunately, evaluating $\bar{J}_k(x, \mu)$ is not difficult, since it is easy to see that in (7) the problem decomposes by periods. Hence an optimal policy for (7) is simply

$$u_t^j = \begin{cases} 1 & r^j > \mu A^j, \\ 0 & \text{otherwise} \end{cases}$$

Note this is in fact a bid-price control with bid prices given by the vector μ for all times $t \leq k$ and all states x. Evaluating the cost under this optimal control yields

$$\bar{J}_k(x, \mu) = \sum_{t=1}^{k} \sum_{j=1}^{n} E(R_t^j - \mu A^j)^+ + \mu x. \tag{8}$$

The partial expectation $E(Z - z)^+$ is a convex function in z for any random variable Z for which the expectation exists. Therefore, $\bar{J}_k(x, \mu)$ is convex in μ. As a result the problem,

$$v_k(x) = \min_{\mu \geq 0} \bar{J}_k(x, \mu) \tag{9}$$

is a convex program. This minimization problem generates the least upper bound from our relaxation. Let μ^* denote an optimal solution to (9). Note that μ^* clearly depends on x and k; that is, $\mu^* = \mu^*(x, k)$. However, to economize on notation we do not include the arguments x and k.

Note that $\frac{d}{dz}E(Z - z)^+ = -P(Z > z)$. Therefore, if the distributions of R_t^j are continuous on $(0, +\infty)$, then an optimal solution μ^* to (9) satisfies the Kuhn-Tucker necessary conditions

$$\sum_{t=1}^{k} \sum_{j=1}^{n} P(R_t^j > \mu^* A^j) A^j - \lambda = x,$$

$$\lambda \mu^* = 0,$$

$$\lambda \geq 0. \tag{10}$$

Since (9) is a convex program, these conditions are also sufficient.

The Kuhn-Tucker conditions (10) are quite intuitive. Note that since $P(R_t^j > \mu^* A^j) = Eu_t^j$, the term $\sum_{t=1}^{k} \sum_{j=1}^{n} P(R_t^j > \mu^* A^j) A^j$ is the vector of expected number of requests for each leg over the remaining horizon k from itineraries whose revenue exceeds the bid prices defined by μ^*. Since $\lambda \geq 0$, the first and second condition in (10) imply that if $\mu^{*i} > 0$, then $\lambda^i = 0$ and the expected number of such request for leg i, across all itineraries, is precisely the capacity x^i; if $\mu^{*i} = 0$, $\lambda^i \geq 0$ and the expected number of such requests for leg i is no more than the capacity x^i. Indeed, (9) is equivalent to the problem of maximizing the expected revenue subject to the constraint that the *expected* number of requests is no more than x (i.e. the constraint $E[\sum_{t=1}^{k} Au_t] \leq x$).

The optimal value obtained by solving (9) also provides an analytical alternative to the ''perfect hindsight'' upper bound, which is used frequently in many practical simulation studies. The perfect hindsight bound is obtained by solving the linear program

$$V_k(x, \omega) = \max \sum_{t=1}^{k} R_t(\omega) u_t(\omega),$$

$$\sum_{t=1}^{k} A u_t(\omega) \leq x,$$

$$u_t(\omega) \in [0, 1]^n, \tag{11}$$

where we use ω to indicate that the optimization is performed using perfect information about the actual realization of demand. From strong duality, we then have

$$\begin{aligned} V_k(x, \omega) &= \min_{\mu(\omega) \geq 0} \max_{u_t(\omega) \in [0,1]^n} \sum_{t=1}^{k} R_t(\omega) u_t(\omega) \\ &\quad + \mu(\omega)(x - A u_t(\omega)) \\ &= \min_{\mu(\omega) \geq 0} \sum_{t=1}^{k} \sum_{j=1}^{n} (R_t^j(\omega) - \mu(\omega) A^j)^+ + \mu(\omega) x \\ &\leq \sum_{t=1}^{k} \sum_{j=1}^{n} (R_t^j(\omega) - \mu^* A^j)^+ + \mu^* x \end{aligned}$$

where again μ^* denotes an optimal solution to (9). Taking expectations on both sides above and using (8) and (9) yields

$$EV_k(x, \omega) \leq v_k(x).$$

Hence, as a bound on optimal expected revenues, (9) is weaker than what one obtains by simulating and averaging (11); however, because it is analytical, requiring only one optimization and no simulation, it is much more computationally efficient.

In summary, (9) provides an upper bound on optimal revenues, and its optimal solutions, μ^*, satisfy the constraints of the original problem in expectation. We next show that if one fixes the set of bid prices at μ^* for all times $t \leq k$, the resulting revenue is asymptotically optimal in a certain scaling of the problem. That is, a fixed bid price policy with bid prices equal to μ^* is in fact asymptotically optimal.

4.2. Asymptotic Analysis of Bid Prices Derived from the Upper Bound Problem

Below, μ^* denotes an optimal solution to (9). Again, we note that μ^* depends on the initial capacity x and the initial time-to-go k. Consider the following fixed-bid-price heuristic:

Fixed-Bid-Price Heuristic (H)

At time k with remaining capacity x, compute μ^* once by solving (9). *Then, for all times $t \leq k$, accept a request for itinerary j with revenue r if and only if there is sufficient capacity to satisfy it and $r > \mu^* A^j$.*

Let $J_k^H(x)$ denote the expected revenue of this heuristic given initial capacity x and time-to-go k. Let θ be a positive integer, and consider a sequence of problems, indexed by θ, with initial capacity vectors θx, time-to-go θk and revenues, denoted $R_t(\theta)$, where

$$R_t(\theta) =_D R_{\lceil t/\theta \rceil}, \tag{12}$$

and $=_D$ denotes equality in distribution. This construction corresponds to splitting each period t in the original problem into θ statistically independent and identical periods in the scaled problem and at the same time increasing the initial capacity by a factor of θ. As a result, the relative values of demand, capacity and time are preserved.

