Equilibrium price dispersion under demand uncertainty: the roles of costly capacity and market structure

James D. Dana, Jr.*

When capacity is costly and prices are set in advance, firms facing uncertain demand will sell output at multiple prices and limit the quantity available at each price. I show that the optimal price strategy of a monopolist and the unique pure-strategy Nash equilibria of oligopolists both exhibit intrafirm price dispersion. Moreover, as the market becomes more competitive, prices become more dispersed, a pattern documented in the airline industry. While generating similar predictions, the model differs from the revenue management literature because it disregards market segmentation and fare restrictions that screen customers.

1. Introduction

In markets with costly capacity and uncertain demand, firms will not necessarily set the same price. When firms set their prices before demand is known, they are indifferent between offering units at higher prices with the expectation that these units will sell with a lower probability (when demand is high) and offering units at lower prices with the expectation that these units will sell with higher probability (when demand is high or low). Equilibrium price dispersion with homogeneous goods and perfect competition was first described by Prescott (1975) in a model of hotel competition and was developed more formally by Eden (1990). Prescott’s model offers an explanation for interfirm (and intrafirm) price dispersion in industries such as airlines, automobile rentals, hotels, restaurants, theaters, and sporting events. Furthermore, it has also been the basis of recent applied research on the real effects of monetary shocks (Lucas and Woodford, 1993; Eden, 1994), wage rigidities (Weitzman, 1989), resale price maintenance (Deneckere, Marvel, and Peck, 1996), stochastic peak-load pricing (Dana, 1999b), and advance purchase discounts (Dana, 1998).1

This article is the first to extend Prescott’s model to monopoly and imperfect competition. By expanding firms’ strategy sets to include price distributions, i.e., sets

---

* Northwestern University; j-dana@nwu.edu.

I would like to thank Severin Borenstein, Robert Gertner, Tom Holmes, Carsten Kowalczyk, Boaz Moselle, Alejandro Manelli, Ed Olimstead, Steve Postrel, Kathryn Spier, Olaf Staffans, and Sang-Seung Yi for helpful comments. I would also like to thank the Editor, Kyle Bagwell, and several anonymous referees. Financial support from the Rockefeller Grant in Economics at Dartmouth College is gratefully acknowledged.

1 See also Eden (forthcoming) for a broader discussion of applications of the model to monetary economics.
of prices and quantity limits at each price, I show that there exists a unique pure-strategy equilibrium in price distributions even when no pure-strategy equilibrium exists in prices (see Bryant, 1980). In other words, the model predicts an equilibrium with intrafirm price dispersion in which each firm offers its output at multiple prices (as opposed to random prices). The oligopoly equilibrium is symmetric and the market price distribution converges to Prescott’s competitive equilibrium as the number of firms approaches infinity. As competition increases, the average price level falls and the degree of price dispersion increases. This result arises for the same reasons that a monopolist typically does not raise its price by the entire amount of an increase in costs, while competitive firms pass through all of their cost increases.

This inverse correlation between price dispersion and market concentration has been observed in the airline industry. In particular, Borenstein and Rose (1994) showed empirically that price dispersion is greater on city-pair routes that are served by a larger number of carriers. Borenstein and Rose attribute this result to price discrimination and argue that point using a monopolistic-competition model with certain demand. However, their empirical results are also consistent with this article’s theory that price dispersion is due to capacity costs (i.e., perishable assets) and demand uncertainty. Furthermore, this model is consistent with other characteristics of the airline industry, including capacity utilization rates (load factors) that are higher for airlines that charge lower fares. For example, in 1993 the major airlines whose yields (prices) were below average had an average load factor (capacity utilization) of 64.25% and a yield of 13.80 cents, while the major airlines whose yields were above average had an average load factor of 59.92% and a yield of 16.40 cents. This is the natural prediction of a model with costly and perishable capacity, price rigidities, and demand uncertainty and is difficult to explain in models that do not have each of these assumptions.

Yield management, now commonly called revenue management, is the use of market segmentation and seat-inventory control (i.e., assigning limits on the availability of seats at each fare) to maximize firm revenues (or profits). Sophisticated revenue management systems are now used by airlines, hotels, car rental companies, commercial shippers, and increasingly in other industries to manage demand uncertainty and to ensure that their products and services are available even when demand is high. Because of customer segmenting restrictions (or “fencing”) such as advanced-purchase discounts, Saturday-night stayover requirements, nonrefundable purchases, and volume discounts, revenue management is usually considered to be a form of third-degree price discrimination. However, in a model in which market segmentation and price discrimination are not feasible, I show that demand uncertainty and the perishable nature of the assets are alone sufficient to explain intrafirm price dispersion and may help explain

2 They consider a variant of Borenstein’s (1985) model where consumers have heterogeneous “travel” (waiting) costs in a circle location model and product variety (departure times) is fixed while they vary the market structure. However, most models of oligopoly price discrimination, including Borenstein (1985), Holmes (1989), and Gale (1993), do not predict a positive correlation between price dispersion and market structure.

3 These two theoretical explanations are not inconsistent with one another. In fact, if the existence of price dispersion for other reasons facilitates airlines’ use of discriminatory restrictions, then they may be complementary.

4 These calculations are based on data from Department of Transportation Form 441. Yield is defined as total operating revenue divided by the number of passenger miles flown by paying customers. Load factor is defined as the number of passenger miles flown by paying customers divided by the number of available seat miles.

5 An alternative explanation of this result is that airlines with greater market power (and hence higher prices) have excess capacity in order to deter entry.
why price dispersion increases with competition. In the next section I describe more carefully the relationship of the model to airline revenue management practices and research.

Another application of the monopoly model presented here is theater and stadium ticket pricing. Although in these examples firms’ products are not homogeneous, ticket sellers face the same tradeoff between selling the marginal unit at a low price for sure and at a high price with uncertainty. Since prices are clearly set in advance, theaters and stadiums reserve some seats at higher prices (or some lower-quality seats at higher margins) to meet the uncertain demand.

The model also applies to retail pricing and inventories. Loss-leader models typically explain low prices as a “bait-and-switch” tactic, but instead one might think of advertised low-priced products as items priced at (or near) cost for which firms then have no incentive to hold speculative inventory. When these items stock out, customers are left to buy similar goods at higher prices, but only because realized demand was high and not because the retailer intended to ration customers that he could profitably serve at the low price.

An important assumption of the model is that prices are rigid; firms must precommit to a menu or schedule of prices before demand is known. This commitment may arise from advertisements or promotions, or because the firm must physically print tickets in advance. Alternatively, it may reflect a strategic commitment on the part of firms not to invest in technologies to make prices more responsive to demand. Regardless of the reason, price commitments enforce a price system in which customers who either arrive early or arrive late when demand is low can buy at lower prices, while customers who arrive late when demand is high pay a higher price. An individual consumer at the time of purchase may only observe the price of each firm’s least-expensive remaining unit, but \textit{ex post} consumers will have paid different prices solely because of the random order in which they were served.

Complete rigidity is much more than I need to assume. It is enough to assume that firms do not immediately change their prices when they receive a new, relevant piece of demand information. In the airline industry, useful information about final demand might be contained only in sales that are very close to departure time after it is too late to alter prices, and the information might not be revealed until sales are finalized. Allowing firms to change their prices in response to changes in competitors’ prices, past sales, market research, or advertising campaigns, or to change the prices for later flights based on information about past flights’ sales, would not alter the results. Note also that if firms learn no additional useful information about the demand state during the trading process, then the price rigidity assumption can be relaxed (see Eden, 1990 and forthcoming).\footnote{This means that during the trading process, the conditional probability of any feasible demand state must always be equal to the \textit{ex ante} probability that the state occurs divided by the \textit{ex ante} probability that any of the remaining feasible states occur.}

The article adds to earlier work on oligopoly models with demand uncertainty. Klemperer and Meyer (1986, 1989) consider models with some market clearing. The latter is unusual because firms’ strategy sets consist of supply functions that restrict the relationship between price and quantity, but these functions are quite different from the price distributions described here. Deneckere and Peck (1995) consider a more closely related model; however, they assume that customers can visit only one firm and do not allow intrafirm price dispersion. Firms in their model compete in both price and availability (see also Carlton, 1978; Peters, 1984; and Dana, 1999a). Schmidt...
(1991) looks at a model in which demand uncertainty and heterogeneous costs (as opposed to sunk capacity costs) lead to price dispersion.

Other monopoly models that predict intrafirm price dispersion are Wilson (1988), Salop (1977), and Dana (1992). Wilson shows that with an increasing marginal cost function and a marginal revenue function that is not everywhere nonincreasing, a monopolist may charge at most two prices and ration the availability of the lower-priced units (his rationing process is the same as that considered here). Salop shows that a monopolist can open multiple sales outlets and offer different prices at each in order to price discriminate against consumers with high search costs. Dana considers a different mechanism by which price rigidities might cause price dispersion. When the monopolist’s optimal price changes with the realization of demand (because the elasticity decreases with realized demand), then the firm might charge more for units that sell only when demand is high.7

Many other oligopoly and competitive models predict price dispersion as a result of randomization. While some can be reinterpreted to predict intrafirm price dispersion, none predict price dispersion as the unique pure-strategy equilibrium of an oligopoly model.8 Of particular interest are articles by Varian (1980), Butters (1977), Burdett and Judd (1983), and related articles by Stahl (1989) and Rosenthal (1980). Stahl and Rosenthal find that prices increase and price dispersion decreases as the market becomes more competitive, in contrast to the result obtained here. The difference is that price dispersion in their models is driven by search and imperfect information (or similar market imperfections) rather than capacity costs and demand uncertainty.

Section 2 reviews the practice of revenue management and its relationship to the model. Section 3 presents a simple two-state example to illustrate the contributions of the article. Section 4 describes the basic model, and Section 5 derives the equilibrium pricing under the different market structures of perfect competition, monopoly, and oligopoly. Section 6 examines the relationship between price dispersion and market structure, and concluding remarks are offered in Section 7.

