

Ramarao Desiraju & Steven M. Shugan

Strategic Service Pricing and Yield Management

The authors investigate pricing strategies based on yield management systems (YMS), such as early discounting, overbooking, and limiting early sales, for capacity-constrained services. They find that YMS work best when price-insensitive customers prefer to buy later than price-sensitive consumers. The authors also identify other conditions favoring the use of YMS.

Service sector¹ growth (e.g., Shugan 1994) and its study suggest that service providers frequently face special problems (e.g., Johnson 1964; Radas and Shugan 1998a, b; Rust, Zahorik, and Keiningham 1995). One such problem is perishability (Parasuraman and Varadarajan 1988). When planes, trains, or ships depart, for example, unused seats are lost forever. Similar problems occur for vacant hotel rooms, unsold concert tickets, idle tables at restaurants, and many other fixed-capacity services.

Linked to the perishability problem is a capacity-constraint problem. When demand peaks, many services face binding capacity constraints that prevent serving additional customers. For example, car rental agencies have a limited number of cars at each location, hotels have a limited number of rooms, and repair shops have a limited number of mechanics. Indeed, nearly all service providers face some form of capacity constraint.²

The combination of perishability and capacity constraints can encourage a business orientation in which service providers focus on filling capacity (Wardell 1989). For example, Amtrak states, "We don't want empty seats" (Wilder 1991). However, as Heuslein (1993, p. 178) reminds us, "Filling up empty seats with discount fares doesn't help if packed planes are still flying at a loss."

To profitably fill capacity, many service providers use complex pricing systems administered by a computer. Such systems, referred to as *yield management systems* (YMS), employ techniques such as discounting early purchases, lim-

iting early sales at these discounted prices, and overbooking capacity (Kimes 1989; Lieberman 1993). The rest of this article refers to those systems that use these techniques as "YMS."³

First developed in the mid-1980s in the airline industry, YMS now are appearing in other services including lodging, transportation, rental firms, hospitals, and satellite transmission (Harris and Peacock 1995). These highly complex, multiperiod pricing systems are both popular and profitable among many airlines. In fact, YMS are now an essential part of the airline business. Commenting on American Airline's disastrous attempt to abandon YMS in favor of simplified pricing, Feldman (1992, p. 54) states, "Attempts to simplify fares have not worked before and there was little reason to expect success." However, attempts at simplification of pricing are finding more success in other industries. For example, railroads and other freight and transportation services are trying to simplify their pricing schedules (Bohman 1992). Companies such as AT&T, American Hawaii Cruises, Premier Hospital Alliance, Yellow Freight System, Novell, CIX, Infonet, and Dow Jones News/Retrieval recently have moved to more simplified pricing (Armbruster 1995; Bryant 1992; Kimball et al. 1994; Snow 1996).

The best action is unclear. Some industries find complex YMS extremely profitable, whereas others do not. Unfortunately, current theory provides little guidance on when traditional YMS should improve profits. Marketers have successfully developed sophisticated mathematical programs to solve dynamic pricing problems (e.g., see the seminal work by Lodish 1980). However, they also need to determine the precise conditions in which traditional YMS will improve profits by better integrating YMS with marketing theory (Karmarkar 1996).

Given a small marginal cost for serving an additional customer, the basic objective of YMS is to adjust price over time to fill all available capacity. A hotel, for example, may charge a rate of \$100 for a date in the distant future. As the hotel fills for that date, YMS raise the price for the limited number of remaining rooms. Sophisticated YMS adjust prices according to the number of early bookings and usually stop taking reservations after exceeding (i.e., overbooking) the number of available rooms.

³The term "Yield or Revenue Management" sometimes is used loosely to refer to any strategy for maximizing revenues.

¹We adopt the U.S. Census's approach of defining *services* as businesses, excluding manufacturing, agriculture, and related activities (e.g., mining and fishing). For example, we include retailing in this definition.

²As with any constraint, temporary actions may add capacity. Restaurants, for example, could add tables temporarily, but permanently relaxing capacity constraints usually involves considerable cost to the service providers.

Ramarao Desiraju is a visiting assistant professor, and Steven M. Shugan is Russell Berrie Foundation Eminent Scholar and Professor of Marketing, Warrington School of Business Administration, Department of Marketing, University of Florida. The authors are listed in alphabetical order for lexicographic convenience; both authors contributed equally to the article. The authors express their gratitude to the *JM* reviewers for their helpful comments and to the Russell Berrie Foundation for its helpful funding.

In practice, YMS tend to fill capacity by partitioning time into discrete periods, charging discount prices in early periods, and reserving capacity by restricting sales in these periods. For example, YMS may divide the time for buying an airline ticket into three periods: (1) 21 days before the flight, (2) 7 to 21 days before the flight, and (3) within 7 days of the flight. Ticket prices often vary by period, for example, \$200, \$250, and \$300, respectively. Finding these prices is difficult and requires complex YMS (e.g., HIRO and AIRMAX) to partition time, set prices, and set limits for sales at each price.

Despite their complexity, YMS depend on the fundamental concepts of multiperiod pricing and market segmentation. Such reasoning suggests that YMS should develop from the multiperiod pricing literature in marketing and economics. However, YMS are built on a foundation more from operations than from marketing or economics. With the notable exception of Harris and Peacock (1995), the marketing literature seldom mentions YMS. The operations literature, in contrast, emphasizes efficient computational tools for the rapid integration of demand information (e.g., Belobaba 1989; Brad and Singh 1996; Smith, Leimkuhler, and Darrow 1992). Although integration is valuable (Blattberg, Glazer, and Little 1994; Glazer 1991), basic marketing and pricing theory is critical.

Thus, it is important to understand YMS at a conceptual rather than computational level. Without that understanding, it is difficult to answer many important questions. For example, we are unable to identify when YMS will improve profits without a full and costly implementation. We are also unable to determine which specific aspects of YMS, such as overbooking or limiting sales, are useful.

This article has two objectives. First, we seek a conceptual foundation for the strategic pricing of capacity-constrained services. Second, we seek precise conditions in which specific strategies (e.g., early discounting and limiting sales) are best. We attempt to present, in a concise and rigorous manner, the conditions in which different strategies will work. We use the term “work” to refer to the long-term objective of maximizing profits.

We view YMS only as tools for implementing an optimal, multiperiod pricing strategy in which each price is a function of forecasted excess capacity. Yield management systems employ a pricing strategy that involves discounting early prices but reserves some capacity for later sale at a higher price.⁴ We determine which factors influence the profitability of this strategy. We ask, for example, when advanced purchase prices should exceed last minute prices and vice versa. We also ask when it is profitable to limit advanced sales and reserve some capacity for last minute arrivals. Still another question we ask pertains to the profitability of allowing sales to exceed available capacity. The answers to these queries should interest service providers that are considering the adoption of YMS. They should also interest academics concerned with the theoretical foundation of YMS and its relationship to market segmentation and multiperiod pricing.

⁴In practice, customers may purchase at a higher price when the discount price becomes unavailable.

The Nature of Arrivals

In this article, we suggest that YMS and the associated strategic pricing problem depend crucially on two factors: customer valuation for the capacity-constrained service and the nature of customer arrivals. Accordingly, we begin by defining two market segments that vary on the basis of customer valuation. Then, we define three classes of service on the basis of the nature of arrivals from consumers in the two segments.

Two Market Segments

We consider a market composed of two segments, in which one of the segments has a higher valuation for the service than the other. The segment with the higher valuation is willing to pay more for the service. We refer to this segment as the *price-insensitive* (PI) segment. We refer to the other segment as the *price-sensitive* (PS) segment.

Three Service Classes

Class A. These services experience early arrivals from the PS segment (i.e., consumers with lower relative valuation or reservation prices). Here, customers who arrive early are willing to pay less than those who buy late.⁵ For Class A services, the PI segment either arrives late or has a greater preference to buy late (i.e., a higher cost of committing to a purchase in an early period).

Examples of Class A services include airlines, hotels, and car rentals. Trade publications suggest that early arrivals are generally individuals or families who are planning vacations or other personal trips far in advance. These customers are usually price sensitive and have low costs of committing to an early purchase. Empirical analysis can determine whether a particular service is of Class A. We believe Class A services usually are found only in the travel industry, in which business travelers, with expense accounts, arrive late.

Class B. These services experience early arrivals from the PI segment (i.e., consumers with higher relative reservation prices). Here, customers who arrive early are willing to pay more than those who buy late.