For the scaled problem, let $J_{\theta k}(\theta x)$ denote the optimal expected revenue and $J_{\theta k}^H(\theta x)$ denote the expected revenue of the fixed-bid-price heuristic. (We show below that the vector μ^* solves (9) for all θ, so the same vector of bid prices is used for each problem in the sequence.) The following result shows that the fixed-bid-price heuristic is asymptotically optimal as the scale of the problem, as measured by θ, increases:

THEOREM 1. *If $R_t^j \le C$ (a.s.), then*

$$\frac{J_{\theta k}^H(\theta x)}{J_{\theta k}(\theta x)} \ge 1 - O(\theta^{-1/2}).$$

In particular,

$$\lim_{\theta\to\infty} \frac{J_{\theta k}^H(\theta x)}{J_{\theta k}(\theta x)} = 1.$$

PROOF. By considering (8) and (12), we have

$$\sum_{t=1}^{\theta k} \sum_{j=1}^{n} E(R_t^j(\theta) - \mu A^j)^+ + \theta\mu x$$
$$= \theta\left[\sum_{t=1}^{k}\sum_{j=1}^{n} E(R_t^j - \mu A^j)^+ + \mu x\right] = \theta\bar{J}_k(x, \mu).$$

Therefore, the vector μ^* that solves (9) is the same for all value of θ. We therefore have by Lemma 1

$$J_{\theta k}(\theta x) \le \theta v_k(x). \tag{13}$$

Now consider the fixed-bid-price heuristic with μ^* as the fixed vector of bid prices. We will construct a sample path bound on the revenue with these bid prices using a coupling argument. To do so, we consider an alternate system which follows the bid-price policy for accepting sales, but has no capacity constraint; rather, in the alternate system we subtract a revenue of C for each set sold in excess of θx_i on each leg i, where C is the uniform upper bound on the itinerary revenues. Consider the net revenues collected in each system for a given sample path of arrivals. Note in the two systems, the same revenues are collected up until the time one or more of the leg capacities is exhausted.

Now, suppose a request arrives for an itinerary j that needs a leg i whose capacity is exhausted. If $R_t^j(\theta) \le \mu^* A^j$, it will be rejected in both systems and no revenues will be collected in either system. If $R_t^j(\theta) > \mu^* A^j$, the request will be rejected in the original system because of the capacity constraint. In the alternate system, it will be accepted but at least one penalty of C will be charged because one or more leg capacities are exhausted. No revenues will be collected in the original system, and the alternate system will realize a loss since $R_t^j(\theta) - C \le 0$. Moreover, the alternate system will have even less capacity remaining because it accepted the request. Hence, it follows that the net revenues in the alternate system are, pathwise, a lower bound on the revenues obtained under the bid-price heuristic. Therefore,

$$J_{\theta k}^H(\theta x) \ge \sum_{t=1}^{\theta k}\sum_{j=1}^{n} E(R_t^j(\theta) - \mu^* A^j)^+$$
$$+ \sum_{t=1}^{\theta k}\sum_{j=1}^{n} P(R_t^j(\theta) > \mu^* A^j)\mu^* A^j$$
$$- C\sum_{i=1}^{m} E(N^i - \theta x^i)^+, \tag{14}$$

where N^i is the number of leg i seats sold under the bid-price heuristic, which is given by

$$N^i = \sum_{t=1}^{\theta k}\sum_{j=1}^{n} \mathbf{1}\{R_t^j(\theta) > \mu^* A^j\} a_{ij}.$$

The first two terms in (14) are the actual revenues collected; the last term is the total penalties charged.

Let $Y_{ijt} = \mathbf{1}\{R_t^j > \mu^* A^j\} a_{ij}$ and note that by (12) and the independence of the vectors $R_t(\theta)$, that $EN^i = \theta \sum_{t=1}^{k}\sum_{j=1}^{n} EY_{ijt}$ and that $\mathrm{Var}(N^i) = \theta \sum_{t=1}^{k}\sum_{j=1}^{n} \mathrm{Var}(Y_{ijt})$. We now use a bound due to Gallego (1992), which states that for any random variable Z with mean μ and finite variance σ^2,

$$E(Z - z)^+ \le \frac{\sqrt{\sigma^2 + (z-\mu)^2} - (z - \mu)}{2}.$$

Applying this bound to the terms in the last sum in (14) and using the fact that

$$EN^i = \sum_{t=1}^{\theta k}\sum_{j=1}^{n} P(R_t^j > \mu^* A^j) a_{ij} \le \theta x^i$$

by the Kuhn Tucker conditions (10) for μ^*, implies that

$$E(N^i - \theta x^i)^+ \leq \frac{\sqrt{\mathrm{Var}(N^i) + (\theta x^i - EN^i)^2} - (\theta x^i - EN^i)}{2}$$

$$\leq \frac{\sqrt{\mathrm{Var}(N^i)} + |\theta x^i - EN^i| - (\theta x^i - EN^i)}{2}$$

$$= \frac{\sqrt{\mathrm{Var}(N^i)}}{2}. \tag{15}$$

Also note that by the Kuhn Tucker conditions (10), $\mu^*(\sum_{t=1}^{\theta k} \sum_{j=1}^{n} P(R_t^j(\theta) > \mu^* A^j)A^j - \theta x) = 0$, so that

$$\sum_{t=1}^{\theta k} \sum_{j=1}^{n} P(R_t^j(\theta) > \mu^* A^j)\mu^* A^j = \theta \mu^* x. \tag{16}$$

