Targeting Teaching

Diagrammatic Approach to Capacity-Constrained Price Discrimination

William S. Reece* and Russell S. Sobel†

This paper presents a diagrammatic solution to the firm’s profit-maximizing price discrimination problem in the face of capacity constraints. Airlines, hotels, and other firms practice yield management, allocating fixed capacity to customer groups paying different prices. In these cases, the firm’s short-run problem is not a decision about production levels, but it is one of allocating a fixed number of output units among customers. Our diagram shows that the conditions for profit-maximizing price discrimination are very different under these circumstances than in the standard model in which the firm is not constrained by capacity.

1. Introduction

News media regularly report consumer frustration and confusion over pricing in travel and tourism. USA Today reports in an article on cruise discounts that “only a chump pays full fare” (Stoddart 1999). The New York Times laments complexity in airline pricing in an article entitled “So, How Much Did You Pay for Your Ticket?” (Wald 1998). The Wall Street Journal reports on techniques to get low airfares in “How Farebusters Play the Airlines” (Keates 1998). This complexity, confusion, and frustration are the inevitable result of a rational process designed to solve a very difficult and common business problem: optimal pricing when the firm has fixed capacity and faces classes of consumers with different demands. In the airline’s cases, once the airline assigns a particular aircraft to a particular flight, the number of seats is fixed. It faces many different kinds of demand for those seats, including business travelers trying to fly on short notice, senior citizens considering a variety of travel alternatives, and college students planning months in advance to return home for a holiday. Under these circumstances, how will the airline price its seats to maximize profits?

This paper examines the firm’s profit-maximizing pricing problem in the face of capacity constraints using a simple diagram suitable for use in classes in intermediate microeconomics, industrial economics, and applied subjects such as the economics of transportation or tourism. The topic has important applications: Airlines, hotels, universities, theaters, and other firms practice price discrimination as part of systems to allocate fixed capacity to customer groups paying different prices. The airline industry was the first major industry to implement formal

---

* Department of Economics, P.O. Box 6025, West Virginia University, Morgantown, WV 26506-6025, USA; corresponding author.
† Department of Economics, P.O. Box 6025, West Virginia University, Morgantown, WV 26506-6025, USA.

The authors acknowledge the helpful comments of two referees.
systems to solve this complex problem. Airlines recognized that if they did not implement controls, a passenger willing to pay only a low price could, by reserving early, take a seat away from a passenger willing to pay much more for that seat. To prevent such revenue losses, the airlines implemented a control process known as yield management. These practices later spread to other areas such as the lodging industry. To illustrate the relevance of the capacity-constrained model, during 1997 the 10 largest U.S. air carriers denied boarding to more than one million ticketed passengers from flights filled to capacity.

Many previous authors have examined price discrimination in yield-management systems. Some of these authors, including many using mathematical programming approaches to carefully address problems associated with demand uncertainty over time, take the prices as predetermined exogenously. Others give standard explanations of price discrimination but do not illustrate the fundamental changes to the problem created by the capacity constraint. Kraft, Oum, and Tretheway (1986) in considering stochastic demand suggest that airlines may simulate various alternative discounts to find the profit-maximizing level taking the full fare as given. This paper shows a diagrammatic solution to a simplified version of the yield-management problem. Our diagram shows how both prices and quantities for each customer group can be determined endogenously, rather than taking prices as predetermined.

A simplified version of optimal capacity allocation with price discrimination can be described as a few standard steps. First, the seller must segment the market into groups of customers with different demands. Second, the seller creates restrictions that separate the categories of service offered to the customer groups. For example, requiring a Saturday night stay will in many cases separate business travelers from leisure travelers. In the third step, the seller establishes a price for each category based on anticipated demand. Finally, the seller allocates its fixed inventory among the categories. For example, if there are 130 coach seats on a particular flight, the airline might create 3 fare categories: deep discount, discount, and full fare, requiring, respectively, 14-day advance purchase and Saturday stay, 7-day advance reservation and Saturday stay, and no restrictions. The airline allocates some portion of the 130 coach seats to each of these categories. Over a period of months, as the flight time approaches the airline may reallocate seats to categories depending on sales. In principle, the prices may remain constant over time, while the availability of fare categories changes as seats in categories with lower fares become filled and the category becomes unavailable. This typically means that the discount and deep discount seats go primarily to leisure fliers and that business fliers who cannot meet the restrictions pay full coach fare.