Examples of Class B services include many bakery services, cellular services, and high-fashion retailers. Arguably, the first customers for these services are willing to pay the highest price. The first customers to arrive at a bakery, early in the day, are the most motivated and want the maximum selection of cakes and pastries, perhaps, for an important dinner party that night. The first customers to buy cellular services or items from high-fashion retailers are the most eager for the items, revealing a greater desire for them. Subsequent arrivals are presumably late because they find less need for the service. Consequently, late arrivals are willing to pay less than early arrivals. Empirical analysis can identify Class B services.

Class C. These services experience arrivals that are uncorrelated with reservation prices. The customers who arrive

⁵To be precise, customers who arrive late prefer to arrive late (i.e., have a higher cost of early commitment). They may arrive early when given the appropriate incentives.

early are likely to have similar reservation prices to those who arrive late. Here, time of arrival is uncorrelated with the segment (PI or PS).

Examples of Class C services include catering, repair, and lawn care services. Here, customer arrivals are generally random with respect to reservation prices. For catering services, customers who arrive late (i.e., want immediate services) may have a great need for the service. However, it is also possible that those customers who book the service far in advance may have a specific desire for a specific date and, consequently, an equal need for the service. Here, the lack of a difference in average reservation prices across periods prevents us from using arrival time to infer willingness to pay for a service.

Table 1 summarizes the three service classes. We now show that marketing strategies should differ by the type of service provider, that is, Class A, B, or C. In the next section, we consider pricing strategies for Class A services in three conditions: (1) predictable demand, (2) partially unpredictable demand, and (3) fully unpredictable demand. With predictable demand, we know the size of each market segment; with partially unpredictable demand, we are uncertain about the size of the PI segment; and with fully unpredictable demand, we are uncertain about the sizes of both the PS and the PI segments. In studying these scenarios, we first consider pricing strategies that do not limit early sales or allow for overbooking of capacity. Subsequently, we identify conditions in which limiting early sales is optimal, and finally, we analyze the trade-offs involved in overbooking capacity. Pricing strategies for services in Classes B and C then are presented in two separate sections. The final section provides a summary and our conclusions.

Pricing Class A Services

Pricing Strategies with Predictable Demand

An important aspect of strategic pricing for capacity-constrained services is the nature of arrivals. We capture the nature of arrivals with a two-period model. Customers arrive in either Period 1 or Period 2. After arrival, they purchase a

ticket for services that usually are consumed in Period 2.⁶ This situation describes many services, including airlines, cruises, sporting events, educational institutions with early admission, and so on. By design, this two-period formulation is simple. Although extensions are straightforward, we seek to remove the complexities often modeled in the operations literature and reveal the underlying forces that drive marketing strategy for capacity-constrained services.

Beyond avoiding needless complexity, we also seek to avoid excessive abstraction. Therefore, to improve the exposition of our model, we frequently use the context of the airline industry as an example. However, tedious, prolonged, and competitive trial and error probably leaves little room for improvement in the airline industry. Therefore, our analysis may be more interesting for new or emerging services such as Internet services, new restaurants, and entertainment services.

For all Class A services, the PS segment arrives in Period 1 (i.e., early arrivals have lower reservation prices). For airlines, early arrivals are often leisure travelers lacking expense accounts, whereas late arrivals are business travelers with expense accounts. Thus, customers who arrive late are willing to pay a higher price, denoted v_{High} , than the lower price, denoted v_{Low} (i.e., $v_{High} > v_{Low}$) paid by customers who arrive early.

In addition to having lower reservation prices (or greater price sensitivity), early arrivals are more flexible. For example, business travelers, in contrast to leisure travelers, may dislike committing to a particular flight because of shorter trip duration (e.g., same weekday travel), uncertainty regarding the ending time for business meetings, or a need to attend unexpected last minute meetings. For Class A entertainment services, some customers may avoid purchasing advance concert tickets because they lessen flexibility. Therefore, the PI segment requires more flexibility. This is an essential prerequisite for the profitability of YMS.

⁶As long as there is no repeat purchase, there is no substantive change to our results if consumers purchasing in Period 1 consume the service at the end of Period 1. Even our subsequent analysis of overbooking does not change in a significant manner.

TABLE 1
Classes of Services

		Price Insensitive Segment	
		Prefer to Arrive Late (High Commitment Cost)	Prefer to Arrive Early (Low Commitment Cost)
Price Sensitive Segment	Prefer to Arrive Early (Low Commitment Cost)	Class A*	Class C
	Prefer to Arrive Late (High Commitment Cost)	Class C**	Class B

*Traditional YMS improve profits mainly for Class A services.

**For class C services, members of each segment may randomly arrive early or late.

TABLE 2
Customer Surplus and Prices by Period

Service Class	Purchase in:	PI Segment	PS Segment
A	Period 1	$v_{High} - c_{High} - p_1$	$v_{Low} - c_{Low} - p_1$
	Period 2	$v_{High} - p_2$	$v_{Low} - p_2$
B	Period 1	$v_{High} - p_1$	$v_{Low} - c_{High} - p_1$
	Period 2	$v_{High} - c_{High} - p_2$	$v_{Low} - p_2$
C	Period 1	$v_{High} - p_1 - c_{Low}$	$v_{Low} - p_1 - c_{Low}$
	Period 2	$(v_{High} - p_1 - c_{High})$ $v_{High} - p_2$	$(v_{Low} - p_1 - c_{High})$ $v_{Low} - p_2$

Note: p_1 = the price charged to those customers who buy in Period 1, p_2 = the price charged to those customers who buy in Period 2, c_{Low} = "low" commitment cost, c_{High} = "high" commitment cost, v_{High} = premium price, v_{Low} = discount price, and $v_{Low} - c_{Low}$ = super-discount price.

We capture a preference for flexibility by positing that late arrivals have a greater cost of commitment and desire flexibility in Period 1 (denoted c_{High}). Conceptually, this cost is the amount an early arrival would pay to wait one period before making a purchase. For many travel services, leisure travelers have a lower cost of commitment. Although they may prefer to book late and retain some flexibility, they are more willing to commit to a purchase in Period 1. Here, leisure travelers have a lower cost of commitment (denoted c_{Low}) in Period 1 than business travelers, namely, $c_{High} > c_{Low} > 0$. Note that neither segment incurs a commitment cost for purchasing in Period 2.

We also require that the difference in commitment costs between the two segments is sufficiently large (i.e., $c_{High} - c_{Low} > v_{High} - v_{Low}$). When that condition is satisfied, we can effectively segment the market into early and late arrivals. If the condition is not satisfied, we are unable to effectively segment the market and YMS fail. Let us consider why.

When early and late arrivals have different commitment costs, service providers can charge a low price in Period 1 and sell exclusively to the PS segment. They can, subsequently, raise the price in Period 2 and charge a higher price to the PI segment. They effectively are using arrival times as a means of price segmentation. That strategy fails when both segments have similar commitment costs, because PI customers can shift their purchases to Period 1. With airlines, for example, the strategy fails when business travelers buy in Period 1.

Although the preceding prerequisites are essential for multiperiod segmentation across time, additional conditions must hold before YMS will work. We now explore those conditions. Let n_L denote the size of the PS segment and n_B be the size of the PI segment. Now providers must decide on prices for Periods 1 and 2, denoted p_1 and p_2 , respectively. Given these prices, customers in each segment decide whether to buy in Period 1, buy in Period 2, or not buy at all.

We use each segment's surplus to predict purchase decisions. For Period 1, PI segment's surplus is the maximum price customers will pay less their costs, namely, $v_{High} - c_{High} - p_1$. Consider the airline example. A business traveler who buys in Period 1 will pay v_{High} , less the cost of commitment, namely $v_{High} - c_{High}$. When the airfare is p_1 , the business traveler enjoys a surplus that consists of what the traveler would have paid, less what was charged, namely,

$v_{High} - c_{High} - p_1$. Note that the surplus becomes greater for a larger valuation v_{High} , a smaller commitment cost c_{High} , or a smaller price p_1 . Table 2 shows the surpluses for business and leisure travelers in Periods 1 and 2. A customer with a positive surplus buys the service, but the service provider incurs an opportunity loss because the customer would have paid more. When the surplus is zero,⁷ the customer buys and the maximum possible price is extracted from that customer. Finally, customers with a negative surplus do not buy.

Table 2 also defines three possible price levels that set customer surpluses to zero. We refer to these as the premium price, the discount price, and the super-discount price. The premium price is the maximum price that the PI segment (e.g., a business traveler) will pay in Period 2, v_{High} . The discount price is the maximum price the PS segment (e.g., a leisure traveler) will pay in Period 2, v_{Low} . Finally, the super-discount price is the maximum price the PS segment will pay in Period 1, $v_{Low} - c_{Low}$. Note that the premium price is greater than the discount price, which is greater than the super-discount price.