Using (15) and (16) in the second and third sums in (14) and using (13), we obtain

$$J_{\theta k}^H(\theta x) \geq \sum_{t=1}^{\theta k} \sum_{j=1}^{n} E(R_t^j(\theta) - \mu^* A^j)^+$$

$$+ \theta \mu^* x - \frac{C}{2}\sqrt{\theta} \sum_{i=1}^{m} \sqrt{\sum_{t=1}^{k} \sum_{j=1}^{n} \mathrm{Var}(Y_{ijt})}$$

$$= \theta v_k(x) - O(\sqrt{\theta}),$$

which completes the proof. □

4.3. Uniqueness of the Asymptotic Bid Prices

We next address the uniqueness of the asymptotically optimal bid prices. Although the upper bound problem (9) is convex in μ, in general it is not strictly convex and hence it may not have a unique solution.

To see this, we can write the function $\bar{J}_k(x, \mu)$ as

$$\bar{J}_k(x, \mu) = g(\mu A) + \mu x. \tag{17}$$

where $g : R^n \to R$ is defined as

$$g(r) = \sum_{t=1}^{k} \sum_{j=1}^{n} E(R_t^j - r^j)^+, \tag{18}$$

and $r = (r^1, \ldots, r^n)$. Now if there exists a t such that $P(R_t^j > r^j) > 0$, $\forall r^j \geq 0$, then g is strictly convex on R^n.

However, even if g is strictly convex, in general $g(\mu A)$ is only (weakly) convex in μ. It is not hard to see that the function $g(\mu A)$ will be strictly convex in μ if and only if $xA = yA$ implies $x = y$, which is equivalent to the condition: $xA = 0$ implies $x = 0$. But this is true if and only if A has rank m. Therefore, we have the following sufficient condition for uniqueness:

PROPOSITION 3. *Suppose for all j there exists a t such that $P(R_t^j > r) > 0$ on $[0, +\infty)$. Further, suppose* rank(A) $= m$. *Then the solution to* (9) (*and hence the vector of asymptotically optimal bid prices*) *is unique.*

In most practical settings, we would expect $n \gg m$ and hence it is highly likely that A will have rank m. In this case, sufficiently large tails on the fare distributions will result in unique asymptotically optimal bid prices.

However, multiple asymptotically optimal bid-price can occur if fare distributions are highly concentrated. For example, one can easily construct situations in which there is a range of bid prices that are high enough to block a low fare class while still being low enough to allow higher fare classes to book; each value produces the same acceptance decision (with probability one) and hence all are asymptotically optimal bid prices. Alternatively, it is possible that because rank$(A) < m$ multiple solutions to (9) exist. As a simple example, consider the case of two legs in series ($m = 2$) with one itinerary ($n = 1$) that traverses both legs. In this case rank$(A) = 1 < m$, and multiple solutions exist. Specifically, it is easy to see in this case $\bar{J}_k(x, \mu)$ depends on μ only through the sum $\mu^1 + \mu^2$.

4.4. Bid Prices and Opportunity Cost

The above observations illustrate an important point; namely, there is not a one-to-one correspondence between optimal bid prices and the opportunity cost of leg capacity. That is, one can generate examples of bid prices that give near optimal accept/deny decisions but at the same time are very poor approximations to the marginal value of leg capacity.

As a simple example of this difference, consider a single-leg problem in which high revenue fare classes arrive strictly *before* low revenue fare classes. In this case, it is clear that it is optimal to accept arrivals in first-come-first-serve (FCFS) order. Therefore, a constant bid price of zero is optimal. On the other hand, the opportunity cost $J_k(x) - J_k(x - 1)$ at each point in time k is certainly not zero. In other words, while it is *sufficient* to compare the revenue to the true opportunity cost

$J_k(x) - J_k(x-1)$ at each point in time to make optimal accept/deny decisions, it is not *necessary* to do so; other threshold values may produce the same accept/deny decision and same optimal revenues, as the value of zero does in this case.

One might argue that the real goal is to make the right accept/deny decision and therefore it is not worth worrying about the difference between optimal bid-price values and opportunity costs; however, in practice a good estimate of opportunity cost is often essential. In particular, for special event requests—especially ad-hoc group bookings—which are typically not part of the forecast, one needs an accurate assessment of opportunity cost to make a good decision.

The point, simply, is that one has to be careful about the interpretation of the bid prices produced by any optimization algorithm. Ideally, we would like the sum of leg bid prices along an itinerary to represent the itinerary's opportunity cost. On the other hand, due to "degeneracy" of the value function, this may not always be achievable. Yet despite this difficulty, Theorem 1 shows that a properly constructed bid-price control rule is still asymptotically optimal against forecasted demand. The algorithmic challenge, therefore, is to construct bid prices which produce near-optimal acceptance decisions against forecasted demand, while simultaneously providing good estimates of the opportunity cost (whenever possible), so that special event (group) requests can be properly evaluated.