2. Revenue management

The term “revenue management” is commonly used to describe most aspects of airlines’ pricing and seat-inventory control decisions; but in reality, revenue managers primarily practice seat-inventory control.9 Formally, revenue management describes a process of setting fares for each route (origin and destination pair) and each set of restrictions (nonstop, time-of-day, day-of-week, refundable, advance purchase, first class or coach, and Saturday-night stayover) and limiting the number of seats available at each fare. In the language of economics, revenue management increases airlines’ profits in three ways. First, it implements peak-load pricing. Second, it implements third-degree price discrimination. That is, fare restrictions screen customers and segment them by their sensitivity to price and potentially by their demand uncertainty.10 And third, it implements an inventory control system for coping with uncertain demand for a perishable asset.

---

7 In this article the elasticity of demand is assumed to depend only on price and not the realization of the demand state.
8 However, an oligopoly version of Salop’s (1977) model with differentiated products would.
9 See Weatherford and Bodily (1992) for an overview of research on revenue management.
10 Dana (1998) shows that competitive firms will use advance-purchase price discounts to screen for customers with more certain demands (because they contribute less to aggregate demand uncertainty), and Courty and Li (1998) show that a monopolist may use control over refunds to screen for customers with more certain demands.
Like most of the literature on seat-inventory control, this article addresses only this last role of revenue management. Peak-load pricing and price discrimination are, after all, primarily pricing problems. If demand were known, then prices alone could be used to accomplish these objectives and there would be no reason to use seat-inventory control. I focus instead on how firms can profit from offering multiple fares and limiting the availability of seats at low fares when they cannot segment the market. In contrast to the traditional operations literature on revenue management, I show that market segmentation is not necessary for revenue management to increase profits. Demand uncertainty, costly capacity (perishable assets), and price rigidities are sufficient to explain the use of seat-inventory control.

This article differs from most research on revenue management (seat-inventory control) in several ways. First, I simultaneously determine prices and inventory levels; the revenue management literature treats fares as exogenous. Second, I explicitly consider strategic interactions between firms, while the revenue management literature typically considers single-firm models. Third, because I ignore market segmentation, I am allowing consumers’ demands to be “diverted” to the next-highest fare when the lower fare stocks out; high-willingness-to-pay consumers first try to buy at low prices, but when low fares are no longer available, they buy at higher prices. An acknowledged shortcoming of the revenue management literature is that it typically treats market segments as independent, both in the sense that demand is not diverted and in the sense that demands are stochastically independent. But since airlines do in fact segment markets, the advantages here of studying a seat-inventory control model with endogenous pricing, strategic interaction, and diversion must be weighed against the cost of ignoring market segmentation.

There are other potential problems in applying this model to the airline industry. I assume that firms irrevocably commit themselves to distributions of prices and seat inventories for each flight before learning demand. At many (but not all) airlines, revenue managers use modern computer technology to calculate and adjust seat-inventory allocations on a daily basis. Price changes (while much less frequent) can also be made quickly. To understand the importance of these issues for the model, we have to look beyond a theoretical discussion of revenue management.

In practice, revenue management has traditionally described several separate functions, if not separate departments, within an airline’s organization. First is the collection of historical (and contemporary) sales data used to generate forecasts. Second is fare setting. That group determines the restrictions that passengers must meet and the prices that tickets will sell for. In practice, fares apply to many flights, and any limits on departure dates or times are specified as restrictions on the fare. These departments closely monitor competitors’ fares on computer reservation systems and quickly match any of their price changes. Third is a potentially computerized system that determines,
given demand forecasts and fares, the optimal limit on the number of seats sold at each fare and then transmits that information to a computer reservation system.

Although now more than ten years old, Belobaba (1987) reports the results of a survey of airlines’ revenue management (seat-inventory control) practices in late 1985. This is the only systematic look at airline practices of which I am aware, and more important, it was done only months before the price data analyzed by Borenstein and Rose was generated. In 1985, many airlines were just beginning to store and manage the enormous amount of historical sales data generated by reservations systems. Before 1985, forecasting was so poor that some airlines were still using a common percentage booking limit for discount seats on all their flights, regardless of the route or season. In 1985, all revenue management systems, whether computerized or not, chose seat inventories for each fare category using static optimizing models, as if they were making a once-and-for-all decision. Today some computerized revenue management systems dynamically optimize booking limits before a flight departs using updated demand forecasts that incorporate information about actual bookings. However, in 1985, static booking limits could be changed only by an experienced revenue manager, using his or her judgment, whenever the computer reservation system, which automatically monitored bookings, brought to his or her attention a booking limit that was about to be reached. Thus booking limits were often relaxed (presumably whenever early bookings at the full fare were below projections) but rarely tightened in response to actual bookings.

While some airlines combine pricing and seat-inventory control in one department, initiatives to change prices are usually made by marketing personnel, often with surprisingly little information. In many cases, pricing decisions do not even make use of historical sales information. Even if prices are optimal given the firm’s best forecasts of demand (as I assume in my model), pricing departments have no systematic method for using actual bookings to revise prices for the remaining seats on that flight. In 1985, seat-inventory control relied on the assumption that prices would not change during the final six weeks before a flight departure, the period in which the bulk of airline reservations are made (Belobaba, 1987).

The subsequent development and adoption of better tools for demand forecasting and computerized dynamic seat-inventory control has no doubt changed airline competition significantly. However, the one-shot selection of prices and quantities early in the history of revenue management does seem to closely mirror the pricing assumption in the model presented here. It is not clear whether airlines today are currently trying to systematically make pricing decisions a function of actual bookings, but there is certainly little evidence that they did so in the mid-1980s.

Though the model has obvious limits, it is nevertheless consistent with stylized facts about airlines. In particular, capacity utilization rates are higher for seat-inventory allocations of low-fare seats. While information about capacity utilization by fare within an airline is proprietary (because seat-inventory control information is proprietary), this is a direct consequence of the algorithms used by their revenue management computers. Also, as mentioned earlier, airlines with lower fares tend to have higher load factors.

---

15 In early 1985, American Airlines introduced the first modern computerized revenue management system that combined demand forecasting and optimizing algorithms for seat-inventory control. That system has been credited with allowing American Airlines to profitably match People’s Express fares across the board, leading to the rapid demise of that airline (which lacked any information technology for forecasting demand or practicing revenue management). In 1985, industry analysts clearly did not understand the potential of American Airlines’ new system.
3. A two-demand-state example

The intuition for much of the article can be seen in the following numerical example with two demand states. Suppose that consumers’ reservation values for homogeneous stadium seats are uniformly distributed on [0, 100] and that either 100 or 400 consumers will demand seats (with equal probability). So the inverse demand is

\[ P_L(Q_L) = 100 - Q_L \]

if demand is low and

\[ P_H(Q_H) = 100 - 0.25Q_H \]

if demand is high.

Also suppose the marginal cost of offering a seat in the stadium is $20. (This may be interpreted as either the marginal cost of stocking inventory or as the shadow cost of capacity.) Firms are required to print tickets with assigned prices before learning demand. Consumers arrive in random order and choose among the unsold tickets, but since the good is homogeneous, consumers will always prefer the cheapest remaining ticket.

First suppose that stadium seats are provided by a perfectly competitive market (Prescott’s (1975) model). Specifically, imagine that each firm offers (and prices) a small number of the stadium’s seats, taking the aggregate distribution of prices and quantities as given. In the unique competitive equilibrium, 80 tickets are offered at $20 and 180 tickets are offered at $40. In the low-demand state, exactly 80 consumers are willing to spend $20 per ticket and all 80 are served. (No remaining consumer is willing to spend $40, so the expensive tickets go unsold.) In the high-demand state, 320 consumers are willing to pay $20 per ticket but only 80 of them are lucky enough to get a $20 seat. Of the remaining 260 consumers, only 180 are willing to pay $40 for a seat (this is because of the proportional rationing assumption) and all of these consumers are served. So 80 tickets at $20 each are purchased whether demand is high or low, while 180 tickets at $40 each are purchased only if demand is high. The expected revenue on tickets at either price is $20, which is equal to the marginal cost of providing a seat.

It may seem odd at first that the unique competitive equilibrium has dispersed prices. Why doesn’t a uniform price equilibrium exist? To see why, suppose that all of the seats are offered at a uniform price of $30 each. The zero-profit condition implies that the \textit{ex ante} probability of sale of each seat is \( \frac{2}{3} \), so in equilibrium firms offer 210 seats (because \( \frac{1}{2}(\frac{70}{210}) + \frac{1}{2}(\frac{140}{210}) = \frac{2}{3} \)). But this cannot be an equilibrium, because a firm could increase its profit either by offering a seat for $29 or by offering one for $50. A $29 seat would be the first seat to sell in either demand state, would be assured of selling, and would generate a profit of $9. A $50 seat would sell in the high-demand state (there is strictly positive residual demand because the demand at $30 is 280, which exceeds the 210 seats available) but not in the low-demand state, so it would earn an expected profit of $5. Similar logic establishes that no competitive equilibrium exists at any other uniform price either.

The competitive equilibrium is shown graphically in Figure 1. In our discussion above, the expected revenue on each seat is equal to $20, the marginal cost. Equivalently, we can think of firms as setting the price of a ticket equal to the marginal cost of a seat divided by the probability of sale. The graph in Figure 1 depicts the competitive equilibrium by doubling the marginal cost curve rather than halving the demand curve. To see that this is a competitive equilibrium, notice that expected revenue is less than or equal to marginal cost at every price (and exactly equal to marginal cost at $20 and $40 where firms’ supply is positive). At prices below $20, profit is clearly negative; at prices between $20 and $40, the expected revenue is less than $20 because the probability of sale is still one-half; and at prices greater than $40, expected revenue is zero.
If the stadium seats are provided by a monopolist, then 40 tickets will be offered at a price of $60 and 90 tickets at a price of $70. To see why, imagine that the monopolist does indeed print 40 tickets with a price of $60. This would be the optimal strategy if demand were $P_L(Q_L) = 100 - 2Q_L$ with certainty because $MR = 100 - 2Q_L = 20 = MC$, as shown in Figure 2. Taking these 40 “cheap” seats as given, it makes sense for the monopolist to offer additional seating; for example, one more seat with a price of $90 would sell half the time, generating an extra expected profit of $25. The question is, how many additional seats should the monopolist offer, and at what price?