Even such a simple yield-management process could be the source of the consumer confusion and frustration so often reported. At some particular point in time, the round-trip airfare from Baltimore to Minneapolis might be twice the fare for the round-trip from Minneapolis to Baltimore, even though both trips involve the same travel, one flight in each direction. This

---

1 See Smith, Leimkuhler, and Darrow (1992) and Kraft, Oum, and Tretheway (1986) for discussions of the history and practices of airlines' yield management.
2 See, for example, Kimes (1989) and Hanks, Cross, and Noland (1992) for discussions of yield management in the lodging industry.
3 U.S. Department of Transportation (1998, p. 23). Of course, many more passengers were unable to purchase tickets for sold-out flights as well.
4 See, for example, Belobaba (1989), Brumelle et al. (1990), Curry (1990), and Weatherford and Bodily (1992).
6 We do not address the issues related to demand uncertainty here because our intent is to present a basic diagrammatic exposition.
would occur, for example, when the discount fare categories are still available for the flight originating in Minneapolis but are no longer available for the flight originating in Baltimore. The following day’s flights will have some different set of apparent anomalies, depending on how airline bookings have progressed in the intervening time. With yield management, at any point in time comparisons of airfares for some set of flights will seem mysterious because some discount fare categories will be closed to further bookings while others remain available. The diagram discussed below illustrates price and quantity setting in yield-management systems.

2. Simple Price Discrimination

One standard explanation of price discrimination is that a firm is able to separate consumers into two or more groups having different demands and to charge different prices to the different groups. Figure 1 illustrates the process for an airline with its expected demand divided into business and leisure segments. As is standard in this type of graphical presentation, we must assume constant marginal cost to avoid a marginal cost (MC) function that depends on the sum of the two quantities. The firm maximizes profits by selling the quantities where marginal revenue (MR) equals MC in each market segment. The firm’s profit-maximizing level of output in the business segment, \( Q_B \), is the quantity that equates MR and MC, \( MR_B = MC_B = MC \). Similarly, \( Q_L \) is the quantity that makes \( MR_L = MC_L = MC \). The firm’s total production will be \( Q_B + Q_L \). At each of these two quantities, the firm sets its price along the demand curve at the levels \( P_B \) and \( P_L \). We have drawn the figure so that business travelers have the less elastic demand and thus pay the higher price.

3. Price Discrimination with Capacity Constraints

The airline seat assignment problem differs from the standard case of price discrimination in an important way: Once the airline has assigned a particular aircraft to serve the route, the number of seats available becomes fixed. In the short run, the airline is not free to choose \( Q_B + Q_L \) exceeding fixed capacity. In such a case, the standard picture does not illustrate the

---

7 We assume throughout that firms set prices based on expected demands. Firms may, of course, change these prices in the future as experience changes expectations about demand.
airline’s problem. Suppose that there is an increase in the business segment demand for travel, shifting the business demand and marginal revenue in Figure 1 to the right. This would result in an increase in $Q_B$. Now, $Q_B + Q_L$ could exceed the plane’s fixed capacity, so that expansion in $Q_B$ would be met with an offsetting decrease in $Q_L$. At this point, however, the relevant marginal cost of satisfying the additional business demand is no longer $MC$ but is rather the forgone revenue from selling the seat to a leisure traveler. As we will see, the airline’s profit-maximizing choice would not be to expand business quantity to this new level and reduce leisure quantity by an offsetting amount. With fixed capacity, the problem is one of allocation between competing demands, rather than a decision about the level of production.

If there were only one customer group, we could impose a capacity constraint by having the marginal cost curve become vertical at capacity. With two groups, the additive nature of the constraint between the two groups requires a different solution. Figure 2 introduces the capacity constraint by constructing the mirror image of the leisure graph in Figure 1 and attaching it to the right-hand side of the business graph, as in Figure 2.\(^8\)

In Figure 2, the length of the horizontal axis is set to the capacity level, for example, the fixed number of coach seats on a particular aircraft. The lengths $Q_B$ and $Q_L$ are forced by construction to sum to the fixed capacity. In this paper, we take $MC_{OP}$ to be the additional operating cost of filling a currently empty seat with an additional passenger on an aircraft with fixed capacity. For example, we can think of $MC_{OP}$ as the per-passenger cost of meals and other flight attendant service. The firm’s optimal decision, as usual, is to choose the level $Q_B$ that makes $MR_B = MC_B$ and the level $Q_L$ that makes $MR_L = MC_L$. In this case, however, the marginal cost of selling an additional business seat is not $MC_{OP}$. This cost becomes irrelevant to optimal price and quantity decisions in cases in which every seat will be sold. If every seat will be sold, the cost of meals and other flight attendant service is constant regardless of the allocation of the seats between the two groups.