5. A Unified View of Bid Price Approximation Schemes

It is quite helpful, to view bid price methods as corresponding to various approximations of the optimal value function. That is, a given approximation method A yields a function $J_k^A(x)$ that approximates $J_k(x)$. The bid prices are then the gradients of $J_k^A(x)$, i.e.

$$J_k(z) - J_k(x - A^j) \approx \nabla_x J_k^A(x) A^j,$$

If the gradient does not exist, then $\nabla_x J_k^A(x)$ above is typically replaced (at least implicitly) by a subgradient of $J_k^A(x)$. This interpretation of approximations schemes raises two important questions: Is $J_k^A(x)$ a good approximation of the value function? And more importantly, is $\nabla_x J_k^A(x) A^j$ a good approximation of the opportunity cost? In this section, we examine these questions for two popular approximation schemes and also the asymptotic bid prices developed in Theorem 1.

Before proceeding, we note that in actual applications, the approximation $J_k^A(x)$ is usually resolved frequently to allow the vector of bid prices to adjust to changes in remaining capacity x and remaining time k. In this section, we assume the approximations are solved for each x and k and compare the resulting bid prices to optimal bid prices.

5.1. Deterministic Linear Program (DLP)

The deterministic linear programming method corresponds to the approximation

$$J_k^{\mathrm{LP}}(x) = \min \sum_{j=1}^{n} ER_j y_j,$$
$$Ay \leq x,$$
$$0 \leq y \leq ED,$$

where $D = (D_1, \ldots, D_n)$ and D_j denotes the demand to come for itinerary j (ED is the expected value of D) and R_j is the (possibly random) revenue associated with itinerary j. In our earlier notation, $D_j = \Sigma_{t=1}^{k} \mathbf{1}\{R_t^j > 0\}$. The decision variables y_j represent a discrete (nonnested), static allocation of capacity to each itinerary j.

If the constraints $Ay \leq x$ are not degenerate (linearly dependent) at the optimal solution, then $\nabla J_k^{\mathrm{LP}}(x)$ exists and is given by the unique vector of optimal dual prices associated with these constraints; if these constraints are degenerate, then there are multiple optimal dual price vectors, each of which is only a subgradient of the function $J_k^{\mathrm{LP}}(x)$.

The most serious weakness of the DLP formulation is that it considers only the mean demand and ignores all other distributional information. As a consequence, the dual values are zero on any leg that has a mean demand less than capacity. Despite this deficiency, Williamson's (1992, Chapter 6) extensive simulation studies showed that with frequent reoptimization, the performance of DLP bid prices is quite good, producing higher revenue than both probabilistic math programming models (see below) and a variety of leg-based EMSR heuristics.

5.2. Probabilistic Nonlinear Program (PNLP)

The probabilistic nonlinear programming method corresponds to the approximation

$$J_k^{\text{PNLP}}(x) = \min \sum_{j=1}^{n} ER_j E \min\{D_j, y_j\},$$

$$Ay \le x,$$

$$y \ge 0,$$

where again D_j and R_j are defined as in the DLP case. As in the DLP, the decision variables y_j represent a discrete, static allocation of capacity to each itinerary j. If the constraints $Ay \le x$ are not degenerate at the optimal solution, then $\nabla J_k^{\text{LP}}(x)$ exists and is given by the unique vector of optimal dual prices associated with these constraints; if these constraints are degenerate, the multiple optimal dual vectors are subgradients of the function $J_k^{\text{LP}}(x)$.

This formulation appears somewhat better than the DLP, in that the term $E \min\{D_j, y_j\}$ in the objective function captures the randomness in demand. However, the assumption of a discrete, static allocations of capacity to each fare class can lead to poor behavior. This behavior was demonstrated empirically in Williamson's (1992, Chapter 6) simulation studies, in which she observed that the PNLP bid prices consistently produced lower revenues than the DLP bid prices. In a further computational comparison between the DLP and PNLP, Talluri (1996) found similar behavior.

To understand the weakness of the PNLP approximation, consider a problem with $m = 1$ leg and n itineraries on the leg, each of which is identical. Assume each itinerary j has demand, $D_j \sim D$, where D is normally distributed with mean μ and standard deviation σ and that each itinerary has the same deterministic revenue r. The PNLP formulation is then

$$J_k^{\text{PNLP}}(x) = \min \sum_{j=1}^{n} rE \min\{D, y_j\},$$

$$\sum_{j=1}^{n} y_j \le x,$$

$$y \ge 0,$$

where $D \sim N(\mu, \sigma)$. By symmetry, the optimal solution is $y_j = x/n$, $j = 1, \ldots, n$ and hence the Kuhn-Tucker conditions imply the optimal dual price λ satisfies

$$\lambda = rP(nD > x) = r\left(1 - \Phi\left(\frac{x - n\mu}{n\sigma}\right)\right), \quad (19)$$

where $\Phi(x)$ is the CDF of the standard normal distribution. The value λ above then forms our estimate of the marginal value of the xth seat.

But since all itineraries are identical, the marginal value in this problem should be the unchanged if we aggregate all n fare classes into one fare class with mean $n\mu$ and variance $n\sigma^2$. Aggregating and applying the PNLP we find the optimal dual multiplier in this case satisfies

$$\lambda = rP(\sum_{j=1}^{n} D_j > x) = r\left(1 - \Phi\left(\frac{x - n\mu}{\sqrt{n}\sigma}\right)\right). \quad (20)$$

If $x \neq n\mu$ and n is large, (19) and (20) give very different estimates of the marginal value. Of course, from first principles we know the true opportunity cost is completely independent of how we aggregate (or disaggregate) these identical itineraries.