To answer this question, it is necessary to derive the residual demand curve when demand is high. Since there are 40 tickets at $60 each for sale, when demand is high only $\frac{1}{4}$ of the 160 customers who are willing to pay $60 are lucky enough to get a cheap ticket. As shown in Figure 3, the residual demand curve is $RD_H(P_H) = (\frac{3}{4})(400 - P_H)$ for prices above $60$. The monopolist’s rule is to set (expected) marginal revenue equal to marginal cost. But since high demand occurs only half the time, the expected marginal revenue is half what it would be if the demand were certain. As in the competitive case, Figure 3 graphically depicts the effect of demand uncertainty by doubling the firm’s marginal cost rather than halving marginal revenue. The monopolist maximizes profits by offering an additional 90 tickets at a price of $70 per seat.

This is, in fact, the optimal pricing strategy for the monopolist. The monopolist does not want to ration the availability of its tickets in any other way, nor adjust the price of the low-priced seats (which would shift the residual demand when demand is high), nor offer price distributions with more than two prices. The monopolist sets

---

16 Equivalently, if the monopolist has a capacity of 130 seats, then its shadow cost of capacity is $20 and its price strategy is the same.

17 The inverse residual demand is $P_H(Q_H) = 100 - (\frac{3}{4})Q_H$. For prices below $60$, the residual demand curve corresponds to the original demand curve because consumers would choose to buy any less-expensive units first.

18 $E[MR] = (\frac{3}{4})(100 - (\frac{3}{4})Q_H) = 20 = MC$.

19 While the results of the article are stated for continuous distributions of demand uncertainty, this distribution of prices is the limit of the optimal price distribution when the distribution of demand uncertainty converges to the discrete distribution.

© RAND 1999.
prices as if there were two separate segments, some consumers who always show up (those who buy in the low-demand state and an equal number who buy in the high-demand state) and some who may or may not show up (the residual demand in the high-demand state). If the monopolist could offer only a single price, it would charge $66 and earn an expected profit of $2,880 versus a profit of $3,000 with the two-price strategy.

This example helps illustrate my main points. First, nonrandom price dispersion is predicted in a simple model with uncertain market demand, costly or constrained capacity, and price commitments. In this example, price dispersion arises for both perfect competition and monopoly, and later in the article I show that price dispersion
arises for oligopolies, at least when there is sufficient demand uncertainty and high-enough capacity costs. The reasoning is straightforward: firms require a higher marginal revenue to hold inventories of seats that are expected to sell with lower probability.

Second, the model predicts *intrafirm* price dispersion. Although the competitive equilibrium may be interpreted as exhibiting either interfirm price dispersion or intrafirm price dispersion, the monopoly outcome exhibits only intrafirm price dispersion, and later in the article I show that the unique oligopoly outcome exhibits intrafirm price dispersion as well.

Finally, the example illustrates that the degree of price dispersion is greater in more competitive market structures, as we can see here by comparing the monopoly and competitive prices. Demand uncertainty reduces the expected marginal revenue of a sale by a fraction equal to the probability that the good will go unsold, so setting expected marginal revenue equal to marginal cost is analogous to setting marginal revenue equal to marginal cost divided by the probability of sale. As the probability of sale drops, prices rise to reflect firms’ lower expected marginal revenue or equivalently higher marginal costs. But because a monopolist usually absorbs part of its cost increases, while competitive firms do not, the monopolist prices are compressed, or less dispersed, relative to the competitive ones. The markup of price over the break-even price is decreasing in the break-even price, so the monopoly prices are less dispersed than the break-even (or competitive) prices.\(^2\)

Another useful benchmark for comparison is the market-clearing price. If prices cleared the market, competitive firms would offer 240 seats and sell them for $0 when demand was low and $40 when demand was high. The monopolist would offer 120 seats and sell them for $50 in the low-demand state and $70 in the high-demand state. Note that in this example, because the uncertainty about demand is so large, the monopolist does not sell all its capacity in the low-demand state, but this is not important for the inferences being drawn from the example. Clearly prices are less volatile for a monopolist than for competitive firms. The reason is the same. A monopolist facing linear demand absorbs part of the increase in its cost (in this case, the shadow cost of capacity is increasing from $0 to $40). But while prices here are volatile, there is no price dispersion when prices clear the market; in each demand state all consumers pay the same price.

Before turning to the continuous-demand uncertainty model, it is useful to see why a pure-strategy equilibrium does not exist for the case of imperfect competition in this discrete example. Imagine a duopoly in which each firm is endowed with 65 seats, so the industry capacity is 130 seats, the same as the monopolist’s equilibrium choice of capacity. This is clearly not the capacity choice that duopolists would choose in equilibrium, but it illustrates why no pure-strategy equilibrium exists. One might expect that the equilibrium of the duopoly game is where each firm offers 20 tickets at $60 and 45 tickets at $70, the same distribution as a monopolist. But this is not an equilibrium. Either firm could lower the price of all 20 of its tickets at $60 and of 20 of its 45 tickets at $70 to a new price of $60 \(\pm \epsilon\). At that price it would capture all the sales in the low-demand state and still sell all its remaining 25 tickets at $70 in the high-demand state. This increases the firm’s profits dramatically because the expected revenue on a $60 seat is much higher than on a $70 seat ($60 versus $35 because the

\(^2\) Here demand is linear, so a monopolist does not pass through all its cost increases. In general, a monopolist will absorb part of a cost increase as long as the elasticity of demand increases sufficiently fast with price. Later in the article, I assume that the demand curve has a finite price intercept that is sufficient to guarantee that the upper bound of prices is the same for all market structures; this also implies that the elasticity of demand increases with price (in a neighborhood of the intercept).

© RAND 1999.
lower-priced tickets are sold twice as often). The price cut increases profits at the expense of the other firm (consumers are better off) by causing the competitor’s $60 seats to be sold only in the high-demand state. It is not hard to show that no pure-strategy equilibrium of this game exists.

4. The model

- In this section I describe the basic model. Demand is parameterized by price, \( p \), and a random state variable, \( e \), drawn from a continuously differentiable cumulative probability distribution, \( F(e) \), with strictly positive and bounded probability density, \( f(e) \), on a support \([e_l, e_r] \). Continuity of \( F(e) \) is necessary only for the oligopoly model. For the competition and monopoly models I also briefly discuss the conceptually simpler case where \( e \) has a discrete distribution with possible values \( \{e_1, e_2, \ldots, e_n\} \) and associated probabilities \( \{f_1, f_2, \ldots, f_n\} \).

For simplicity, I consider the case in which demand is multiplicatively separable, e.g., \( D(p, e) = g(e)D(p) \), which without loss of generality can be written \( D(p, e) = eD(p) \).\(^{21}\) I also assume the inverse demand function has a finite intercept, \( \bar{p} = \inf\{p | D(p) = 0\} \), which is clearly the same for all \( e \). Without this assumption the model would predict infinite prices.

I assume that \( D(p) \) is twice continuously differentiable for all \( p \in [0, \bar{p}] \). Furthermore, there are no income effects, and for all \( p \in [0, \bar{p}] \), \( D_p(p) < 0 \) and \( D_{pp}D - 2(D_p)^2 < 0 \), where the derivatives at \( \bar{p} \) are defined appropriately using limits. The second condition is identical to the standard assumption that the marginal revenue function is strictly decreasing in output. The elasticity of demand, \( \eta(p) = -pD_p(p)/D(p) \), is independent of \( e \) because demand is multiplicatively separable. Let \( p^\alpha(c) \) denote the monopoly price for a firm with known demand, \( eD(p) \), and constant marginal cost, \( c \), and note that both \( p^\alpha(c) \) and the marginal revenue function are also independent of \( e \).

I assume that demand uncertainty has full support, by which I mean that demand is zero in the lowest demand state, or \( D(c, e) = 0 \). For multiplicatively separable demand this clearly implies \( e = 0 \). This assumption can be relaxed for the monopoly or perfect competition cases. But in the oligopoly case this assumption, like the assumption that \( F(e) \) is continuous, eliminates mass points in the first-order conditions and is necessary for the existence of a pure-strategy equilibrium.

Production exhibits constant returns to scale. Firms have two costs: a strictly positive unit cost (or shadow cost) of capacity, \( \lambda \), incurred on all units, and a unit marginal cost of production, \( c \), incurred only on the units that are sold. I also assume \( \bar{p} > \lambda + c \).

Firms’ strategies are distributions of prices and quantities.\(^{22}\) The most general way to represent the set of such distributions would be the set of nonnegative measures on the positive real numbers, where the measure over the entire range of prices would equal the firm’s total output or capacity. I have considered only those distributions that can be represented by a Lebesgue integrable density, \( q(p) \), that is strictly positive on some finite domain, \( [p, \bar{p}] \), and zero everywhere else, and a finite number of mass points. To simplify the exposition, however, I maintain throughout the assumption that

\(^{21}\) The class of multiplicatively separable demand functions is a special case of a more general set of demand functions, \( D(p, e) \), which are continuous, nonincreasing in \( p \) (strictly decreasing when \( D(p, e) > 0 \)), and nondecreasing in \( e \) (strictly increasing when \( D(p, e) > 0 \)); however, no closed-form solution exists for the general class of demand functions (see Dana, 1996).

\(^{22}\) The assumption that prices and quantities do not depend on the state of demand, either directly or indirectly through the pattern of sales, is implicit. Other forward contracts, such as priority service pricing contracts considered by Harris and Raviv (1981), Chao and Wilson (1987), Wilson (1989), and Spulber (1992), are also ruled out.

© RAND 1999.
there are no mass points. Let $q(p) = \sum_i q_i(p)$ denote the market price distribution, equal to the sum of the individual firms’ price distributions, and let $Q(p) = \sum_i Q_i(p)$ denote the cumulative market price distribution where $Q(p) = \int_p^\infty q_i(p) \, dp$. In the propositions below a price distribution is said to be unique when the cumulative market price distribution is the same for all equilibrium values of $q(p)$.