If we are the service provider, the following sequence of events describes our problem:

Period 1 begins, we have capacity T , and we announce our first period price, p_1 . We sell D_1 tickets at that price.

Period 2 begins, and we have $T - D_1$ capacity (e.g., seats) remaining for sale. We announce our second period price, p_2 .

We sell D_2 tickets at the p_2 price during the second period.

We deliver the service (e.g., the plane departs) with $D_1 + D_2 \leq T$ capacity taken (e.g., seats filled).

We begin by exploring YMS that consist of an early low price followed by a higher price but that do not limit sales. By assuming that segment sizes are known, we obtain Result 1. All proofs are in the Appendix.

⁷We adopt a common practice in the modeling literature: When consumers are indifferent between buying and not buying (as in the zero surplus case), they buy. The usual rationale is that the firm can offer an infinitesimal discount on the zero surplus price, and consumers then make a positive surplus. Consequently, purchase is induced; however, because the discount is infinitesimal, profits are essentially unaffected.

Result 1: With no uncertainty about the size of the PS and PI segments, traditional YMS (without limiting sales) are no better than (a) a single-price strategy charging the premium price, v_{High} , when the PS and PI segments are both sufficiently large (precisely, $n_L \geq T$, $n_B \geq v_{\text{Low}}T/v_{\text{High}}$) or (b) a single-price strategy charging the discount price, v_{Low} , when the PS segment is large and the PI segment is small. Precisely, $n_L \geq T$, $n_B < v_{\text{Low}}T/v_{\text{High}}$.

Here, having different prices in each period is no more profitable than having a single price. Therefore, YMS selling capacity at different prices (but not limiting sales) fail to improve profits. Before proceeding to more complex situations, we consider the intuition underlying this result in the case of an airline.

When the airline knows that the number of business travelers is sufficiently large to fill the plane, it should reserve all seats for these travelers and charge the premium price, v_{High} , in both periods. It is more profitable to sell all seats at that premium price in the second period than to sell at a lower price in Period 1, because the higher cost of commitment faced by business travelers must be overcome. Therefore, YMS employing multiperiod pricing strategies provide no additional profits over a simple single-price strategy.

When the PI segment becomes sufficiently small, it is profitable to sell capacity to the PS segment. In Result 1b, the size of the PI segment is not sufficiently large to make it profitable to sell tickets to only the PI segment. It is more profitable to sell to a mix of passengers in Period 2. That strategy involves selling at the discount price in both periods, which is superior to selling only to the PS segment in Period 1 at the super-discount price.

The benefits of traditional YMS can come from an insufficiently large PI segment to fill capacity. The advantages of multiperiod YMS, therefore, must come from exploiting this factor. Moreover, without this factor, the more complex and creative sales-limiting strategies, which we discuss subsequently, are unnecessary because price can be sufficiently high to select only the PI segment.

In Result 1a, the size of the PS segment is sufficiently large to fill available capacity. The size of the PI segment is sufficiently large to make it profitable to sell tickets to only the PI segment rather than just the PS segment in Period 1 or a mix of the PS and PI segments in Period 2. It is more profitable because selling at the premium price provides greater revenue than selling either to the PS segment in Period 1 at the super-discount price or to a mix of both segments in Period 2 at the discount price.

Pricing Strategies with Partially Unpredictable Demand

We consider now the case when the service provider is uncertain about the size of the PI segment. We refer to this case as partially unpredictable demand. In this section, we seek to understand how partial uncertainty affects pricing strategy. When we are also uncertain about the PS segment size, we have fully unpredictable demand. The next section considers that case.

Let the size of the PI segment (e.g., number of business travelers) in the market be a random variable (denoted with the bold n_B). For simplicity, let n_B take a high (denoted n_{BH}) or low (denoted n_{BL}) value. If the PI segment size always exceeds capacity, we should sell at the premium price during both periods, fill capacity with the PI segment, and make the maximum possible profit.

Therefore, the problem becomes interesting only when we are unsure about whether the demand from the PI segment is sufficient to fill available capacity; that is, the PI segment size can be either less than ($n_{\text{BL}} < T$) or greater than ($n_{\text{BH}} > T$) capacity. Therefore, we focus on the situation when we expect that the PI segment size (denoted $E[n_B]$) is less than the available capacity (i.e., $E[n_B] < T$). An airline, for example, may expect the number of business travelers to be less than the capacity of the plane.

Using standard dynamic programming, we work backwards (just as we did for Result 1). We first determine the best price in Period 2. Knowing that price, we determine the price in Period 1. There are two possible prices for Period 2. The first is the premium price, v_{High} . With that price, we expect to sell only to PI customers in Period 2, but we also expect to have capacity remaining. The second possible price for Period 2 is the discount price, v_{Low} . With that price, we fill capacity in Period 2 because both segments will buy at the discount price.

We see that the best price in Period 2 depends on the number of seats not sold in Period 1, namely, $T - D_1$. Let $x_0 = T - v_{\text{High}}E(n_B)/v_{\text{Low}}$. Then, the premium price is best when $D_1 > x_0$. The discount price is best when $D_1 < x_0$. Working backwards, we then can determine the Period 1 price that maximizes profits. After that analysis, we obtain Result 2.

Result 2: Result 1 holds when there is uncertainty about the size of the PI segment. More precisely, Result 1 holds substituting $E(n_B) < T$ for n_B .

Again, we find that multiprice YMS can offer no advantage over a single price. Result 1's intuition applies here.⁸ Uncertainty about the size of the PI segment does not change the value of adopting YMS. However, a limitation on the size of both the PS segment and the PI segment can make YMS valuable, as we now demonstrate.

Before considering that possibility, it is worth reiterating that we are unable to sell to the PS segment in Period 1 because doing so requires lowering the price to the super-discount price and selling all seats at that price. We therefore adopt a single-price strategy, which sometimes makes it unprofitable to fill all available capacity (i.e., Result 1a).

Pricing Strategies with Full Unpredictable Demand

The prior section allowed the PI segment size to be uncertain. However, the PS segment size (i.e., n_L) also may be uncertain. With this situation, we obtain Result 3.

⁸Note that Result 2 assumes that the service provider's profits are independent of the variance of the uncertain size of the PI segment. Were it dependent, Result 2 would require modification.

TABLE 3
Best Pricing Strategies for Class A Services

		PI Segment Size	
		Small	Large
PS Segment Size*	Small	YMS** (with regular markets) or discount price	YMS
	Large	Discount price	YMS** (with sales restrictions) or premium price

*The precise definitions of small and large vary by result. See individual results for precise conditions.

**The YMS strategy is a super-discount price in Period 1 and a premium price in Period 2.

Result 3: Define a market as regular when the combined segments, PI and PS, can fill capacity, but each segment alone cannot. Precisely, $E(\mathbf{n}_B) < T$, $E(\mathbf{n}_L) < T$, and $E(\mathbf{n}_B) + E(\mathbf{n}_L) \geq T$. In a regular market, traditional YMS (without limiting sales) are no better than (a) a single-price strategy charging the premium price, v_{High} , when both the PS and the P segments are sufficiently large (precisely, $E[\mathbf{n}_L] \geq v_{High}[T - E(\mathbf{n}_B)]/v_{High} - v_{Low} + c_{Low}$, and $E[\mathbf{n}_B] \geq v_{Low}T/v_{High}$) or (b) a single-price strategy charging the discount price, v_{Low} , when the PS segment is sufficiently large and PI segment is sufficiently small. Precisely, $E(\mathbf{n}_L) \geq T(v_{High} - v_{Low})/v_{High} - v_{Low} + c_{Low}$, and $E(\mathbf{n}_B) < v_{Low}T/v_{High}$. Otherwise, YMS should sell at the super-discount price (i.e., $v_{Low} - c_{Low}$) in Period 1 and at the premium price (i.e., v_{High}) in Period 2.

Corollary: The value of YMS over single-price strategies increases with increases in the capacity, T , and decreases with increases in the PS segment's cost of commitment (c_{Low}).

Result 3 (along with the Corollary) reveals the dynamics of multiperiod pricing for many capacity-constrained services that face low demand conditions. Most important, Result 3 shows that multiperiod YMS (without limiting sales) are only best in restrictive conditions.

In a regular market, the expected size of the PI segment is, by itself, too small to fill capacity. In such market conditions, YMS are profitable only when the expected size of the PS segment is much smaller than capacity. The small size ensures that, though we sell seats at a lower price in the first period to the PS segment, there is sufficient capacity left to accommodate the PI segment in Period 2. Subsequently, it becomes profitable to serve the PS segment in Period 1 at the super-discount price and fill the remaining capacity in Period 2 with the PI segment at the premium price.