While on the surface this seems like a contrived example, it is not unreasonable to expect a similar type of behavior in large hub-and-spoke networks. For example, if many uncongested in-bound legs have connecting passengers traveling on a single congested outbound leg and passengers pay comparable revenues, then the situation is quite similar to the example above. That is, one would like to treat all passengers as a single fare class (i.e. "nest" the fare classes) but the PNLP allocates space to each separately, resulting in a distorted estimate of the marginal value of the leg. In contrast, the DLP method does indeed posses this "nesting" property, since, in the above example, aggregating all n fare classes does not change the resulting DLP bid price.

5.3. Prorated EMSR

Another method for computing estimates of bid prices is to use a prorated expected marginal seat revenue (PEMSR) scheme. Originally proposed in Williamson (1992), PEMSR schemes involve allocating a portion of the revenue of each itinerary to the legs of the itinerary. One then solves m leg-level problems using the expected marginal revenue (EMSR) heuristic proposed by Belobaba (1989). The resulting EMSR values from each leg are then used as bid prices.

Specifically, let $\alpha = (\alpha_1, \ldots, \alpha_m)$ be a non-negative real vector. For each itinerary j, define new revenues, one for each leg in the itinerary, by

Table 2 Problem Data for Iterative Allocation Example

Time (k)	Itin. (A_j)	Fare	Prob.
2	AB	\$100	0.5
	CD	\$100	0.5
1	ABC	\$1000	0.5
	BCD	\$1000	0.5

$$R_{ij} = \frac{\alpha_i}{\Sigma_{i \in A_j} \alpha_i} R_j, \quad i \in A_j.$$

Next, treat each leg i independently as if it received demand D_j, but with reduced revenue R_{ij} and solve the corresponding leg-level EMSR.[1] The approximation to the value function is then

$$J_k^{\text{PEMSR}}(x) = \sum_{i=1}^{m} J_i(x_i, \alpha),$$

where $J_i(x_i, \alpha)$ denotes the expected revenue of leg i under the allocation α.

Williamson (1992) investigated several methods for determining the allocation α, including prorating based on mileage, number of legs and the relative revenue value of local demand on each leg. Her conclusion is that none of these fixed allocations is very robust in general. Indeed, it is not hard to see that if one leg of an itinerary is highly congested and all others have abundant capacity, then the revenue of the itinerary should be entirely allocated to the congested leg. Depending on the realization of demand, however, the congested leg could be any of the legs on the itinerary; hence, no fixed allocation scheme can be expected to work well in all cases.

An intriguing idea along these lines, again appearing in Williamson (1992, p. 107) but not pursued fully there, is to prorate revenues using an iterative loop. That is, first obtain the marginal values based on some initial proration scheme and EMSR calculations. Then, use these marginal values to do the next round of proration and EMSR calculations. Repeat these iterations until the marginal values (hopefully) converge. The hope here is that the bid-prices will converge to a near-optimal set of bid-prices. Unfortunately there is little theoretical justification behind this idea.

Even if the bid prices converge to some value (it is not entirely clear if convergence is guaranteed), they do not necessarily converge to a good set of bid-prices. For example, consider a three-leg line network, with nodes A, B, C and D. Each of the three legs, AB, BC and CD, has one remaining seat. Suppose $t = 2$ and we have data for itinerary arrivals as shown in Table 2. If we start with an allocation of fares for $t = 1$ using equal weights, prorate the fares in period $t = 1$ by these weights, and then compute the expected marginal value of each leg we get $\mu_{\text{AB}} = 250$, $\mu_{\text{BC}} = 500$ and $\mu_{\text{CD}} = 250$.

The results of repeated applications of this procedure are shown in Table 3. Note that the bid prices converge to $\mu_{\text{AB}} = 0$, $\mu_{\text{BC}} = 1000$ and $\mu_{\text{CD}} = 0$. However, by inspection of the data in Table 2, it is clear that we want to reject both of the itineraries arriving in period $t = 2$, so we need $\mu_{\text{AB}} > 100$ and $\mu_{\text{CD}} > 100$. Such a policy yields an expected revenue of \$1,000. Because the iterative proration scheme produces zero bid prices for legs AB and CD, it accepts both of the itineraries in period $t = 2$, generating an expected revenue of only \$600.

One problem with the iterative proration scheme is that once the problem is broken up into leg-level problems, all network information is lost. Additional problems can arise due to the fact that many EMSR methods assume fare classes have ordered arrivals, usually with lower fare classes arriving before high fare classes, see Belobaba (1987, 1989), Brumelle et al. (1990), and Curry

Table 3 Example of Convergence of Iterative Proration Scheme ($t = 1$)

Iter.	μ_{AB}	μ_{BC}	μ_{CD}
0	250.0	500.0	250.0
1	166.7	666.7	166.7
2	100	800	100
.	.	.	.
.	.	.	.
.	.	.	.
∞	0	1000	0

[1] As formulated in our model, there is no particular order of arrival. Typically, some assumption on the arrival order, like lower fare fare classes book before higher, is made as an approximation and then one applies an EMSR scheme that does not require monotonically increasing fares, e.g., Robinson (1991).

(1989). In a proration scheme, the regular ordering of arrivals is often destroyed, and certainly any assumption of low prorated revenues arriving strictly before high prorated revenues becomes untenable.