Because prices do not adjust to clear the market, I must specify how goods will be rationed in equilibrium. I imagine that output is sold sequentially in order of price, with the cheapest units going to the consumers who buy first, so the rationing rule determines the order in which consumers’ demands are met. Note that the rationing rule describes how aggregate demand is rationed and does not depend on individual firms’ price distributions; to describe the rationing process it is sufficient to consider Prescott’s model where infinitesimal firms each charge a single, heterogeneous price. The two most commonly used rationing rules are proportional rationing and parallel rationing. I assume proportional rationing that is consistent with homogeneous consumers with elastic demands queuing in random order and purchasing according to their demand at the available prices. It is also consistent with heterogeneous consumers with unit demands queuing in random order and purchasing as long as their valuation exceeds the price of the available units.

The residual demand is the quantity demanded at price $p$ given that all the available units with prices less than $p$ have already been sold. Under proportional rationing, the residual demand curve for $p \geq \bar{p}$ in state $e$, given $D(\bar{p}, e) > 0$, is $\max\{R(p, e; \bar{p}, q), 0\}$, where

$$RD(p, e; \bar{p}, q) = D(p, e) - \int_p^{\bar{p}} \frac{D(p, e)}{D(r, e)} q(r) \, dr. \quad (1)$$

Residual demand is further defined to be zero when $D(\bar{p}, e) = 0$.

The intuition for equation (1) is easiest to see in the case of a discrete price distribution. Suppose there are $Q$ units sold at a price $\bar{p}$. If $Q > D(\bar{p}, e)$, then there is excess supply in state $e$, and the residual demand is zero at any price $p \geq \bar{p}$. However, if $Q < D(\bar{p}, e)$, then the residual demand at any price $p \geq \bar{p}$ is $D(p, e) - Q / D(\bar{p}, e)$, or equivalently $D(p, e) / D(\bar{p}, e) - Q / D(\bar{p}, e)$. The probability that a customer who is willing to pay $\bar{p}$ actually receives one of the $Q$ available units is $Q / D(\bar{p}, e)$, and the fraction of the customers willing to pay at least $\bar{p}$ whose demand is unsatisfied at price $\bar{p}$ is $1 - Q / D(\bar{p}, e)$. Note that this probability is independent of $p$, so $1 - Q / D(\bar{p}, e)$ is also the fraction of the customers willing to pay at least $p$ whose demand is unsatisfied.

Equation (1) defines the analogous residual demand function for more general price distributions, $q(p)$. The fraction of the demand that can be satisfied by a single unit sold at price $r$ is $1 / D(r, e)$. Since we are interested in the residual demand at price $p$, $D(p, e) / D(r, e)$ is the fraction of the unit (or units) priced at $r$ that are sold to consumers who would have been willing to pay the price $p$. Note that because proportional rationing does not change the relative proportion of consumers with different demands, this proportion can be calculated using the original demand curve instead of the residual demand curve (at price $p$ given that all of the units below $r$ have already been sold).

---

23 The more general proofs are in an earlier version of the article, available from the author.

24 The terms proportional and parallel rationing take their names from the effect that each rule has on the shape of the residual demand curve (see Beckman, 1965; Davidson and Deneckere, 1986). Parallel rationing is sometimes called efficient rationing because it maximizes consumer surplus, but parallel rationing is actually very inefficient in this context. Because prices are dispersed, the rationing rule affects which customers are rationed not only when there is excess demand but also when there is excess supply.
The residual demand function determines the number of units sold in each state. Suppose that in state \( e \) all the units priced below \( p \) have already been sold. If \( RD(p, e; p, q) = 0 \), then there are still consumers who are willing to pay \( p \) or more for the good. I define the “market-clearing” price, \( \rho(e, q) \), as the price of the highest-priced unit that sells in state \( e \), although markets do not clear in the traditional sense. So \( RD(\rho(e, q), e; \rho(e, q), q) = 0 \) implicitly defines the market-clearing price \( \rho(e, q) \) as a function of \( e \) given the price distribution \( q \).

From (1), using the fact that demand is multiplicatively separable, we have

\[
RD(p, e; p, q) = eD(p) \left[ 1 - \int_{\rho}^p \frac{q(r)}{eD(r)} \, dr \right].
\]

So, \( RD(p, e; p, q) = 0 \) implies that either \( eD(p) = 0 \) or \( 1 - \int_{\rho}^p \frac{q(r)}{eD(r)} \, dr = 0 \).
Clearly, \( eD(p) = 0 \) only if \( p \geq \overline{p} \) or \( e = 0 \). Hence, for all \( e > 0 \), either the market-clearing price \( \rho(e, q) \) is defined implicitly by \( 1 - \int_{\rho}^p \frac{q(r)}{eD(r)} \, dr = 0 \) or all the available units sell and there is excess demand at \( \overline{p} \).

I define \( \epsilon(p, q) \) as the minimum demand state in which all the units offered at prices \( p \) and below are sold, also using \( RD(p, e; p, q) = 0 \). So \( \epsilon(p, q) \) is uniquely defined for all \( p \in [p, \overline{p}] \) by

\[
\epsilon(p, q) = \int_{\rho}^p \frac{q(r)}{D(r)} \, dr
\]

and is continuous and weakly increasing.

Equation (3) can also be used to define \( \rho(e, q) \) for all \( e \leq \epsilon(\overline{p}, q) \). Without loss of generality, \( \overline{p} \leq \overline{p} \), so I define \( \rho(e, q) = \overline{p} \) for all \( e > \epsilon(\overline{p}, q) \), since any unit priced at \( \overline{p} \) or less would be sold in those states. So \( \rho(e, q) \) is uniquely defined for all \( e \) in \([0, \overline{p}]\) and is strictly increasing on \([0, \epsilon(\overline{p}, q)]\). Note that if \( q(p) = 0 \) on some subset \([p_1, p_2]\) of \((p, \overline{p})\), then \( \epsilon(p, q) \) is constant on \([p_1, p_2]\) and \( \rho(e, q) \) is discontinuous at \( \epsilon(p_1, q) \). However, we shall see that in equilibrium \( q(p) \) is strictly positive on \((p, \overline{p})\).

5. Equilibrium pricing

- Perfect competition. Imagine that competitive firms choose how much to sell and at what price (or prices), and that consumers, arriving in random order, costlessly observe the lowest prices still available from each firm and buy from the firm (or firms) offering the lowest prices. Since the market does not clear in the traditional sense, it is necessary to begin with a definition of a competitive equilibrium in this context.

Definition. A market price distribution \( q^*(p) \), with an associated probability of sale function \( y^*(p) \), is a competitive equilibrium if (1) there exist functions \( q^i \) such that (i) \( \forall i, q^i \) maximizes firm \( i \)'s profit given \( y^*(p) \) and (ii) \( q^* = \sum_i q^i \), and (2) \( y^*(p) \) is equal to the probability that any unit of output offered at a price \( p \) will sell given the market price distribution \( q^*(p) \).

In the standard Walrasian model of perfect competition, we assume that each firm chooses output to maximize its profit, taking the equilibrium price as given. Here, since there are many prices and quantities, I have generalized the Walrasian assumption. I assume that firms believe they are unable to affect the terms of trade facing other firms,
that is, the set of possible prices and the associated probabilities of sale. This is equivalent to assuming that each firm chooses its price distribution believing that it has no effect on the aggregate price distribution.

In a competitive equilibrium, the profit from offering a single unit must be nonpositive at each price. Otherwise, because of the constant-returns-to-scale assumption, any firm could increase its profits by offering additional units at that price. The probability of sale for a unit at price \( p \) given the market price distribution \( q(p) \) is \( y(p, q) = 1 - F(e(p, q)) \), so a firm that offers units at price \( p \) will earn an expected profit \( (p - c) y(p, q) - \lambda \) per unit of capacity. In a competitive equilibrium, \( (p - c) y(p, q) - \lambda = 0 \) for all \( p \in [\underline{p}, \bar{p}] \), and \( (p - c) y(p, q) - \lambda \leq 0 \) for all \( p \notin [\underline{p}, \bar{p}] \), so

\[
p = c + \frac{\lambda}{1 - F(e(p, q))}
\]  

(4)

for all \( p \in [\underline{p}, \bar{p}] \), and \( p \leq c + \lambda [1 - F(e(p, q))] \) for all \( p \in [\underline{p}, \bar{p}] \). \(^{26}\)

Any price that is offered in equilibrium must be equal to marginal cost plus the unit cost of capacity divided by the probability that a unit offered at that price will be sold. The latter term can be interpreted as an “effective” cost of capacity, since it is equal to the revenue (net of marginal cost) that the firm must earn if the good is sold in order to cover the unit capacity cost it incurs whether or not the good is sold. Of course, this definition of cost is somewhat ad hoc, since it implies that cost is endogenous and depends on the price chosen by the firm. An alternative interpretation of the zero-profit condition is simply that the expected net revenue per unit of capacity, \( (p - c) y(p, q) \), must equal the marginal cost of capacity for all prices. In the competitive equilibrium, as in other models of price dispersion (such as Butters (1977) and Burdett and Judd (1983)), firms are indifferent between setting a higher price with a smaller probability of sale and setting a lower price with a higher probability of sale.

If instead of having a direct cost of capacity, \( \lambda \), firms had a binding industry capacity constraint of \( K \), then the competitive equilibrium would still be characterized by (4); however, in this case \( \lambda \) would denote the shadow cost of capacity and be endogenously determined. If the capacity constraint is equal to the level of capacity that would have been chosen if the direct cost of capacity were \( \lambda \), then the shadow cost of capacity is \( \lambda \), and the equilibrium price distributions are the same. This implies that the equilibrium price distribution is the same whether firms choose quantity and price sequentially or simultaneously.

**Proposition 1 (Prescott).** If capacity is either costly or constrained, the competitive equilibrium price distribution is strictly dispersed and is uniquely defined by

\[
q_\lambda (p) = D(p) \frac{\lambda}{(p - c)^2} \left[ F^{-1} \left( 1 - \frac{\lambda}{p - c} \right) \right]
\]  

(5)

on the support \([ c + \lambda, \bar{p}] \).

---

\(^{25}\) This clearly demonstrates that the competitive equilibrium cannot contain mass points; the condition cannot hold for all \( p \in [\underline{p}, \bar{p}] \) if \( v(p, q) \) is discontinuous at any point \( p_0 \in [\underline{p}, \bar{p}] \).

\(^{26}\) Equation (4) is the first-order condition obtained from a perfectly competitive firm’s profit maximization. To see this, consider the oligopolist’s profit function, equation (11) below, with \( \rho(e, q) \) and \( e(p, q) \) fixed.
Proof. See the Appendix.