In Figure 2, looking from left to right, the marginal cost of an additional business passenger

\(^8\) Layson (1988) introduces a figure like our Figure 2, except that the points $v$, $s$, and $t$ are all at a single point because he does not impose a capacity constraint. Our Figure 4, which shows the long-run situation without the capacity constraint, is basically the same as his Figure 1.
is $MC_{op}$ only up until the point $s$. At levels of business passengers beyond the point $s$, the marginal cost of an additional business passenger is the marginal revenue forgone from leisure passengers. The point $s$ is the point at which marginal operating cost, $MC_{op}$, intersects the leisure marginal revenue curve. Given that the airline has decided to make the flight, the true marginal cost curve for business passengers is the flat line at $MC_{op}$ up to the point $s$, and beyond $s$ marginal cost is the line $MR_L$ for higher levels of $Q_B$. Similarly for leisure passengers, looking now from right to left, the marginal cost curve is the flat line at level $MC_{op}$ up to the point $t$ and the line $MR_B$ for higher levels of $Q_L$. The complexity of the seat allocation and pricing problem results from this unusual marginal cost function, which in part is the marginal revenue function for the other group.

The equalities $MR_B = MC_B$ and $MR_L = MC_L$ are satisfied at the point $v$. Dropping straight down from $v$ to the quantity axis divides the capacity optimally between business and leisure passengers. In this model, prices are simultaneously determined with the optimal allocation of capacity. To find the prices, move up from the point $v$ to the leisure demand curve to find the leisure price and to the business demand curve to find the business price. Thus, $P_L$ and $P_B$ are the airline’s optimal discount fare (leisure) and full fare (business). Once the airline sells $Q_L$ tickets at the discount fare, that fare class is no longer available. There are, however, $Q_B$ tickets available at full fare.$^{10}$

This diagram highlights the fact that if a firm is operating in the short run with a binding capacity constraint, the marginal cost of serving an additional passenger in an otherwise empty seat, $MC_{op}$, is irrelevant to the pricing and allocation decisions of the firm. At capacity, the true marginal cost is the forgone revenue from the other group. Returning to Figure 2, assume that the price of meals and other passenger-related costs change, shifting $MC_{op}$ either up or down. As long as $MC_{op}$ remains below the intersection of the two marginal revenue curves, it will not change the firm’s profit-maximizing prices or the optimal allocation of seats. Thus, when the firm operates with a binding capacity constraint, changes in the cost of labor and other inputs in the short run do not affect its prices.

This diagram shows another important implication of the capacity constraint: Changes in one group’s demand affect the price charged to the other group. Because prices in this model allocate capacity, a higher demand by one group raises the price to the other group. This can be seen in the above model by examining the impact of a shift in one of the demand curves. This is in contrast to the standard model without capacity constraints. In the standard model, a change in the demand of one group affects the price charged to the other group only if marginal cost is upward sloping in total production. In the face of binding capacity constraints, however, even with constant marginal operating cost, a change in the demand of one group affects the price charged to the other group.

4. Price Discrimination with Nonbinding Capacity Constraints: Flying with Empty Seats

With this diagram, it is easy to demonstrate the circumstances under which the airline would plan to fly the aircraft with some empty seats. The airline will decline additional pas-

---

$^9$ See Botimer (1996) for an extensive treatment of the welfare implications of airline yield management.

$^{10}$ This diagram illustrates a solution that could be shown mathematically using the Kuhn-Tucker conditions for constrained optimization. See, for example, Varian (1992, pp. 503–5).
sengers and plan to fly with empty seats if the marginal revenue from the additional tickets sold (not the price of the next ticket) is less than the marginal cost $MC_{Op}$. Figure 3 shows this situation.

Once again, the profit-maximizing passenger levels will be the levels that make $MR_B = MC_B = MC_{Op}$ and $MR_L = MC_L = MC_{Op}$. These are shown as $Q_B$ and $Q_L$ in Figure 3. In this figure, the marginal revenue curve for business intersects the marginal cost curve for business before marginal cost begins rising along the leisure marginal revenue curve. The same is true for the marginal revenue and marginal cost of leisure passengers. The distance between $Q_L$ and $Q_B$ is the number of empty seats. The airline will not go beyond $Q_L$ or $Q_B$ because the additional revenue from selling an additional seat is less than the additional cost of having another passenger on the plane. In effect, the problem now collapses back to the standard model because the capacity constraint is not binding. In the case of a nonbinding capacity constraint, changes in $MC_{Op}$ have direct effects on both the prices charged and the total number of seats sold. In the next section, we shall see that the firm’s optimal long-run decision about capacity is to make the capacity constraint exactly binding, with a level of $MC_{Op}$ that is well below the intersection of the marginal revenue curves. The situation depicted above will occur in the short run only if $MC_{Op}$ rises by a substantial amount.