5.4. Asymptotic Bid Prices

The asymptotic analysis provides an alternative approximation approach. Indeed, note from (9) that $\nabla_x v_k(x) = \mu^*$, so we can view the upper bound $v_k(x)$ as an approximation of $J_k(x)$ with μ^* its (sub)gradient. The approximation (9) is somewhat unique in that it is solved directly in the space of the bid prices μ, whereas in the DLP and NLP methods, μ is a dual value of a problem whose primal variables are inventory allocations.

The approximation (9) has the ''nesting'' property because the objective function in (9) sums all arriving itineraries j whose revenue exceeds the fixed thresholds μA^j. As a result, it does not suffer from the discrete allocation problem of the PNLP method. For example, suppose two itineraries j_1 and j_2 are entered as separate columns of A, but in reality $A^{j_1} = A^{j_2}$ and each has the same fare distribution. If these two itineraries were combined into a new itinerary (by adding the probabilities of arrival in each period together) the asymptotic bid prices would not change, which is the correct behavior.

At the same time, the approximation (9) suffers from the same weakness as the DLP in that its value only depends on the first moment of demand. Indeed, if, as above, we let D_j denote the demand to come for itinerary j and R_j denote the random revenue associated with itinerary j, then (9) becomes

$$v_k(x) = \min_{\mu \geq 0} \sum_{j=1}^{n} ED_j E(R_j - \mu A_j)^+ + \mu^T x,$$

which only depends on ED_j.

The problem here is that the asymptotic analysis is too ''coarse'' to capture some of the second-order stochastic effects which the PNLP captures. As a result, as in the DLP case, the asymptotic bid-prices will turn out to be zero if the mean demand on a leg is strictly less than its capacity. This is indeed correct behavior asymptotically, but for demands with high variance near capacity, the actual bid price could be significantly larger than zero, possibly even more than some of the fares of the lower fare classes. Indeed, under appropriate scaling, one can show that, as fare variances tend to zero, the asymptotic bid prices from (9) approach the DLP bid prices. One can then combine this result with Theorem 1 to show that as fare variances tend to zero, the DLP bid prices are also asymptotically optimal.

However, the asymptotic approximation, unlike the DLP and PNLP methods, accounts for variability in itinerary revenues, which provides a significant advantage in approximating optimal bid prices when fares vary significantly within itinerary/fare-classes.

As a simple example of this effect, consider a single-leg problem with only one fare class. Suppose the actual fares R vary and have distribution $F(r)$ with mean ER. If the mean demand, ED, is significantly higher than capacity, x, the bid price produced by DLP and PNLP methods both approach ER. That is, reducing the leg capacity by one almost certainly results in a lost sale, which each of these models values at ER. Under an optimal bid-price policy, however, the bid price rises above ER as the demand/capacity ratio increases. This occurs because, with many requests to choose from, it is optimal to be selective and accept only the higher fares within the fare class (i.e. fares in the ''right tail'' of the distribution) rather than accepting all fares. From (9), one can show that the optimal bid price in the above example tends to a value r^* satisfying

$$ED(1 - F(r^*)) = x.$$

Depending on the distribution of F, the value of r^* can be significantly higher than ER; hence the DLP and PNLP methods will under-estimates the optimal bid price. A complimentary effect can occur when the demand/capacity ratio is small; in this case, the DLP and NLP models may tend to over-estimate the optimal bid price.

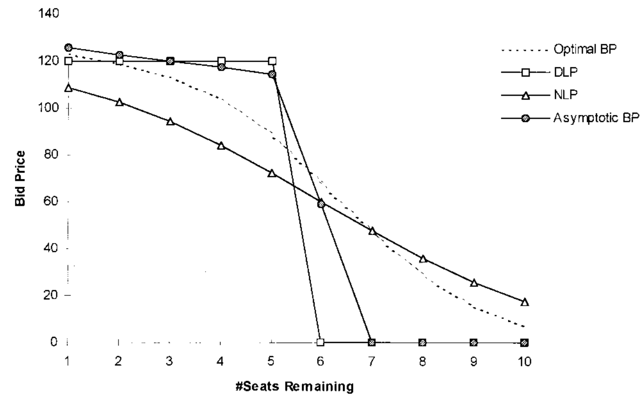
5.5. Numerical Examples

In this section, we illustrate the above qualitative behavior by computing bid prices for the DLP, PNLP and asymptotic approximations for a simple, 3-fare-class, single-leg problem. (The PEMSR method is not included because it reduces to ordinary EMSR in this case.) These bid prices are compared to the optimal bid prices obtained by dynamic programming. The examples are constructed primarily to illustrate the behavior

Table 4 Data for Numerical Examples (in order of arrival)

	Ex. 1	Ex. 2	Ex. 3
Class 3			
Fare Mean	80	120	120
Fare Std. Dev.	4	6	36
Class 2			
Fare Mean	120	120	120
Fare Std. Dev.	6	6	36
Class 1			
Fare Mean	400	120	120
Fare Std. Dev.	15	6	36

Figure 2 Bid Prices for Example 2



of the bid prices produced by each method rather than to mimic real data.

The problem has 30 time periods. The three fare classes arrive sequentially, with fare class 3 arriving in periods 1–10; fare class 2 arriving in periods 11–20; and fare class 1 arriving in periods 21–30. (Strict low-before-high fare order.) The probability of arrival in each period is 0.2 for all three fare class, resulting in a mean demand of 2 with a standard deviation of 1.26 for all three classes. Fares in each class are normally distributed. The three examples were generated by varying the fare means and standard deviations as shown in Table 4.