Similarly, if capacity is either costly or constrained, and there are \( m \) discrete demand states, the competitive equilibrium price distribution is strictly dispersed and is uniquely defined by

\[
p_i = c + \frac{\lambda}{m} \sum_{j=1}^{m} f_j
\]

and

\[
q_{c,i} = eD(p) \left[ 1 - \sum_{j=1}^{i-1} \frac{q_{c,j}}{eD(p)} \right].
\]

Proposition 1 generalizes Prescott (1975) (and Eden, 1990) by considering a continuum of demand states. Note also that many different equilibrium price distributions for individual firms yield the same aggregate equilibrium price distribution. It is natural to suppose that each firm sets a single price, as Prescott did, in which case firms specialize in different prices and in serving different market niches. In some of the examples suggested earlier there does appear to be some price specialization, e.g., discount hotels and discount car rentals, although those firms’ products are often differentiated in ways other than price.

However, Proposition 1 is equally consistent with the interpretation that firms choose symmetric price distributions, \( q_c(p)/n \). This distinction is important because I find that in the oligopoly model, only a symmetric equilibrium exists. That is, imperfect competition implies intrafirm price dispersion and not specialization in prices.

Three other observations are worth noting. First, there is excess demand when the demand state is between \( e(\overline{\varphi}, q) \) and \( \overline{\varphi} \), so stockouts do occur. If not, then there would be some unit that was sold only in state \( \overline{\varphi} \), but the profit on that unit,

\[
(p - c)(1 - F(\overline{\varphi})) - \lambda = -\lambda,
\]

is necessarily negative because the unit would be sold with probability zero.

Second, the competitive equilibrium price distribution is not continuous in \( \lambda \) at \( \lambda = 0 \). As \( \lambda \to 0 \) the support of the price distribution is actually increasing, even though the mean of the distribution is converging to \( c \). At \( \lambda = 0 \), the support of the distribution collapses to the point \( c \), and the equilibrium is the standard competitive equilibrium with a uniform price, \( p = c \).

Third, the competitive equilibrium is not Pareto efficient. There are clearly deadweight losses that arise in each state of nature. This is manifested in two ways. First, goods may not be allocated to the consumers who value them most (because of the proportional rationing rule), and second, capacity may not be fully utilized even when there are consumers whose willingness to pay exceeds marginal cost. Of course, it may not be relevant to compare the equilibrium here to one that might arise if complete forward contracts existed, since any pricing policy implemented through regulation would undoubtedly be subject to most of the same price rigidities and rationing as in the unregulated marketplace.
Monopoly. Here I examine the optimal pricing strategy of a monopolist. The monopolist chooses a price distribution, \( q(p) \) to maximize

\[
\pi(q) = \int_{e(p,q)}^{e(\bar{p},q)} \int_{e}^{e(p,q)} (p - c)q(p) \, dp \, de + \int_{e(p,q)}^{e(\bar{p},q)} (p - c)q(p) \, dp \, de - \lambda \int_{e}^{\bar{e}} q(p) \, dp.
\]

The first term is the firm’s expected net revenue when the state is between \( e(p, q) \) and \( e(\bar{p}, q) \). The inner integral is the net revenue earned from each unit sold in state \( e \), and the outer integral takes the expectation over demand states between \( e(p, q) \) and \( e(\bar{p}, q) \). These are the states in which the monopolist sells only some of its capacity, i.e., that priced below \( \rho(e, q) \). The second term is the firm’s expected net revenue when \( e \) is large enough so that all its units are sold. The last term is the firm’s direct cost of capacity.

The first-order condition obtained from (6) is

\[
[1 - F(e(p, q))](p - c) - \lambda \int_{e(p,q)}^{e(\bar{p},q)} (\rho(e, q) - c) \frac{D(\rho(e, q))}{D(p)} f(e) \, de = 0,
\]

for all \( p \in [p, \bar{p}] \), and is derived in the Technical Appendix.\textsuperscript{27} Intuitively, the monopolist sets the expected net revenue from an additional unit at price \( p \) equal to the cost of adding another unit of capacity plus the expected loss in revenue that results if a higher-priced unit goes unsold as a consequence of adding the additional unit at price \( p \).

If the monopolist faces a capacity constraint instead of a direct cost of capacity, then the only difference is that \( \lambda \) is the shadow cost of capacity and is determined endogenously. The equilibrium value of \( \lambda \) in the profit maximization when capacity is constrained by \( K \) is equal to the direct cost of capacity that would have induced the firm to choose capacity \( K \) when it chooses prices and capacity simultaneously.

Defining \( y_m(p) = 1 - F(e(p, q)) \) to be the probability that a unit of output offered at price \( p \) is sold, the first-order condition is solved for a closed-form solution given in Proposition 2.

**Proposition 2.** If capacity is either costly or constrained, then the monopoly equilibrium price distribution is strictly dispersed and is uniquely defined by

\[
y_m(p) = \frac{\lambda D_p(p)}{D_p(p)(p - c) + D(p)}
\]

and

\[
q_m(p) = -D(p) \frac{y'_m(p)}{f(F^{-1}(1 - y_m(p)))}
\]

on the support \( [p_m, \bar{p}] \), where \( p_m \) is defined using (8) by \( y_m(p_m) = 1 \).

**Proof.** See the Appendix.

\textsuperscript{27} The Technical Appendix is available at http://www.kellogg.nwu.edu/faculty/dana/.
Note first that (8) can be written in a more familiar form as

$$\frac{p - c - \frac{\lambda}{\gamma_\mu(p)}}{p} = \frac{1}{\eta(p)},$$

(10)

where $\eta(p)$ is the elasticity of demand at price $p$. The monopolist’s markup can be thought of as the difference between price and marginal cost plus the “effective marginal cost” of capacity, where the latter is the cost of capacity divided by the probability of sale. The firm sets a higher price on a unit that sells only in higher-demand states because its effective cost is higher, or equivalently, the firm sets a higher price because conditional on demand being high enough to sell that unit, the firm’s opportunity cost of capacity is higher. The monopolist’s markup, as defined in (10), declines as price increases as long as the elasticity of demand is decreasing in price.

Also notice that if $\lambda$ is zero, then the monopolist’s price distribution collapses to a single price $p_m$, which from (8) is clearly the standard monopoly price. This is true despite the demand uncertainty because demand is multiplicatively separable (so marginal revenue is invariant to changes in the demand state).

The assumption that $\epsilon$ has full support can easily be relaxed. If demand is strictly positive at $p_m$ in demand state $e$, then the monopoly price distribution will have a mass point of size $D(p_m, e)$ at $p_m$ but will otherwise be described by (8) and (9) as above.

If the distribution of demand uncertainty is discrete, as in the example given in Section 3, then Proposition 2 and (10) suggest that

$$\frac{p_i - c - \frac{\lambda}{\sum_{j=1}^{m} f_j}}{p_i} = \frac{1}{\eta(p_i)},$$

and

$$q_{m,i} = e_iD(p_i) \left[1 - \sum_{j=1}^{i-1} \frac{q_{m,j}}{e_iD(p_j)}\right],$$

where $\sum_{j=1}^{m} f_j$ is the probability that a unit priced at $p_i$ will sell in equilibrium. While a formal proof of this result is not given here (a complete proof requires that the distribution of $\epsilon$ be generalized to allow mass points while still allowing the firm to choose continuous price distributions), the discrete price distribution above is the limit of the continuous distribution in Proposition 2 as the distribution $F(e)$ converges to the discrete distribution given by \{e_1, e_2, . . . , e_m\} and \{f_1, f_2, . . . , f_m\}. And in particular, units that are priced to sell in even the lowest-demand state (those with price $p_1$) are priced as if demand were certain demand. This suggests that the prices derived in the example are indeed optimal.

\[\square\]

**Oligopoly.** In this subsection I derive a Nash equilibrium in price distributions when there are $n$ identical firms. I consider two cases: (1) firms choose price distributions given a cost of capacity, and (2) firms choose price distributions subject to a capacity constraint. The capacity-constrained model extends the classic Bertrand-Edgeworth model considered by Dasgupta and Maskin (1986) and Allen and Hellwig
(1986a, 1986b) by considering uncertain demand and allowing firms’ strategy spaces to be increased to include price distributions.

Given a cost of capacity, $\lambda$, an oligopolist chooses $q'(p)$ to maximize

$$
\pi(q'|q^{-}) = \int_{\epsilon(p,q)}^{\sigma(p,q)} (p - c)q'(p) \, dp \int_{\epsilon(p,q)}^{\sigma(p,q)} f(e) \, de + \int_{\sigma(p,q)}^{\sigma(p,q)} (p - c)q'(p) \, dp \int_{\epsilon(p,q)}^{\sigma(p,q)} f(e) \, de
$$

\[ - \lambda \int_{\epsilon(p,q)}^{\sigma(p,q)} q'(p) \, dp, \]

where $q^{-} = (q_{i})_{i=1}^{n}$ and $q = q^{i} + q^{-}$. As in the monopoly case (see the description of (6) above), the first term is the firm’s expected net revenue when the demand state is between $\epsilon(p, q)$ and $\epsilon(p, q)$. The second term is the firm’s expected net revenue when $e$ is large enough so that the entire market capacity is sold. The final term is the firm’s direct cost of capacity.