5. A Price Discriminating Firm’s Optimal Choice of Capacity

In this section, we extend the model to consider the firm’s choice of capacity. A hotel adding rooms and an airline adding planes are both making long-run capacity adjustments. This section shows the firm’s optimal choice of capacity and its implications for the probability that the constraint will be binding in the short run for most firms.

Extending the model to the choice of optimal capacity requires several additional features. Let $MC_{Cap}$ be the marginal cost of expanding capacity by one unit. With variable capacity, the

---

11 The distinction between marginal revenue and price is very important, as $MC_{op}$ is generally low for these firms. Selling 50 seats at $100, rather than 49 at $102, results in marginal revenue of only $2$ on the additional seat.

12 We assume that capacity can be changed by one unit. If capacity is not always available in small units, the optimal capacity may not be feasible. We also abstract from the time dimension, which would require using present discounted values of future costs and revenues.
cost of serving an additional customer becomes the sum of these two costs, or \( MC_{Op} + MC_{Cap} \). Graphically, a change in capacity simply alters the length of the horizontal axis in the model, with the demand and marginal revenue curves staying attached to their respective vertical axes. The optimal capacity choice for the firm will be to expand or contract capacity until this combined marginal cost equals marginal revenue. Because the condition for profit maximization is to equate marginal revenue for both groups, we may write the condition for optimal capacity as solving \( MC_{Op} + MC_{Cap} = MR_B = MR_L \). Thus, optimal capacity is at the minimum level such that the capacity constraint exactly binds. Figure 4 shows a firm expanding its capacity to satisfy this condition.

Note that the decision about the size of capacity involves the marginal cost of capacity, while the allocation decision for a given capacity does not. A firm that has chosen optimal capacity should generally operate in the short run with a situation where \( MC_{Op} \) is substantially below the intersection of the marginal revenue curves (by the amount of \( MC_{Cap} \), as was shown in Figure 2. Thus, changes in the operating cost for a marginal passenger would have to rise by more than the marginal cost of capacity (to a point above the intersection of the marginal revenues) to result in any change in the short-run price decisions of the firm. Only in the long run, through capacity adjustments, will prices fall with lower operating costs. An increase in operating costs will also have no impact on short-run prices as long as the change is less than the marginal cost of capacity (leaving it below the intersection of the marginal revenue curves). Above this point, increases in marginal operating costs would result in the firm’s increasing prices and choosing to have unused capacity.

---

13 Our discussion is consistent with the envelope theorem, because \( MC_{Op} \) is not short-run marginal cost at optimal capacity. It is the additional cost of filling an otherwise empty seat when there is excess capacity. The envelope theorem makes it clear that at optimal capacity the firm incurs both capacity and operating costs to produce a unit of output. See, for example, Varian (1992 pp. 70–1).

14 Consider the mathematical optimization problem for the firm. There would be a profit function maximized subject to a capacity constraint. Setting the derivative of profit with respect to capacity equal to zero would show optimal capacity. Since capacity appears in the constraint, this derivative would be the Lagrangian multiplier, \( \lambda \). Thus, setting \( \lambda = 0 \) by choosing the minimum level of capacity to make the constraint nonbinding solves the optimal capacity problem. Note, however, that with \( \lambda = 0 \) in the long run, the capacity constraint disappears from the equation. Thus, long-run pricing is identical to pricing with no capacity constraints.
6. Conclusions

Yield management involves establishing price categories, prices, and quantities so that low-revenue customers do not take capacity away from high-revenue customers. The essence of this problem is allocational rather than being one of production levels, because the firms operate under capacity constraints. This paper has developed a diagram of price discrimination under capacity constraints that allows for joint determination of price and quantity for each customer group. Our model suggests that a firm’s optimal choice of capacity results in the constraint’s being exactly binding in the short run. Under this optimal choice, capacity is chosen such that it equates marginal revenue with the sum of the marginal operating cost and the marginal cost of capacity. Thus, in the short run, the marginal cost of filling an otherwise empty seat is substantially lower than marginal revenue. At capacity, the true short-run marginal cost of serving one customer group is the forgone revenue from the other group, not the marginal operating cost. As long as the forgone revenue exceeds the marginal operating cost, changes in input prices have no impact on market prices in the short run.

References