Figure 1 show the bid prices produced by each approximation method for Example 1. The values are those for the beginning of the horizon ($k = 30$ time units remaining) and are graphed as a function of the remaining seat capacity. The dashed line shows the optimal bid price. Note that the all three approximation methods underestimate the optimal bid price when the remaining capacity is less than the mean total demand of 6. When remaining capacity is greater than the mean total demand of 6, the PNLP overestimates the optimal bid price, while the DLP and asymptotic approximation produce a bid price of zero as noted above.

Figure 1 Bid Prices for Example 1

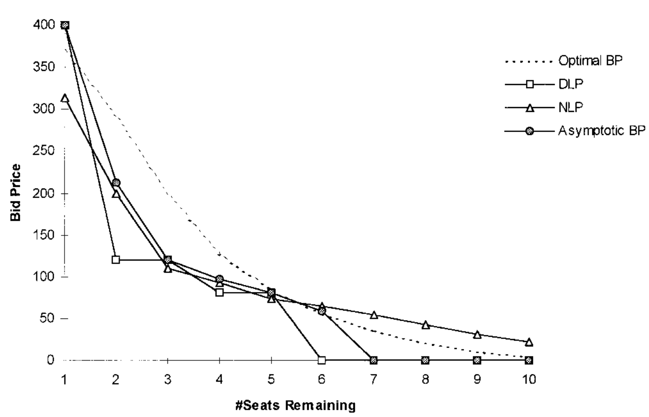
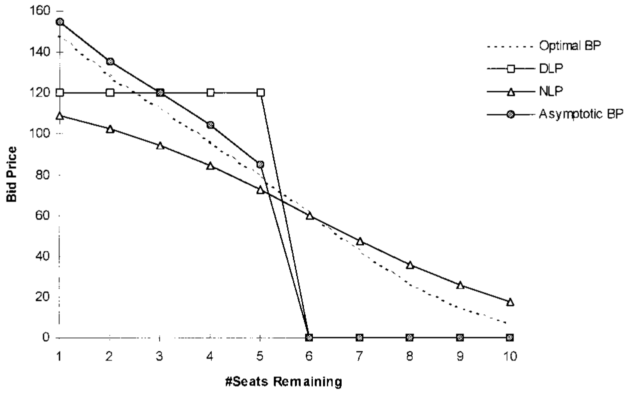


Figure 2 shows the bid prices for Example 2, again with 30 time units remaining. Since all fares are equal in this example, it highlights the nesting properties of the approximation methods. Note in this case the DLP and the asymptotic methods produce very similar values, with both methods overestimating the bid price at low remaining capacities and underestimating bid prices when capacity exceeds the mean total demand of 6. The PNLP method has the opposite behavior, underestimating the bid price at low capacity and overestimating it at high capacity. We note that if the three fare classes were aggregated (e.g. they were treated as one fare class arriving uniformly over period 1 to 30 with probability 0.2 in each period), then the bid price produced by PNLP would in fact be optimal, while the bid prices produced by the DLP and asymptotic method would be unchanged.

Figure 3 shows the effect of fare variance. The data for this example are the same as in Example 2, except that the fare variance has been increased in each fare class. Note this change has no effect on the DLP and PNLP bid prices as expected. However, the optimal and asymptotic bid prices change significantly at low remaining capacities, both approaching approximately

Figure 3 Bid Prices for Example 3



$150 as capacity tends to zero, rather than $120 as in Example 2. This reflects the optimality of accepting higher-than-average fares as capacity becomes highly constrained. Note that the asymptotic approximation provides a significantly better estimate of marginal value at low capacities in this case.

6. Conclusions

Bid prices are an appealing practical method for network revenue management. Our analysis confirms that, though not optimal, bid-price controls are provably near optimal in certain cases, provided the right bid prices are used. This result should be reassuring to the many users in the airline, hotel and broadcasting industries who are making the transition to bid-price technology. Our analysis also identifies cases where bid-price controls can be suboptimal and it sheds new light on the reasons why they can be suboptimal, in particular when there is a type of degeneracy in the value function. There are interesting questions concerning precisely how to detect such degeneracies and correct for them.

While asymptotic analysis is arguably a crude form of analysis, we believe that a good test of any algorithm for calculating bid-prices is that it have good asymptotic properties. That is, having good performance asymptotically ensures that the bid prices are capturing the correct ''first-order'' revenues. In this sense, among the two (DLP and PNLP), the DLP method seems to have best asymptotic properties. Asymptotic performance may explain why, for all its apparent simplicity, the LP bid-prices seem to generate more revenue than PNLP bid-prices in simulation experiments (see Talluri 1996 and Williamson 1992). Of course, by incorporating demand variance properly, it may be possible to obtain significantly more revenue than the LP approach.

Finally, we suggest that, in future research, bid price schemes be viewed as approximations of the value function. From this unified point of view, one can better understand the various strengths and weaknesses of each scheme and begin to make useful comparisons among various approaches. At the same time, more work is needed to understand the performance of approximation schemes on realistic data sets.[1]

[1] We thank Renwick Curry, Guillermo Gallego, and Rick Stone for helpful feedback on an earlier draft of this paper. We also thank the associate editor and two anonymous referees for their helpful feedback and suggestions.