Clearly, YMS become more valuable as capacity increases. With sufficiently large capacity, we can serve the entire PS segment while leaving room for the PI segment. An opposing force, however, arises from the cost of commitment for the PS segment. Result 3 indicates that increases in the cost of commitment for the PS segment favor a single-price strategy because the PS segment (e.g., leisure traveler segment) is now less willing to purchase in Period 1. To persuade these consumers to purchase in advance, we must lower the price in Period 1 to the point that such sales

offer no advantage. Thus, we should make all sales in Period 2 when customers do not incur a cost of commitment.

Analogous considerations hold when the combined PI and PS segments are insufficient to fill available capacity. This case is represented in the small/small cell in Table 3. In this case, the value of YMS over single-price strategies depends on whether the high valuation of the PI segment, namely, $(v_{High} - v_{Low})E(\mathbf{n}_B)$, overcomes the cost of getting the PS segment to buy in Period 1, namely, $c_{Low}E(\mathbf{n}_L)$. Therefore, the profitability of YMS in such markets depends on the relative gain in revenues from the PI segment's higher reservation price balanced against the loss in profit from the PS segment's commitment cost.

Result 3 reveals that, when the expected PS segment size is sufficiently large, a single-price strategy dominates the YMS strategy previously outlined. Furthermore, when capacity is sufficiently small, a YMS strategy becomes unnecessary. In such cases, what (single) price is charged depends on the expected size of the PI segment. We generally charge a premium price when the expected PI segment size is large relative to the expected PS segment size; then, we sell only to the PI segment (e.g., business travelers) at the premium price. We generally charge a discount price when the reverse is true; then, we sell to both PI and PS segments at the discount price. Either choice factors in the price sensitivities of each segment. For example, a large reservation price for the PS segment (i.e., a large v_{Low}), other factors held constant, favors the discount price. The high valuation of the PS segment makes it unprofitable to target only the PI segment. In these cases, we sell all capacity at the discount price (v_{Low}) in both periods.

Upon examining these results closely, we begin to discern the enormous advantage of implementing limited sales. On the surface, it seems that profits must go up if we can sell to all the PI consumers at the premium price and fill the remaining capacity with the PS consumers at the super-discount price. The next section refines this intuition and identifies the conditions that favor a limited sales strategy.

Implementing Limited Sales

Although packaged goods manufacturers seldom find it profitable to limit their sales, if we are a Class A service, we may find it profitable to limit the quantity of a service avail-

able for sale. We consider the strategy of limiting sales in Period 1. We limit sales in Period 1 to reserve some capacity for Period 2. We set a limit of D_1 , which is less than capacity T , and sell up to D_1 at the super-discount price. In general, we should leave exactly the capacity we can sell in Period 2. For example, an airline should reserve the number of seats that it expects can be sold to business travelers in Period 2. Implementing this constrained sales strategy can improve profits and provide more situations in which YMS improve profits.

Result 4: Consider a regular market: (a) If the PI segment is sufficiently small (i.e., $E[n_B] \leq c_{Low}T/v_{High} - v_{Low} + c_{Low}$), then the single discount price strategy is best; (b) If the PI segment is sufficiently large (i.e., $E[n_B] > c_{Low}T/v_{High} - v_{Low} + c_{Low}$), then YMS with limited sales are best. The first period optimal price is the super-discount price, $v_{Low} - c_{Low}$, and the second period optimal price is the premium price, v_{High} .

Note that our regular market condition limits the minimum expected size for the PI segment in Result 4a.

Table 3 summarizes these results.⁹ Consistent with prior results, we find that YMS provide no more profit than a single discount price when the PI segment is small and the PS segment is sufficiently large to fill the remaining capacity. Unlike prior results, we find that limiting sales now makes YMS profitable when both the PI and the PS segments are large. In this situation, restricting sales enables us to serve the PS segment.

Result 4 is key to understanding why particular strategies work for only Class A services. Class A services should examine the expected number of PI (high valuation) customers in Period 2. Airlines, for example, should consider the expected number of business travelers (i.e., $E[n_B]$). When that number is sufficiently large, but insufficient to completely fill capacity, it is profitable to sell the capacity to all these customers at the premium price, namely, v_{High} . The remaining capacity should be sold in Period 1 to low valuation customers at the super-discount price, $v_{Low} - c_{Low}$. However, these sales are limited in number to the expected excess capacity, namely, $T - E[n_B]$.

When the expected number of high valuation customers is small, selling at the premium price in Period 2 fills little capacity. It is more profitable to sell to both segments in Period 2 because doing so fills capacity while avoiding losses associated with overcoming the cost of commitment. This strategy is best, even though no premium-priced tickets are sold. For an airline facing many leisure travelers and few business travelers, it is better to fill all seats at the discount price, namely, v_{Low} , in Period 2 than to fill most seats at the super-discount price in Period 1 and some at the premium price in Period 2.

⁹However, the definitions of large and small in Table 3 vary for different results. For precision, therefore, please refer to the specific results.

Overbooking

Thus far, we have assumed that all customers who purchase in Period 1 will use the service in Period 2. However, some of these customers may fail to use the service in Period 2. These customers are called "no-shows."

Various service industries deal with no-shows differently. However, few Class A services force customers to completely forfeit their purchase price. Airlines, for example, may extract a penalty from a no-show but still honor the ticket for a future flight. Therefore, no-shows create problems. They create an opportunity loss for capacity. Aircraft, for example, leave with empty seats that can never be used again.

We must remember, however, that filling empty seats should not be the objective. An airline, for example, may find it more profitable to fill only half the plane with business travelers at the premium price than to fill the entire plane with a mix of passengers all paying the super-discount price. Therefore, we must always consider the relevant profits when comparing different pricing strategies.

Historical records enable service providers to estimate the number of no-shows (denoted C). When that number is greater than zero, we should sell more tickets than the capacity can accommodate. Airline ticket sales, for example, should exceed the number of seats on the flight.¹⁰ This practice is known as overbooking.

When airlines overbook, for example, they sell tickets for all seats (i.e., T), plus a number of overbooked seats (denoted δ). Here, total tickets sold equal $\delta + T$, and we expect C customers will fail to show in Period 2. The optimal overbooking (i.e., δ) may depend on several factors, including the penalty paid when insufficient capacity prevents us from honoring a ticket.

We consider two cases. The first case assumes that C is known or estimated accurately. The second case allows C to be uncertain, taking values in the range from 0 to some maximum C_{max} . (This latter case assumes a uniform distribution.) Consistent with our prior analysis, we allow for sales-limiting strategies. We consider the situation in which the expected number of business travelers is such that the airline finds it profitable to reserve seats in Period 1 to accommodate business travelers in Period 2, selling to both types of travelers.

To examine the forces affecting overbooking of capacity further, we also consider multiple periods of service. These periods of service may occur in a day. For example, theaters may have multiple show times, bus stations often have buses departing at different times, and airlines have multiple departure times.

Let A and B denote two times on a single date in which A is earlier in the day. For example, an airline may have an earlier and later departure to the same destination. Business

¹⁰It is easy to prove this statement because revenues decrease by C times the corresponding fare (less any small revenue from cancellation charges). The total profits will be less than if all the seats are filled.

travelers (i.e., the PI segment) choose the flight that provides them the best time. Some business travelers prefer the early departure A, whereas others prefer the later departure B. The same situation holds for leisure travelers. We apply the prior analysis to A and B (e.g., each departure) to allocate tickets appropriately and profitably in the two periods between the two segments (PI and PS). We then examine the pattern of overbooking.

We assume that there are costs associated with overbooking customers who are unable to obtain services at the expected time. These customers require both compensation and alternative services. For example, airline passengers require compensation and alternative travel arrangements. The following notation defines those costs. Let S be the cost of shifting an overbooked passenger from an earlier service (e.g., flight) to a later service. Let F be the cost of shifting an overbooked passenger to the next day. We expect F to be much larger than S , given the inconvenience and expense associated with accommodating the customer for the increased delay. We also expect that the cost of shifting a passenger to a later service (i.e., S) is less than the average price (e.g., fare for the flight). Finally, we expect that the cost of shifting the overbooked passenger to the next day (i.e., F) is greater than the average price.

Result 5: When limiting sales in Period 1 is best (see Result 4), then, (a) With only one period of service and when cancellations C are known with certainty, we should overbook by exactly the expected number of no-shows. As that number increases, overbooking should increase and we should charge lower average prices and sell more tickets to the PS segment. When the number of cancellations C is unknown and uniformly distributed in the range 0 to C_{\max} , we should overbook less than the expected number of no-shows. (b) With two periods of service (each period facing the same demand) and when C is known with certainty, overbooking for the earlier time period should exceed the number of no-shows.