The first-order condition obtained from (11), and derived in the Technical Appendix, is

$$
[1 - F(\epsilon(p, q))](p - c) - \lambda - \int_{\epsilon(p,q)}^{\sigma(p,q)} \frac{q'(\rho(p, q))}{q(p, q)} (\rho(p, q) - c) \frac{D(\rho(p, q))}{D(p)} f(e) \, de = 0.
$$

\[ (12) \]

In the case where firms face a capacity constraint I assume that the firms have identical capacities, $K/n$. I also assume that these capacity constraints are binding. The oligopolist maximizes its expected profit, the first two terms of (11), subject to

$$
\int_{\epsilon(p,q)}^{\sigma(p,q)} q'(p) \, dp = \frac{K}{n}.
$$

\[ (13) \]

The first-order condition obtained from this maximization is

$$
[1 - F(\epsilon(p, q))](p - c) - \lambda - \int_{\epsilon(p,q)}^{\sigma(p,q)} \frac{q'(\rho(p, q))}{q(p, q)} (\rho(p, q) - c) \frac{D(\rho(p, q))}{D(p)} f(e) \, de = 0,
$$

\[ (14) \]

where $\lambda_{i}$ is the Lagrange multiplier for firm $i$ associated with the constraint, (13). Note that since $q'(p)$ is zero outside of the support of the firm $i$, (14) holds if we define $\overline{p} = \sup_{i} [\overline{p}_{i}]$, so $\overline{p}$ is the upper bound of the support of the industry distribution. By evaluating (14) at $\overline{p}$, it is easy to show that $\lambda_{i} = \lambda_{j}$ for all $i$ and $j$. Hence given a symmetric capacity constraint, there is a unique endogenous shadow cost of capacity $\lambda$ for all firms, and in equilibrium firms’ prices will be the same as if $\lambda$ were the exogenous cost of capacity. However, the interpretation of $\lambda$ is clearly different in (12) and (14).

Using $n$ to denote the $n$-firm oligopoly solution, I define $\gamma_{n}(p) = 1 - F(\epsilon(p, q))$ to be the probability of sale as before and $q_{n}(p)$ to be the price distribution of a typical firm. Equations (12) and (14) can be rewritten as a singular, linear differential equation and solved for the solution given in Proposition 3.
Proposition 3. If capacity is costly or constrained, and if $p_n = p^m(c)$, then the oligopoly equilibrium price distribution is strictly dispersed and is uniquely defined by

$$y_n(p) = \frac{n}{n-1} \int_p^\pi \lambda D_p(r)(r - c)D(r)^{1/(n-1)} \, dr$$

and

$$q_n(p) = -D(p) \frac{y_n(p)}{f(F^{-1}(1 - y_n(p)))}$$

on the support $[p^m, \bar{p}]$, where $p_n$ is defined by $y_n(p_n) = 1$.

Proof. See the Appendix.

Corollary. If the cost of capacity, $\lambda$, is sufficiently large, in particular if $\lambda > p^m(c) - c$, then $p_n = p^m(c)$, since $p_n > c + \lambda$, so the condition in Proposition 3 is satisfied.

This is not surprising, since under the proportional rationing rule a similar condition must hold in the case of certain demand for the Bertrand-Edgeworth game to have a pure-strategy equilibrium in prices (see Davidson and Deneckere, 1986). If $\lambda$ is small, then the model undoubtedly has mixed-strategy equilibria, and it is likely that the mixing would be over price distributions that are strictly dispersed and contain mass points. While I have not proved that $p_n = p^m(c)$ is a necessary condition, numerical simulations I have performed suggest that when $\lambda$ is small, the solution to (15) and (16) constitutes a saddle point and is not a local maximum.

Unlike Propositions 1 and 2, Proposition 3 relies on the assumptions that $e$ has full support and that $F(e)$ is continuous. We saw in the example in Section 3 that pure-strategy equilibria of the oligopoly game did not exist when $F(e)$ contained mass points, at least for the case of constrained capacity. However, if $e$ is positive, so $e$ does not have full support, then firms’ first-order conditions will yield mass points at the lowest bound of their price distributions, and the same nonexistence argument would apply. For the simultaneous choice of prices and capacity the problem is potentially worse. Gertner (1985) showed that even with certain demand, if firms can simultaneously increase output and undercut their rival’s price, then no pure-strategy equilibrium can exist.

Proposition 3 proves uniqueness of the equilibrium given the restriction that firms’ strategies contain no mass points. While I have proved that the pure-strategy equilibrium derived above is still an equilibrium when firms’ strategy sets are enlarged, I have not proved that the equilibrium is unique in the larger strategy space.

The oligopoly equilibrium has the intuitive property that as the number of firms approaches infinity, the equilibrium approaches the competitive equilibrium.

Proposition 4. Let (15) and (16) define $q_n(p)$ as a continuous function of $n$ when $p_n \equiv p^m(c)$. Then $q_n(p) = \lim_{n \to \infty} q_n(p)$ and $q_n(p) = \lim_{n \to 1} q_n(p)$.

The proposition can be proved directly using (15) and L’Hôpital’s rule and can be

---

28 Another difficulty with Proposition 3 is that the local second-order condition is difficult to check either analytically or numerically. The second-order condition and a proof that local second-order conditions for a special case are provided in the Technical Appendix. The proof of Proposition 3 in the Appendix is based instead on the Weierstrass theorem.
seen intuitively using (A9) in the Appendix. Since the definition of competitive equilibrium does not impose any restrictions on individual firms’ price distributions and since \( q'(p) = q_n(p)/n \), Proposition 4 suggests that the competitive equilibrium can be interpreted as a symmetric equilibrium with intrafirm price dispersion. It also suggests that many of the applied results obtained by extending the Prescott model might also hold under the assumption of imperfect competition.

The most natural assumption about the timing of firms’ decision making with price rigidities and demand uncertainty is that firms choose capacity, then prices, and then demand is realized. In the perfect-competition and monopoly models, this timing yields the same result as those presented here when capacity and price are chosen simultaneously. In the oligopoly model this is not the case. Oligopolists will not choose the same capacity when capacity is chosen sequentially as they will when capacity is chosen simultaneously. The intuition is simple (and is the same as the intuition in models of sequential choice with known demand, such as Kreps and Scheinkman (1983), Osborne and Pitchik (1986), and Davidson and Deneckere (1986)). When capacity and prices are chosen simultaneously, a deviation that reduces a firm’s capacity will have no effect on the equilibrium pricing of the other firms. In the sequential game, however, when a firm deviates by reducing capacity in the first stage, that will increase the shadow cost of capacity for every other firm in the pricing subgame and cause other firms to set higher prices. Proposition 3 describes a price and quantity game that is much closer in spirit to Bertrand competition, while the sequential capacity, price, and quantity game is much closer in spirit to Cournot competition.

6. Price dispersion and market structure

The following proposition demonstrates that prices become more dispersed as markets become more competitive. Two qualifications deserve mention: first, the result is obtained by comparing the supports of the distributions and not their variances; second, the result relies on the maintained assumption that the distribution of \( e \) is continuous.

**Proposition 5.** Holding the cost of capacity constant, the support of the market price distribution is greatest under perfect competition and shrinks as the number of firms decreases.

**Proof.** See the Appendix.

To illustrate Proposition 5, I calculated the equilibrium price distribution for each of the three market structures when demand is linear in price and demand uncertainty is uniformly distributed. Specifically, I assumed \( D(p, e) = (1 - p)e \), \( e \) is uniformly distributed on \([0, 1]\), \( c = .1 \), and \( \lambda = .5 \). The densities are shown in Figure 4 for monopoly, perfect competition, and for oligopoly with \( n = 2 \), \( n = 4 \), and \( n = 10 \). Figure 5 is a graph of the cumulative price distributions for the same example, shown on the standard price-quantity axes of supply and demand functions. The demand curve in state \( e = ½ \) is included as a reference point. Note that from Figure 5, it is clear that prices are more dispersed the more competitive is the market.

I conjecture that the variance of the price distribution also increases with \( n \) when demand is linear in price. I have shown this numerically for the example above, but I do not think that any restrictions on the distribution of \( e \) are required, and none are used in the proof of the following statement of this conjecture for the limiting cases of monopoly and perfect competition.

**Proposition 6.** If demand is multiplicatively separable and linear in price, then the variance of the equilibrium price distribution under perfect competition is strictly greater than the variance of the price distribution under monopoly pricing.

© RAND 1999.
Proof. See the Appendix.

Proposition 5 exploits the fact that the intercept of the inverse demand function is finite and Proposition 6 the fact that demand is linear. The latter assumption, by eliminating convexity in the demand, guarantees that a monopolist will raise prices less when its costs increase than a competitive market. The former assumption limits the convexity of demand in a similar but less restrictive way.

While several theories have been developed in the last decade that explain price dispersion in imperfectly competitive markets, few other models predict an inverse relationship between industry concentration and price dispersion. Yet this is precisely
the relationship found by Borenstein and Rose (1994) in the airline industry and discussed earlier.

Shepard (1991) examines the retail gasoline market and finds evidence of price discrimination by retailers who sell both self-service and full-service gasoline. Although Shepard tests and rejects a variant of the zero-profit condition given in equation (2) and concludes that multiproduct firms are price discriminating, her results raise many questions about the mechanism by which these firms price discriminate. Pricing in the gasoline market may reflect a combination of traditional price discrimination and the capacity-cost-based price dispersion discussed here.

7. Conclusion

I demonstrate that price rigidities and demand uncertainty lead not only to interfirm price dispersion, as has been previously pointed out, but also to intrafirm price dispersion. I find a unique equilibrium price distribution for monopoly and imperfectly competitive markets when demand is uncertain and firms’ strategy sets are enlarged to include distributions of prices. I also show that price dispersion increases with the number of firms, in contrast to the relationship predicted by typical models of price discrimination. The result is a useful theory of intrafirm price dispersion that can explain much of the price variation in relatively competitive markets that has typically been attributed to price discrimination, especially in those markets where firms practice some form of revenue management. Although the assumptions of the model are restrictive—demand is uncertain, prices are rigid, and goods are not easily stored—some industries, such as airlines, hotels, and automobile rentals, seem to satisfy the spirit of these assumptions and exhibit an unusual amount of price dispersion. While much of the intrafirm price dispersion seen in these industries has other characteristics of price discrimination, such as the airlines’ use of travel restrictions (e.g., Saturday-night stayovers) on discount fares and advance purchase requirements, this article offers a complementary explanation of price dispersion.

Appendix

Proofs of Propositions 1, 2, 3, 5, and 6 follow.