References

Barnhart, C., K. T. Talluri 1997. Airline operations research. *Systems for Civil and Environmental Engineering*: *An Advanced Text Book* Charles ReVelle and Arthur McGarrity, eds. John Wiley and Sons, New York, 435–470.

Belobaba, P. P. 1987a. Air travel demand and airline seat inventory management. Unpublished Ph.D. Dissertation. MIT, Cambridge, MA.

—— 1987b. Airline Yield management: An overview of seat inventory control. *Transportation Sci*. **21** 63–73.

—— 1989. Application of a probabilistic decision model to airline seat inventory control. *Oper*. *Res*., **37** 183–197.

Bertsekas, D. P. 1995. *Dynamic Programming and Optimal Control*, *Volume I*, Athena Scientific, Belmont, MA.

Brumelle, S. L., J. I. McGill, T. H. Oum, K. Sawaki, M. W. Trethaway 1990. Allocation of airline seats between stochastically dependent demand. *Transportation Sci*., **24** 183–192.

——, J. I. McGill 1993. Airline seat allocation with multiple nested fare classes. *Oper*. *Res*., **41** 127–137.

Chatwin, R. E. 1992. Optimal airline overbooking. Ph.D. Thesis, Stanford University, Palo Alto, CA.

Curry, R. E. 1989. Optimal airline seat allocation with fare classes nested by origins and destinations. *Trans*. *Sci*. **24** 193–204.

—— 1992. Real-time revenue management: Bid price strategies for origins/destinations and legs. *Scorecard*, 2Q 1992, Aeronomics Inc., Atlanta, GA.

Diamond, M., R. Stone 1991. Dynamic yield management on a single-leg. Unpublished Manuscript, Northwest Airlines.

Dror, M., P. Trudeau, S. P. Ladany 1988. Network models for seat allocation of flights. *Tran*. *Res*. **22B** 239–250.

Feng, Y., Guillermo Gallego 1995. Optimal starting times for end-of-season sales and optimal stopping times for promotional fares. *Management Sci*. **41** 1371–1391.

Gallego, G. 1992. A minmax distribution free procedure for the (Q, R) inventory model. *Oper. Res. Letters* **11** 55–60.

——, G. van Ryzin 1994. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Sci*. **40** 999–1020.

——, —— 1997. A multi-product, multi-resource pricing problem and its applications to network yield management. *Oper. Res* **45** 24–41.

Glover, F., R. Glover, J. Lorenzo, C. McMillan 1982. The passenger-mix problem in the scheduled airlines. *Interfaces* **12** 73–79.

Janakiram, S., S. S. Stidham, A. Shaykevich 1994. Airline yield management with overbooking cancellations and no-shows. Technical Report No. UNC/OR/TR 94-9, Department of Operations Research, University of North Carolina, Chapel Hill, NC.

Lee, T. C., M. Hersh 1993. A model for dynamic airline seat inventory control with multiple seat bookings. *Trans. Sci*. **27** 252–265.

Littlewood, K. 1972. Forecasting and control of passengers. *12th AGIFORS Symposium Proceedings*. 95–128.

Luenberger, D. G. 1969. *Optimization by Vector Space Methods*. John Wiley & Sons, New York.

McGill, J. I. 1989. Optimization and estimation problems in airline yield management. Unpublished Ph.D. Dissertation, University of British Columbia, Vancouver, B.C.

Robinson, L. W. 1991. Optimal and approximate control policies for airline booking with sequential nonmonotonic fare classes. *Oper. Res*. **43** 252–263.

Rothstein, M. 1971. An airline overbooking model. *Trans. Sci*. **5** 180–192.

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.

Smith, B., J. Leimkuhler, R. Darrow, J. Samuels 1992. Yield management at American Airlines. *Interface* **22** 8–31.

Soumis, F., A. Nagurney 1993. A stochastic, multiclass airline network equilibrium model. *Oper. Res*. **41** 721–730.

Talluri, K. T. 1996. Revenue management on a line network. Working Paper, UPF, Barcelona, Spain.

——, G. J. van Ryzin 1996. A new provably optimal bid-price technology. *Proceedings of the AGIFORS Reservations and Yield Management Conference*, Zurich, Switzerland.

Wang, K. 1983. Optimum seat allocation for multi-leg flights with multiple fare types. *Proceedings of the AGIFORS 23rd Annual Symposium*, 225–246.

Weatherford, L. R., S. E. Bodily 1992. A taxonomy and research overview of perishable-asset revenue management: Yield management, overbooking and pricing. *Oper. Res*. **40** 831–844.

——, ——, P. Pfeifer 1993. Modeling the customer arrival process and comparing decision rules in perishable asset revenue management situations. *Trans. Sci*. **27** 239–251.

Williamson, E. L. 1992. Airline network seat control. Ph.D Thesis, MIT, Cambridge, MA.

Wollmer, R. D. 1986. A hub-spoke seat management model. Unpublished internal report, McDonnell Douglas Corporation, Long Beach, CA.

—— 1986. An airline reservation model for opening and closing fare classes. Unpublished internal report, McDonnell Douglas Corporation, Long Beach, CA.

—— 1992. An airline seat management model for a single leg route when lower fare classes book first. *Oper. Res*. **40** 26–37.

Young, Y., R. van Slyke 1994. Stochastic knapsack models of yield management. Technical Report 94–76, Polytechnic University.

Accepted by Linda V. Green; received September 1996. *This paper was with the authors for* 5 *months for* 2 *revisions*.