To understand Result 5, recall the conditions favoring a limiting sales strategy. The PI segment is not sufficiently large to fill capacity, so we sell some tickets at the super-discount price but limit the number sold. We sell the remaining tickets to the PI segment in Period 2. As the expected PI segment size grows (i.e., larger $E[n_B]$), it is profitable to reserve more capacity for them; consequently, we decrease the limit on the number of tickets sold in Period 1. When overbooking frees additional capacity for sale, all of that capacity must, therefore, be sold to PS customers. We try to ensure sufficient sales to all PI customers, regardless of whether we overbook. Consequently, airline overbooking, for example, creates additional ticket sales exclusively to leisure travelers. As Result 5 states, the average fare on a plane with more overbooked seats should be less than the average fare on a plane with fewer overbooked seats.

The actual sequence of events, however, suggests that airlines sell to vacationers in Period 1, leaving some empty seats. Then, in Period 2, the airline fills the remaining seats

with business travelers. The time delay makes it falsely appear that business travelers are the marginal purchasers who get overbooked tickets. However, comparing the composition of ticket buyers on two planes (facing identical demand conditions), one plane with more overbooked seats than the other, we find that the number of tickets sold to business travelers is the same on both flights, but the number of tickets sold to vacation travelers is higher on the one with more overbooking.

Next, we consider the situation in which C is uncertain. Here, though the service provider is risk neutral, the loss function is asymmetric. The service provider incurs a direct cost (S or F) when the expected value of C overestimates C by one unit. The service provider, however, loses the opportunity cost (i.e., a discount price) when the expected value of C underestimates C by one unit. Therefore, uncertainty about C causes less overbooking when the discount price is smaller than S . However, overbooking also can exceed C . When service providers offer multiple delivery times, their best overbooking and pricing strategies change. Mathematically, Result 5b states that first period tickets sold should exceed $T + C - E[n_B]$. An airline, for example, should sell more than C seats over expected capacity for an earlier flight. The reverse is true for the later flight.

Unlike our prior findings, here overbooking may exceed the expected number of no-shows for some time periods. These results suggest why airlines may heavily overbook particular flights. When overbooking, airlines hope to fill more seats without incurring a penalty. That penalty is less severe for earlier flights in the day, compared with the last flight of the day, to a particular destination. Earlier flights enable airlines to shift overbooked passengers profitably to later flights. When demand conditions change, however, such overbooking is not always best because the average price of a seat on either plane is contingent on demand conditions for both flights.

Finally, we note that when different service providers offer one service at different times, the overbooking penalty also may involve the additional penalty of shifting customers to competitors, as well as the opportunity cost of losing future customer loyalty.

Pricing Class B Services

We now consider the optimal pricing strategy when the nature of arrivals is reversed. Here, the customers who arrive early are less price sensitive (i.e., PI segment) than the customers who arrive late (i.e., PS segment). This occurs, for example, at a rock concert or for retailers carrying high-fashion items. For these examples, the buyers who have the greatest valuation for the service, perhaps those who want the best selection, want to wear the latest fashions, or are unwilling to risk being sold out, buy in the first period.

Consider, for example, a rock concert by a particular performer. The concert organizers have a limited number of tickets to sell, and that number represents their capacity. Concertgoers can buy concert tickets in either Period 1 or 2.

Those customers who tend to buy in Period 1 are those who have the higher valuation for the performers at the concert. This segment of customers not only may be willing to make a greater commitment to attend the event than those who buy tickets late, but also may value the concert experience more. These avid concert fans are often willing to wait in line for longer periods to ensure that they obtain seats for the concert. These customers are the PI segment. They are analogous to the business travelers, but they prefer to buy early because they have a small or no cost of commitment for early purchase.

For another example, consider a bakery. Early in the day, the bakery has the best selection of goods, and all goods represent the bakery's capacity. Early arrivals may be most concerned about having the best selection, perhaps for some special event that evening. These customers (i.e., the PI segment) may both be willing to make a greater commitment and have greater valuation for the goods. As the day progresses, later arrivals may have less valuation for the product, and the bakery, facing perishability, may offer discounts near closing time. Our two-period model also approximates this situation, with the exception of overbooking.

For a final example, consider a retailer carrying high-fashion and seasonable merchandise. This retailer orders a limited number of items at the beginning of the season and must sell these items before the season's end. The inventory of these high-fashion items represents the retailer's limited capacity. Retailers often find it difficult to order additional stock because of both the uniqueness of the items and the season's limited duration.

The retailer's most fashionable customers arrive early in the season. With a preference for fashion, they want to buy in Period 1. Their behavior implies a low (or zero) cost of commitment associated with buying early, that is, in Period 1. In contrast, customers who are willing to be less fashionable or prefer to be more conservative have a lower valuation, v_{Low} . They also may have a high cost of commitment c_{High} for buying in Period 1; for example, they may be afraid of buying a high-fashion item that is not yet popular or may never become popular. Table 2 provides the surplus for each customer segment as a function of when they purchase.

Result 6: For Class B services, the best pricing strategy is to adopt a premium price in Period 1, $p_{premium} = v_{High}$, and a discount price in Period 2, $p_{discount} = v_{Low}$. Sales-limiting strategies offer no advantage.

The result indicates that, unlike Class A services, initially Class B services should price high to skim the market and then lower the price to penetrate the market. They should not employ sales-limiting strategies. These results are similar to traditionally optimal marketing strategies for the pricing of durable goods. Therefore, YMS strategies are irrelevant for Class B services.

Remember that a sales-limiting strategy is profitable for Class A services only because those services have limited capacity, and customers who arrive in Period 2 will pay higher prices. More profitable sales are possible in Period 2, so some capacity is reserved for Period 2. Although Class B

services also may have limited capacity, the requirements on customer arrivals are noticeably absent. Thus, limited sales strategies are usually not profitable for Class B services.

Pricing Class C Services

For Class C services, the arrival of customers is random, and both the PI and PS segments arrive both early or late. Here, both segments have the same cost of commitment (see Table 2).

Result 7: For Class C services, when the PI segment is (a) sufficiently large (i.e., $[v_{Low}/v_{High}]T < E[n_B]$), a single-price strategy is best (the premium price v_{High}), (b) otherwise, a multiple-price strategy may be best, but prices may increase or decrease depending on whether sales in Period 1 are lower or higher than expected demand.

Hence, when the PI segment size is large, there is no advantage to changing price over time or limiting sales. Single-price strategies perform as well as both multiperiod pricing strategies and limiting sales.

In regular markets, multiperiod pricing strategies can be more profitable, but traditional YMS are inappropriate. Unlike pricing strategies for Class A and Class B, we cannot predict the direction of price changes. Given our expectations about n_B and n_L , we set our first period price. When Period 1 sales are low, we should probably lower our price in Period 2. When Period 1 sales are high, we should raise our Period 2 price. Here, closed loop feedback strategies are likely to be useful (Lodish 1980).

Conclusions

Our objective was to investigate the concepts underlying YMS and the conditions necessary for YMS to improve profits. In the process, we examined the strategic pricing of capacity-constrained services. Although many computer-aided techniques exist for pricing out capacity, few strategic principles have appeared in the literature.

We step back from specific algorithmic pricing tools and attempt to understand when those tools might be applicable. We seek to present, in a concise and rigorous manner, precise conditions in which different pricing strategies will be profitable. Therefore, we focus on ultimate objectives such as profit maximization, rather than short-term objectives such as filling capacity. We are particularly interested in conditions in which traditional YMS strategies, such as discounting early sales, limiting sales, and overbooking (i.e., selling more tickets than available capacity physically would allow), provide greater profits. The strategic implications of our study are many.

Our analysis shows that costly, complex, multiperiod YMS are far more profitable when a service provider faces different market segments that arrive at different times to purchase the service. One segment must be willing to pay more for the service (e.g., the PI segment here) but also have a high cost of making an early purchase. The other segment must be more price sensitive (e.g., the PS segment) and have a lower cost associated with making an early purchase. To facilitate our analysis, we defined

Class A services as those facing this situation (i.e., early arrivals are more price sensitive and have a lower cost of commitment than customers who arrive late). Although common for some transportation services, this situation is less common for other services such as financial, personal, and retailing services, in which early arrivals in the season may seek the latest clothing fashions and be willing to pay more than late arrivals. The profit potential of YMS may be limited, to a large extent, to the travel industry in which business travelers, with expense accounts, find it more costly to make early purchases than leisure travelers do.

We find that, for Class A service providers, it is often best to restrict sales in early periods to reserve capacity for later periods. Price should start low and increase over time, thereby employing traditional YMS methods. We find sales-limiting strategies improve profits when the PI segment is large but less than total capacity. The PS market segment must have a low cost of commitment. In such cases, it is profitable to reserve capacity for late arrivals. When that cost is high, it is best to sell only in the latter period.