Proof of Proposition 1. The equilibrium condition, equation (4), follows from profit maximization. Profit maximization also implies that in a competitive equilibrium, $(p - c)(1 - F(\epsilon(p, q))) - \lambda \leq 0$ for all $p \notin [\underline{p}, \overline{p}]$. Combining equations (3) and (4), we have

$$\int_{\underline{p}}^{\overline{p}} q_0(r) \, dr = F^{-1}\left(1 - \frac{\lambda}{p - c}\right).$$

(A1)

$\forall p \in [\underline{p}, \overline{p}]$, and differentiating (A1) yields (5) $\forall p \in (\underline{p}, \overline{p})$. Equation (3) implies that $\epsilon(p, q) = 0$. So $F(\epsilon(p, q)) = 0$, and from (4) we have $p = c + \lambda$. So (5) also defines $q(p)$, since the right-hand side of (A1) is zero at $p$. Equation (5) and $\lambda > 0$ imply that $q(p) > 0$ for all $p \in [\underline{p}, \overline{p}]$. Suppose $\overline{p} < \overline{p}$. From (3) for all $p \in [\overline{p}, \overline{p}]$, we have $\epsilon(p, q) = \epsilon(\overline{p}, q)$, so

$$(p - c)[1 - F(\epsilon(p, q))] - \lambda > (\overline{p} - c)[1 - F(\epsilon(\overline{p}, q))] - \lambda = 0,$$

or $p > c + \lambda[1 - F(\epsilon(p, q))]$, which is a contradiction of the condition for profit maximization. Q.E.D.

Proof of Proposition 2. If $\overline{q}$ is an extremum of (6), then the Gateaux differential, $\delta \pi(q; h')$, of the monopolist’s profit function must be equal to zero for all Lebesgue integrable functions $h'$. Or, in terms of the standard calculus, $d\pi(q' + ah')|_{a=0} = 0$ for all $h'$. The first-order condition (7) is derived from (6) using this methodology; the derivation is included in the Technical Appendix. The first-order condition, (7), can be written as

© RAND 1999.
\[ y_n(p) = \frac{\lambda}{p - c} - \frac{1}{(p - c)D(p)} \int_\pi^\rho (r - c)D(r)y_n'(r) \, dr. \]  
(A2)

Multiplying both sides of (A2) by \((p - c)D(p)\) and differentiating with respect to \(p\) yields (8), since the terms involving \(y_n'(p)\) cancel. Since \(\lambda > 0\), \(y_n(p)\) is clearly continuously differentiable. Differentiating (3), we have

\[ q(p) = D(p)\frac{de(p, q)}{dp}, \]  
(A3)

and differentiating the identity \(y_n(p) = 1 - F(e(p, q))\) yields

\[ \frac{de(p, q)}{dp} = -\frac{y_n'(p)}{f(F^{-1}(1 - y_n(p)))}, \]  
(A4)

so the equilibrium price distribution is given by (8) and (9).

Clearly \(\pi \leq \overline{\pi}\). Suppose that \(\pi < \overline{\pi}\). Then, as in Proposition 1, (3) implies that \(e(p, q) = e(\overline{\pi}, q)\) for all \(\pi < p \leq \overline{\pi}\). Clearly (7) holds at \(\pi\), but \(\overline{\pi} \leq \overline{\pi}\) implies that the left-hand side of (7) is strictly positive for all \(p\) such that \(\pi < p \leq \overline{\pi}\), which is a contradiction. Hence \(\pi = \overline{\pi}\).

Equation (3) implies that \(e(p, q) = 0\). So \(y_n(p) = 1 - F(e(p, q)) = 1\). Hence, using (8), \(\pi_n\) is defined implicitly by \(y_n(p) = 1\), or

\[ p = c + \lambda - \frac{D(p)}{D_f(p)}, \]  
(A5)

so \(\pi_n > c + \lambda\). From (8),

\[ y_n'(p) = \frac{\lambda(D_nD - 2(D_p)^2)}{(D_p)^2(p - c + \frac{D(p)}{D_f(p)})} < 0, \]  
(A6)

for all \(p \in (p, \overline{\pi})\). Equations (A6), (9), and \(\lambda > 0\) imply \(q(p) > 0\) for all \(p \in (p, \overline{\pi})\), and \(\lim_{p \to \overline{\pi}} q(p) = 0\).

The local second-order condition, also derived in the Technical Appendix, is

\[ \frac{d^2\pi}{dp^2} = \int_\pi^\rho \left[ D(p)D_f(p) + (p - c)D_p(p)q(p)\right]\left[ \left( r - c \right) D(r) dy_n'(r) \right] dr < 0, \]  
(A7)

for all continuous functions \(h(q)\). The second-order condition holds if \((p - c) + D(p)D_f(p) > 0\) for all \(\pi_n < p \leq \overline{\pi}\), which follows from \((p_n - c) + D(p_n)D_f(p_n) = \lambda\) and from observing that \((p - c) + D(p)D_f(p)\) is increasing in \(p\) (since \(2(D_p)^2 - D_nD_f > 0\) by assumption). Hence the monopoly solution is a local maximum. \(\Box\)

Proof of Proposition 3. The first-order condition (12) is derived in the Technical Appendix. Adding equation (12) for \(i = 1 \text{ to } n\) and dividing by \(n\) yields

\[ [1 - F(e(p, q))(p - c) - \lambda - \frac{1}{n} \int_{e(p, q)}^{e(p, \overline{\pi})} (\rho(e, q) - c)D(p(e, q)) \, de] = 0. \]  
(A8)

Using \(y_n(p) = 1 - F(e(p, q))\) and changing the variable of integration yields

\[ y_n(p)(p - c)D(p) - \lambda D(p) = -\frac{1}{n} \int_\pi^\rho (r - c)D(r)y_n'(r) \, dr, \]  
(A9)

Finally, differentiating (A9) yields the following linear differential equation:

\[ y_n''(p) = \frac{n}{(p - c)D(p)} - \frac{1}{n - 1} \frac{D(p) + (p - c)D_f(p)}{(p - c)D(p)} y_n'(p). \]  
(A10)

Clearly \(\overline{\pi} \leq \overline{\pi}\). Suppose that \(\pi < \overline{\pi}\). Once again, (3) clearly implies that \(e(p, q) = e(\overline{\pi}, q)\) for all
The first-order condition is strictly positive for all \( p \) such that \( p < \varrho \), which is a contradiction. Hence \( p = \varrho \).

The first-order condition, (A10), has a singularity at \( p = \varrho \) since each of the right-hand-side coefficients explode at \( \varrho \). However, this equation does have a unique solution. To obtain the solution (15), I first consider the differential equation (A10) with the modified initial-value condition \( y(\varrho - \delta) = \lambda(\varrho - c) \), where \( \delta \) is a small, positive number. This modified initial-value problem has the following unique solution, derived using standard methods:

\[
y_{p}(p) = \frac{\lambda}{(\varrho - c)} \left[ (\varrho - \delta - c)D[\varrho - \delta] \right]^{(n-1)} - \frac{n}{n-1} \int_{c}^{\varrho} \lambda D_{p}(r)[(r - c)D(r)]^{(n-1)} dr
\]

Equation (15) is the limit of (A11) as \( \delta \) goes to 0. Moreover, it is easily verified that (15) is a continuous solution to (A10) with bounded derivatives. In particular, differentiation of \( y_{p}(\varrho) = \lambda(\varrho - c) \) using L'Hôpital's rule yields \( y_{p}'(\varrho) = -(2n/(2n - 1))(\varrho(\varrho - c)^{2}) \). So existence is proved by construction.

Equation (15) is also the unique solution to (A10). The proof follows from a more general proof of existence and uniqueness in Dana (1996). However, a more direct proof follows. Let \( y_{p}'(p) \) and \( y_{q}'(p) \) be any two continuously differentiable solutions of (A10) with the initial-value condition \( y_{p}(\varrho) = \lambda(\varrho - c) \). Let \( y_{p}'(p) \) and \( y_{q}'(p) \) be the unique solutions to (A10) with the initial-value conditions \( y_{p}'(\varrho - \delta) = y_{p}'(\varrho - \delta) \) and \( y_{q}'(\varrho - \delta) = y_{q}'(\varrho - \delta) \) respectively, and note that these solutions depend continuously on \( \delta \). By construction, \( y_{p}'(p) = y_{p}'(p) \) and \( y_{q}'(p) = y_{q}'(p) \) for all \( p \in [\varrho, \varrho - \delta] \). Since (A10) is a linear first-order differential equation, it follows that

\[
y_{p}'(p) - y_{q}'(p) = \lim_{\delta \to 0} [y_{p}'(p) - y_{q}'(p)] = \lim_{\delta \to 0} [y_{p}'(\varrho - \delta) - y_{q}'(\varrho - \delta)] = 0
\]

for all \( p \), so \( y_{p}'(p) = y_{q}'(p) \) are identical.

As in Propositions 1 and 2, (3) implies \( \epsilon(p, q) = 0 \). So using (15), \( p \) is defined implicitly by \( y_{p}(p) = 1 - \epsilon(p, q) = 1 \).

To show that \( q(p) \geq 0 \), I will first show that \( y_{p}'(p) = 0 \). Equation (A10) can be rewritten as

\[
y_{p}'(p) = \frac{n}{n-1} D_{p}(p) y_{p}(p) + \frac{1}{n-1} + \frac{\lambda D_{p}(p)}{n-1} (p-c)D(p).
\]

Multiplying (A12) by \( (p-c) \) and substituting for \( y_{p}(p)(p-c) \) in the first term using (A9) yields

\[
(p-c)y_{p}'(p) = \frac{1}{n-1} D_{p}(p) \int_{A}^{B} (r-c)D(r)y_{p}'(r) dr - \frac{n}{n-1} y_{p}'(p).
\]

Differentiating (A13) and isolating \( y_{p}'(p) \) yields

\[
y_{p}'(p) = \left[ \frac{1}{n-1} D_{p}(p) + \frac{2n-1}{n-1} \right] y_{p}'(p) + \frac{1}{n-1} \frac{[D_{p}(p)D(p) - 2D_{p}(p)]}{(p-c)D(p)} \int_{A}^{B} (r-c)D(r)y'(r) dr.
\]

Let \( \beta = \inf \{ \beta \in [\varrho, \varrho] \mid y'(p) < 0, \forall p \in ([\beta, \varrho]) \} \) and note that (15) implies that \( y'(p) \) is continuous on \([\varrho - \varrho, \varrho)\). Clearly,

\[
y'(\varrho) = -\frac{2n}{2n-1} \frac{\lambda}{(\varrho - c)^{2}} < 0,
\]

so by continuity, \( \beta < \varrho \). Also by continuity, either \( y'(\beta) = 0 \) or \( \beta = \varrho \). If \( \beta > \varrho \), then by (A14) \( y'(\beta) > 0 \), since the first term is zero when \( y'(\beta) = 0 \) and the second term is positive since \( D_{p}(p)D(p) - 2D_{p}(p)^{2} < 0 \) by assumption and \( y'(p) < 0, \forall p \in ([\beta, \varrho]) \) by construction. But by continuity, \( y'(\beta) > 0 \) and \( y'(\beta) = 0 \) imply that for some \( p > \beta \), \( y'(p) > 0 \), which is a contradiction. Therefore, \( \beta = \varrho \), and \( y'(p) < 0, \forall p \in ([\varrho, \varrho]) \). And from (16) this clearly implies that \( q(p) > 0, \forall p \in ([\varrho, \varrho]) \).

\[
\text{© RAND 1999.}
\]
Since \( q_i(p) \) is unique, it follows from (14) that the firm’s price distribution, \( q^*(p) \), is also unique and is given by \( q^*(p) = q_i(p)/n \). This is easily seen by performing a change of variables on (14) (as was done in equation (A2)) and differentiating with respect to \( p \). A unique expression for \( q^*(p) \) is obtained in terms of \( q_i(p) \), hence \( q^*(p) = q_i(p) \) and \( q^*(p) = q_i(p)/n \).