We also find that increased capacity, lower commitment costs for the PS segment, a greater reservation price (or lower price sensitivity) for the PS segment, and a larger PS segment increase the profitability associated with early discounting and limiting early sales. Overbooking for Class A services usually lowers the average price charged and increases profits when no-shows are expected. Overbooking only exceeds the number of no-shows when overbooked customers can be serviced at a later time. A provider may overbook when offering service at different times; for example, an airline may have several flights to the same destination on the same day. Furthermore, when more no-shows are expected, a lower average price should be offered and the amount of overbooked capacity increased.

Unlike Class A services, Class B service providers experience early arrivals from customers with higher reservation prices and lower costs of commitment than late arrivals. Class B providers should start with high prices and then lower prices over time until they fill available capacity. For these services, discounting early purchases and limiting sales both fail to improve profits; consequently, traditional YMS offer no advantage. They should, instead, follow the strategies often employed by high-fashion retailers that must sell all their stock by the end of the season and should mark down their capacity until it is all sold.

Finally, Class C providers experience arrivals that are random with respect to reservation prices. Customers who arrive early are likely to have similar reservation prices to those who arrive late. Class C providers often will find none of the techniques associated with YMS profitable. They find that these strategies offer no advantage because they are unable to use arrival times to identify their customer's willingness to pay. They should start at the "best" price, given their expectations about demand, and then adjust it up or down each period after observing demand.

Our work highlights the importance of linking the tools employed by service providers to the relevant marketing concepts. We hope our work will stimulate additional research in this area. Many problems still remain. For example, further research should consider the impact of competition, additional market segments, channels of distribution, different quality levels, and signaling on YMS.

Appendix

Proof of Result 1. See the Proof of Result 2; the proof technique is identical for Results 1 and 2.

Proof of Result 2. $E(\mathbf{n}_B)$ denotes the expected size of the PI segment, $E(\mathbf{n}_L)$ the expected size of the PS segment, and T the capacity of the plane, with $T \leq E(\mathbf{n}_L)$ and $E(\mathbf{n}_B) \leq T$. Furthermore, the service provider's choice of prices in either period are $v_{\text{High}} > v_{\text{Low}} > v_{\text{Low}} - c_{\text{Low}} > v_{\text{High}} - c_{\text{High}}$. Given this, let x tickets be sold in Period 1; then $T - x$ tickets remain at the beginning of Period 2. If $p_2 = v_{\text{High}}$, then $\min\{E[\mathbf{n}_B], T - x\}$ will purchase; if $p_2 = v_{\text{Low}}$, then $T - x$ will purchase. Note that $v_{\text{High}}E(\mathbf{n}_B) = v_{\text{Low}}(T - x)$ whenever $x = T - v_{\text{High}}E(\mathbf{n}_B)/v_{\text{Low}}$. Define $x_0 = \max\{T - v_{\text{High}}E[\mathbf{n}_B]/v_{\text{Low}}, 0\}$. Then,

$$p_2^* = \begin{cases} v_{\text{Low}} & \text{if } x \leq x_0 \\ v_{\text{High}} & \text{if } x > x_0 \end{cases}$$

Given this optimal price, consider total profits:

$$\pi = \begin{cases} p_1x + v_{\text{Low}}(T - x) & \text{if } x \leq x_0 \\ p_1x + v_{\text{High}}E(\mathbf{n}_B) & \text{if } x_0 < x \leq T - E(\mathbf{n}_B) \\ p_1x + v_{\text{High}}(T - x) & \text{if } T - E(\mathbf{n}_B) < x \leq T \end{cases}$$

The optimal first period price is $v_{\text{Low}} - c_{\text{Low}}$. Then,

$$\pi = \begin{cases} (v_{\text{Low}} - c_{\text{Low}})x + v_{\text{Low}}(T - x) & \text{if } x \leq x_0 \\ (v_{\text{Low}} - c_{\text{Low}})x + v_{\text{High}}E(\mathbf{n}_B) & \text{if } x_0 < x \leq T - E(\mathbf{n}_B) \\ (v_{\text{Low}} - c_{\text{Low}})x + v_{\text{High}}(T - x) & \text{if } T - E(\mathbf{n}_B) < x \leq T \end{cases}$$

First consider the case in which $x \leq x_0$. Because $\pi = v_{\text{Low}}T - xc_{\text{Low}}$, the best choice of x is zero. Therefore, the service provider will set $p_1 = p_2 = v_{\text{Low}}$ and sell the full capacity in Period 2. Denote the profit to the provider in this case as $\pi_{\text{discount}} = v_{\text{Low}}T$.

Now consider the cases in which $x > x_0$. In the region $T - E(\mathbf{n}_B) < x \leq T$, the profits are decreasing in x , and the service provider wants to select the lowest value for x , $T - E(\mathbf{n}_B)$. In the region $x_0 < x \leq T - E(\mathbf{n}_B)$, the profits are increasing in x . Therefore, the largest value of x , that is, $\max\{T - E(\mathbf{n}_B), 0\}$, will be chosen. In either case, when $E(\mathbf{n}_B) < T$, the service provider wants to sell $T - E(\mathbf{n}_B)$ tickets at a super-discount price of $v_{\text{Low}} - c_{\text{Low}}$ in Period 1 and then sell the remaining tickets to the PI segment consumers in Period 2 at a premium price v_{High} . If such a scheme could be implemented, the optimal profits, denoted π_{cap} , would be $\pi_{\text{cap}} = (v_{\text{Low}} - c_{\text{Low}})T + E(\mathbf{n}_B)(v_{\text{High}} - v_{\text{Low}} + c_{\text{Low}})$. However, the service provider cannot implement this solution—when the provider sets the super-discount price in the first period, all tickets are bought by the PS consumers in the first

period, which results in a profit of $\pi_{\text{superdiscount}} = (v_{\text{Low}} - c_{\text{Low}})T$. Comparing π_{discount} and $\pi_{\text{superdiscount}}$, the provider prefers π_{discount} . Next, comparing π_{discount} with the profits from selling the seats at v_{High} to only the PI segment in Period 2 gives the statement of Result 2. QED

Proof of Result 3. Denote the expected number of consumers of each segment by E_i , $i \in \{B, L\}$, where $E_L < T$, $E_B < T$, and $E_L + E_B > T$. Suppose x tickets are sold in Period 1 and $T - x$ tickets are available in Period 2. Let x_B denote the number of tickets bought by the PI customers in Period 1 and x_L those bought by the PS customers, with $x_B + x_L = x$. At $p_2 = v_{\text{Low}}$, all consumers are eligible to buy, and $E_B + E_L - x > T - x$. All remaining capacity will be sold and Period 2 revenues are $v_{\text{Low}}(T - x)$.

At $p_2 = v_{\text{High}}$, only the PI segment is eligible to buy. Because $E_B - x_B$ is the expected number of PI customers in the market in Period 2, the key inequality to study is whether E_B is less than $T - x + x_B (= T - x_L)$. If $E_B \geq T - x_L$, then all tickets will be sold and the revenues are $v_{\text{High}}(T - x)$. If $E_B < T - x_L$, then some tickets will be left unsold and the revenues are $v_{\text{High}}(E_B - x_B)$.

Comparing the revenues from these two prices, we find that whenever $E_B \geq T - x_L$, the optimal price, $p_2^* = v_{\text{High}}$. Otherwise, the optimal price depends on the relative magnitudes of $v_{\text{Low}}(T - x)$ and $v_{\text{High}}(E_B - x_B)$. Therefore, the two period profits are

$$\pi = \begin{cases} p_1 x + v_{\text{Low}}(T - x) & \text{if } 0 \leq E_B < x - x_L \\ + \frac{v_{\text{Low}}}{v_{\text{High}}}(T - x) & \\ p_1 x + v_{\text{High}}(E_B - x_B) & \text{if } x - x_L + \frac{v_{\text{Low}}}{v_{\text{High}}}(T - x) \\ \leq E_B < T - x_L & \\ p_1 x + v_{\text{High}}(T - x) & \text{if } E_B \geq T - x_L \end{cases}$$

Note that if $p_1 = v_{\text{High}}$, or if $p_1 = v_{\text{Low}}$, then no one buys in the first period. If $p_1 = v_{\text{High}} - c_{\text{High}}$, then all will be eligible to buy. Finally, if $p_1 = v_{\text{Low}} - c_{\text{Low}}$, then only the consumers from the PS segment will buy. Therefore, to maximize profits, the optimal first period price is $v_{\text{Low}} - c_{\text{Low}}$.

Suppose (1) $0 \leq E_B < x - x_L + (v_{\text{Low}}/v_{\text{High}})(T - x)$ holds (i.e., the number of PI consumers is quite small), then total profits are $(v_{\text{Low}} - c_{\text{Low}})x + v_{\text{Low}}(T - x)$. The best x is zero, and the service provider sells all its tickets at a discount price, v_{Low} ; (2) $x - x_L + (v_{\text{Low}}/v_{\text{High}})(T - x) \leq E_B < T - x_L$ holds, then the profits are $(v_{\text{Low}} - c_{\text{Low}})x + v_{\text{High}}E_B$, because no PI customer purchases a ticket in the first period because of the optimal prices. The optimal x equals $T - E_B$; and (3) $E_B \geq T - x_L$ holds, then profits are $(v_{\text{Low}} - c_{\text{Low}})x + v_{\text{High}}(T - x)$, and the optimal x is as small as possible and equals $T - E_B$.

However, if there are no limiting sales, then all PS customers (and not just $T - E_B$) will purchase at the super-discount price. This would cut into the service provider's profits because $E_L > T - E_B$. Therefore, if the provider priced naively, its profits would be $(v_{\text{Low}} - c_{\text{Low}})E_L + v_{\text{High}}(T - E_L)$. These profits must be compared with the

profits from selling (1) all tickets at v_{Low} and (2) only to the PI segment at v_{High} . The statement of Result 3 follows directly from these comparisons. QED

Proof of Result 4. Recalling the proof of Result 2 and comparing π_{cap} with π_{discount} , we have $(v_{\text{Low}} - c_{\text{Low}})T + E(\mathbf{n}_B)(v_{\text{High}} - v_{\text{Low}} + c_{\text{Low}}) \geq v_{\text{Low}}T$. This reveals that, if the expected number of PI customers is sufficiently large (i.e., $E(\mathbf{n}_B) \geq Tc_{\text{Low}}/[v_{\text{High}} - v_{\text{Low}} + c_{\text{Low}}]$), then $\pi_{\text{cap}} \geq \pi_{\text{discount}}$. Otherwise, $\pi_{\text{cap}} < \pi_{\text{discount}}$. QED

Proof of Result 5. Suppose E_B is such that the service provider implements limited sales (see Result 4); then, E_B tickets are expected to be sold to the PI customers in Period 2. Now let the provider oversell or overbook by δ tickets. Given a total capacity T , the provider now sells $T + \delta - E_B$ tickets at the super-discount price. Using p_L and p_H to denote the lower and higher prices, respectively, of a limited sales strategy, the average price, p_{avg} , paid per ticket as a function of δ is $p_{\text{avg}}(\delta) = [p_L(T + \delta - E_B) + p_H E_B]/(T + \delta) = p_L + [(p_H - p_L)E_B]/(T + \delta)$. The first derivative, $\partial p_{\text{avg}}/\partial \delta = -[E_B(p_H - p_L)]/(T + \delta)^2$. This is decreasing in δ .

Now suppose there are C cancellations. For each cancellation, the service provider pays back the cost of the ticket but obtains a small cancellation fee, α . If there are fewer cancellations than overbookings, the provider pays a rather hefty fine of F , $F \gg p_{\text{avg}}$ per each overbooked ticket that could not be accommodated. Then, the service provider's profits are

$$\pi = \begin{cases} p_{\text{avg}}(T + \delta - C) + \alpha C & \text{if } \delta \leq C \\ p_{\text{avg}}(T + \delta - C) + \alpha C - F(\delta - C) & \text{if } \delta > C \end{cases}$$

If C is known with certainty, then for a given level of cancellations, π is concave and increasing in δ up to C ; when δ goes beyond C , π is again concave but decreasing in δ . Therefore, the optimal number of overbookings equals C . Because average price of a ticket is decreasing in the number of overbooked tickets, the provider with fewer overbookings has the higher average price.

Suppose C is a random variable with a density function, $f(C)$. Furthermore, let C be uniformly distributed over the interval $[0, C_{\text{max}}]$, where C_{max} is the maximum number of cancellations. Then, the expected value of C is $1/C_{\text{max}} \int_0^{C_{\text{max}}} C dC = C_{\text{max}}/2$. Furthermore, the expected value of $F(\delta - C)$ is $1/C_{\text{max}} \int_0^{\delta} F(\delta - C) dC = \delta^2/2C_{\text{max}}$. Because the expected profits are quadratic and concave in δ and ignoring effects on the margin, the optimal δ is $\delta^* = p_{\text{avg}} C_{\text{max}}/2$. Note that δ^* decreases with F and increases with C_{max} and p_{avg} .

For ease of exposition, we present the analysis in the context of a risk-neutral airline that offers two flights, one in the morning and the other in the evening, on a given day. We assume that the demand conditions at one period of service are independent of the demand conditions at the other. The following notation will be used: The two flights are represented by X_i , $i \in \{1, 2\}$, where X_1 is the earlier flight. The average price of a seat on X_i is denoted by $p_{\text{avg}}^i(\delta)$, where δ is the number of overbooked seats on that flight. The capacity of each flight is identical and equals T ; C_i denotes the number of expected cancellations on the i th flight, and $E_i B$

denotes the expected number of PI customers on the i th flight.

If the risk-neutral airline overbooks by exactly the number of cancellations on each of the two flights, its profits are $\pi^* = [p_{avg}^1(C_1) + p_{avg}^2(C_2)]T + \alpha(C_1 + C_2)$, where $p_{avg}^i(C_i) = p_L + [(p_H - p_L)E^i_B]/(T + C_i)$. Suppose that X_2 overbooked δ seats, $\delta < C_2$. Now, the airline can shift $C_2 - \delta$ overbooked passengers (who show up) from X_1 to X_2 and incur the lower cost S (than F) when overbooking X_1 beyond C_1 . Note that when the first plane is overbooked by B , the expected profit from that plane, π_1 , is

$$\pi_1 = \begin{cases} p_{avg}^1(B)(T + B - C_1) + \alpha C_1 & \text{if } B \leq C_1 \\ p_{avg}^1(B)(T + B - C_1) + \alpha C_1 - S(B - C_1) & \text{if } C_1 < B \leq C_1 + C_2 - \delta \\ p_{avg}^1(B)(T + B - C_1) + \alpha C_1 - S(C_2 - \delta) & \text{if } C_1 + C_2 - \delta < B \\ -F(B - C_1 - C_2 + \delta) & \text{if } C_1 + C_2 - \delta < B \end{cases}$$

Our interest is in the case when $B = C_1 + C_2 - \delta$, because the second plane is underbooking by $C_2 - \delta$. In these conditions, the profit from X_2 is $\pi_2 = p_{avg}^2(\delta)(T + \delta - C_2) + \alpha C_2$. Total profit is $\pi_1 + \pi_2 = [p_{avg}^1(C_1 + C_2 - \delta) + p_{avg}^2(\delta)]T + \alpha(C_1 + C_2) + [(p_{avg}^1 - S) - p_{avg}^2](C_2 - \delta)$. The airline will move overbookings from X_2 to X_1 only when $(\pi_1 + \pi_2)$ exceeds the benchmark profits, π^* . Defining $\Delta = \pi_1 + \pi_2 - \pi^*$, we have $\Delta = [p_{avg}^1(C_1 + C_2 - \delta) - p_{avg}^1(C_1) + p_{avg}^2(\delta) - p_{avg}^2(C_2)]T - [p_{avg}^2(\delta) + S - p_{avg}^1(C_1 + C_2 - \delta)](C_2 - \delta)$.

Substituting for the average prices and computing $\partial\Delta/\partial\delta$, we have $\partial\Delta/\partial\delta = S + (p_H - p_L)\{[C_2E^2_B/(T + \delta)^2] - [C_1E^1_B/(T + C_1 + C_2 - \delta)^2]\}$. Note that this derivative is decreasing in δ . Setting $S = 0$ and $C_1 = C_2 = C$, the optimal δ is $\delta^* = C[2\sqrt{E^2_B}/(\sqrt{E^1_B} + \sqrt{E^2_B}) + T[(\sqrt{E^2_B} - \sqrt{E^1_B})/(\sqrt{E^1_B} + \sqrt{E^2_B})]$. The properties of δ^* follow from the derivatives, $\partial\delta^*/\partial E^1_B = -(C + T)\sqrt{E^2_B}/\sqrt{E^1_B}(\sqrt{E^1_B} + \sqrt{E^2_B})^2 < 0$ and $\partial\delta^*/\partial E^2_B = (C + T)E^1_B/E^2_B(E^1_B + E^2_B)^2 > 0$. Finally, note that in this context, the airline does not have an incentive to overbook X_1 by δ seats, $\delta < C_1$, and overbook X_2 by $C_2 + C_1 - \delta$ because such overbooked (i.e., $C_1 - \delta$) passengers must be compensated F , which is higher than the price of the ticket. The statement of Result 5 follows.

Proof of Result 6. There are three cases to consider here: (1) The number of available PI segment consumers is greater than or equal to T . In this case, the best price to offer is v_{High} ; (2) The number of available PI segment consumers is less than T . Then, in Period 1, the price will be v_{High} to extract the surplus from the PI segment. In Period 2, the price will be reduced to v_{Low} to sell off the remaining capacity to the PS segment; and (3) The total number of PI and PS consumers in the market is less than the available capacity. Even here, the optimal pricing strategy is to offer the service at a higher price in the first period v_{High} , extract the surplus, and lower the prices in the second period to v_{Low} . Furthermore, there is no necessity to limit sales in these types of settings because of the higher prices that can be imposed on the less price sensitive consumers who arrive early. As a consequence, the retailer has no incentive to limit the sales to such consumers. QED

Proof of Result 7. This proof follows directly from the arguments in the text, comparing the profits from the various pricing strategies reveals the statement of Result 7. QED

REFERENCES

- Armbruster, William (1995), "Forward Air Prepares 7.4 Percent Rate Increase," *Journal of Commerce*, (January 26), Transportation Section, 2b.
- Belobaba, P. Peter (1989), "Application of Probabilistic Decision Model to Airline Seat Inventory Control," *Operations Research*, 37 (March-April), 183-98.
- Blattberg, Robert, Rashi Glazer, and John Little, eds. (1994), *The Marketing Information Revolution*. Boston, MA: Harvard Business School Press.
- Bohman, Ray (1992), "Railroads Simplify Boxcar Pricing," *Traffic Management*, 31 (October), 25.
- Brad, Anca and Ajay Singh (1996), "Path-Based Demand Forecasting at United Airlines," presented at the Fall 1996 INFORMS Meeting, Atlanta, Georgia.
- Bryant, Adam (1992), "Many Companies Try to Simplify Pricing," *The New York Times*, (October 18), 4:6.
- Feldman, Joan M. (1992), "The Price of Retribution: American Airlines' Fare Cuts, Coming on Top of Restructuring, Contributed to the Industry's Summer of Big Losses," *Air Transport World*, 29 (December), 54.
- Glazer, Rashi (1991), "Marketing in an Information-Intensive Environment: Strategic Implications for Knowledge as an Asset," *Journal of Marketing*, 55 (October), 1-20.
- Harris, Frederick H. and Peter Peacock (1995), "Hold My Place, Please," *Marketing Management*, 4 (Fall), 34-46.
- Heuslein, William (1993), "Carl Icahn Is Out. Can Al Checchi and Gary Wilson Be Far Behind?" *Forbes*, (January 4), 178.
- Johnson, Eugene M. (1964), *An Introduction to Service Marketing Management*. Newark, DE: University of Delaware.
- Karmarkar, Uday S. (1996), "Integrative Research in Marketing and Operations Management," *Journal of Marketing Research*, 33 (May), 125-27.
- Kimball, James G., Jennifer Lawrence, Russell Shaw, Ira Teinowitz, and Sandra Pesmen (1994), "The Top 100," *Business Marketing*, (October 3), 57.
- Kimes, Sheryl E. (1989), "Yield Management: A Tool for Capacity-Constrained Service Firms," *Journal of Operations Management*, 8 (October), 348-63.
- Lieberman, Warren H. (1993), "Debunking the Myths of Yield Management," *Cornell Hotel & Restaurant Administration Quarterly*, 34 (February), 34.
- Lodish, Leonard M. (1980), "Applied Dynamic Pricing and Production Models with Specific Application to Broadcast Spot Pricing," *Journal of Marketing Research*, 17 (May), 203-11.
- Parasuraman, A. and P. Rajan Varadarajan (1988), "Future Strategic Emphases in Service Versus Goods Businesses," *Journal of Services Marketing*, 2 (4), 57-66.
- Radas, Sonja and Steven M. Shugan (1998a), "Seasonal Marketing and Timing New Product Introductions," *Journal of Marketing Research*, 35 (August), 296-315.

—— and —— (1998b), "Managing Service Demand: Shifting and Bundling," *Journal of Services Research*, 1(August), 47-64.
 Rust, Roland T., Anthony J. Zahorik, and Timothy L. Keiningham (1995), "Return on Equity (ROQ): Making Service Quality Financially Accountable," *Journal of Marketing*, 59 (April), 58-70.
 Shugan, Steven M. (1994), "Explanations for the Growth of Services," in *Service Quality: New Directions in Theory and Practice*, Roland T. Rust and Richard L. Oliver, eds. Newbury Park, CA: Sage Publications, 223-240.

Smith, Barry C., John F. Leimkuhler, and Ross M. Darrow (1992), "Yield Management at American Airlines," *Interfaces*, 22 (January-February), 8-31.
 Snow, Charlotte (1996), "New Supply Deals Strive for Simplicity," *Modern HealthCare*, (October 28), 49.
 Wardell, David, (1989), "Yield Management: Understanding and Exploiting It," *Travel Weekly*, 48 (March 20), 75.
 Wilder, Clinton (1991), "Amtrak Follows Airline Route," *Computer World*, (February 4), 8.

**U.S. Postal Service Statement of Ownership,
 Management, and Circulation (Required by 39 U.S.C. 3685)**

1. Title of Publication: *Journal of Marketing*
2. Publication No. 0022-2429
3. Filing Date: October 1, 1998
4. Issue Frequency: Quarterly
5. No. of Issues Published Annually: 4
6. Annual Subscription Price: \$75
7. Complete Mailing Address of Known Office of Publication: 250 South Wacker Dr., Suite 200, Chicago, IL 60606-5819
8. Complete Mailing Address of the Headquarters or General Business Offices of the Publisher: 250 South Wacker Dr., Suite 200, Chicago, IL 60606-5819
9. Full Names and Complete Mailing Address of Publisher, Editor, and Managing Editor:
 Publisher: Jack Hollfelder, 250 South Wacker Dr., Suite 200, Chicago, IL 60606-5819; Editor: Robert Lusch, Univ. of Okla., Coll. of Bus. Adm., 307 W. Brooks, Rm. 9, Norman, OK 73019-0450; Managing Editor: Francesca Van Gorp Cooley, 250 S. Wacker, Suite 200, Chicago, IL 60606-5819
10. Owner: American Marketing Association, 250 South Wacker Dr., Suite 200, Chicago, IL 60606-5819
11. Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1% or More of Total Amount of Bonds, Mortgages, or Other Securities (If there are none, so state): None
12. For completion by nonprofit organizations authorized to mail at special rates. The purpose, function, and nonprofit status of this organization and the exempt status for federal income tax purposes: Has Not Changed During Preceding 12 Months
13. Publication Name: *Journal of Marketing*
14. Issue Date for Circulation Data Below: October, 1998
15. Extent and Nature of Circulation:

	Average no. copies each issue during preceding 12 months	Actual no. copies of single issue published nearest to filing date
a. Total No. Copies (Net Press Run)	10,025	10,500
b. Paid and/or Requested Circulation		
(1) Sales through dealers and carriers, street vendors, and counter sales (not mailed)	0	0
(2) Paid or Requested Mail Subscriptions (Include Advertisers' Proof Copies/Exchange Copies)	7,908	8,955
c. Total Paid and/or Requested Circulation (Sum of 15b(1) and 15b(2))	7,908	8,955
d. Free Distribution by Mail (Samples, Complimentary, and Other Free)	1,507	935
e. Free Distribution Outside the Mail (Carriers or Other Means)	0	0
f. Total Free Distribution (Sum of 15d and 15e)	1,507	935
g. Total Distribution (Sum of 15c and 15f)	9,415	9,890
h. Copies Not Distributed		
(1) Office Use, Leftovers, Spoiled	610	610
(2) Return from News Agents	0	0
i. Total (Sum of 15g, 15h(1), and 15h(2))	10,025	10,500
Percent Paid and/or Requested Circulation (15c/15g x 100)	84.0% paid circulation	90.5% paid circulation

16. This Statement of Ownership will be printed in the January, 1999, issue of this publication.
17. Signature and title of editor, publisher, business manager, or owner: Jack Hollfelder, Publisher. I certify that the information furnished on this form is true and complete. I understand that anyone who furnishes false or misleading information on this form or who omits material or information requested on the form may be subject to criminal sanctions (including fines and imprisonment) and/or civil sanctions (including multiple damages and civil penalties).