The second-order condition, also derived in the Technical Appendix, is

\[
\frac{d^2\pi}{da^2}_{|a=0} = -2\left(n - \frac{1}{n}\right) \int_{p_1}^{p_2} \int_{e_1}^{e_2} h'(p) \left(\frac{\partial\phi(e, q)}{D(p, e)} - c \frac{D(p(e, q), e)}{q_i(e, q)}\right) h'(p) dp de \\
+ \frac{1}{n} \int_{p_1}^{p_2} \int_{e_1}^{e_2} \left[\frac{\partial D(p(e, q), e)}{D(p, e)} - \frac{D(p(e, q), e)}{q_i(e, q)}\right] \left[\frac{h'(p)}{D(p, e)}\right]^2 f(e) de
\]

for all continuous functions \( h'(p) \). The second term is negative if \( p_2 > p^*(c) \), since then

\[
D(p(e, q), e) + (p(e, q) - c)D_e(p(e, q), e) < 0 \quad \forall \ e;
\]

however, the first term cannot be signed in general. After a change of variables and some manipulation, the first term can be written as

\[
-2\left(n - \frac{1}{n}\right) \int_{p_1}^{p_2} \int_{e_1}^{e_2} k(p, r) h'(p) h(r) dp dr,
\]

where \( k(p, r) = (r - c)/(D(r)f(e,r, q)) \), \( p \leq r \leq \overline{p} \), and \( k(p, r) = (p - c)/(D(r)f(e,r, p)) \), \( r \leq p \leq \overline{p} \). Equation (A17) is negative if and only if \( k(p, r) \) is a positive operator. If demand is linear, the distribution of demand uncertainty is uniform, and \( \lambda > (\overline{p} - c)/2 \), then it is possible to explicitly solve the eigenvalue problem associated with \( k(p, r) \) and prove that \( k(p, r) \) is a positive operator (the details are available in the Technical Appendix). Note that with linear demand \( p^*(c) = (\overline{p} + c)/2 \) and \( \lambda > (\overline{p} - c)/2 \) implies \( p_2 > p^*(c) \), so the second term is also negative and the second-order condition is satisfied. But aside from this special case I have not been able to sign (A17); however, I prove that the solution is a maximum using the Weierstrass theorem.

I have verified that the first-order condition has a unique solution \( q^* = q^*(p)/n \). Since the objective function is continuously differentiable, it follows that if a maximum exists, it must satisfy the first-order condition. To prove that a maximum must exist, first note that the solution is bounded and differentiable with a bounded derivative (this is a standard property of the solution to first-order differential equations, which can be applied for prices below \( \overline{p} \), and at \( \overline{p} \) we have \( q(\overline{p}) = 0 \) and \( q'(\overline{p}) = 0 \). Choose constants \( \bar{k}_1 \gg \max(\sup q^*(p)/n, \inf \lim_{p \rightarrow \overline{p}} \int_{e_1}^{e_2} dD(r)/D(r)) \) and \( \bar{k}_2 \gg \sup q^*(p)/n \), and consider the set of continuous functions \( \{q(p) = \int_{e_1}^{e_2} q(r) dr | [p, \overline{p}] \} \) that (i) are equicontinuous with Lipschitz constant \( \bar{k}_2 \) and (ii) satisfy \( 0 \leq q(p) \leq \bar{k}_1 \) for all \( p \). This set of functions is compact, and by the Weierstrass theorem there exists an element of this set \( q^*(p) \) that maximizes the oligopolist’s profit given the competitors’ strategies,

\[
q^*(p) = q^*(p)/n, \quad \forall \ j \neq i.
\]

I now prove that \( q^*(p) \) must satisfy the firm’s first-order condition (almost everywhere) and hence that \( q^*(p) = q^*(p)/n \). The first-order condition evaluated at \( q^*(p) \) is

\[
FOC(p) = \frac{\partial}{\partial p} [1 - F(e(p, q^*))] (p - c) - \lambda - \int_{e_1}^{e_2} \frac{\partial q^*(r)}{q^*(r)} - \lambda \int_{e_1}^{e_2} \frac{\partial q^*(r)}{q^*(r)} \frac{D(r)}{D(p)} f(e(r, q^*)) \frac{d(e(r, q^*))}{dr} dr,
\]

where \( \lambda = \overline{\lambda} + \lambda_i^* (r) \), and \( \lambda_i^* (r) = (n - 1)/(q^*(p)/n) \).

First I show that \( \lim_{p \rightarrow \overline{p}} \frac{\partial q^*(p)}{\partial p} = 0 \). Clearly, \( \lim_{p \rightarrow \overline{p}} \frac{\partial q^*(p)}{\partial p} = 0 \), since otherwise \( q(\overline{p}, q^*) > \overline{q}, q(\overline{p}) = 0 \), and \( \overline{\lambda} > 0 \) is a contradiction of profit maximization. So if \( \lim_{p \rightarrow \overline{p}} \frac{\partial q^*(p)}{\partial p} > 0 \), then \( FOC(p) > 0 \), and \( FOC(p) = 0 \) on some neighborhood of \( \overline{p} \). But that implies \( q^*(p) = q^*(p)/n \) (almost everywhere) in some neighborhood of \( \overline{p} \), so \( \lim_{p \rightarrow \overline{p}} \frac{\partial q^*(p)}{\partial p} = 0 \), which is a contradiction.

Next I will show that \( \frac{\partial q^*(p)}{\partial p} = q^*(p)/n \) (almost everywhere) in some neighborhood of \( \overline{p} \). This is equivalent to showing that \( FOC(p) = 0 \) on some neighborhood of \( \overline{p} \). Suppose \( FOC(\overline{p}) < 0 \). Then there exists a price \( \overline{p} \leq \overline{p} \) such that \( FOC(\overline{p}) = 0 \), \( \gamma_i(\overline{p}) = \lambda_i(\overline{p}) \), and both \( FOC(p) < 0 \) and \( q^*(p) = 0 \) for all \( p \in (\overline{p}, \overline{p}) \). However, \( \frac{\partial q^*(p)}{\partial p} = 0 \) implies \( \gamma_i(\overline{p}) = ((n - 1)/n)\gamma_i(\overline{p}) \), where \( \gamma_i(\overline{p}) = 1 - F(e(p, q^*)) \), and from the equilibrium conditions for \( q^*(p) \) we have \((n - 1)/n)\gamma_i(\overline{p}) > -\lambda_i (\overline{p})^2 \) for all \( p \in (\overline{p}, \overline{p}) \), so
\[(\gamma_n(p) - \lambda(p - c)(p - c)\]

and equivalently \(\gamma_n(p)(p - c) - \lambda \) are increasing in \(p\), which implies \(\text{FOC}(p) > 0\) for all \(p \in (\beta, \bar{p}]\), which is a contradiction (the firm can increase its profit by increasing sales at prices greater than \(\beta\)). Hence \(\text{FOC}(\bar{p}) = 0\) and \(\gamma_n(\bar{p}) = \lambda(\bar{p} - c)\). It immediately follows that \(\bar{q}(p) > 0\) on a neighborhood \((\bar{p} - \epsilon, \bar{p})\) of \(\bar{p}\), and so \(\bar{q}(p) = q^*(p)/n\) and \(\gamma_n(p) = \gamma_n(p)\) on the same neighborhood.

I will now show that \(\bar{q}(p) = q^*(p)/n\) more generally. Suppose that \(\bar{q} \neq q^*/n\), and let \(\beta = \inf\{p; \bar{q}(r) = q^*(r)/n, \forall r \geq p\}\).

Consider the derivative of \(\text{FOC}(p)D(p)\) with respect to \(p\),

\[
[\text{FOC}(p)D(p)]' = \left(1 - \frac{\bar{q}'(p)}{\bar{q}(p) + q^* - \sigma(p)}\right)\gamma_n(p)D(p)(p - c) + \gamma_n(p)(D(p) + D_p(p)(p - c)) - \lambda D_p(p). \tag{A19}
\]

where \(\gamma_n = 1 - F(\epsilon(p, \bar{q}, q^* - \sigma))\). From (16) we have

\[
\left(1 - \frac{\bar{q}'(p)}{\bar{q}(p) + q^* - \sigma(p)}\right)\gamma_n(p) = -\left(\frac{\bar{q}'(p)}{\bar{q}(p) + q^* - \sigma(p)}\right)f(\epsilon(p, \bar{q}, D(p))\left(\frac{\bar{q}(p) + q^* - \sigma(p)}{D(p)}\right)
\]

\[
= \frac{n - 1}{f(\epsilon(p, q^*))}D(p). \tag{A20}
\]

So using \([\text{FOC}(p)D(p)]' = 0\) for all \(p\) when evaluated at \(q^*(p)/n\), we have

\[
[\text{FOC}(p)D(p)]' = (\gamma_n(p) - \gamma_n(p))(D(p) + D_p(p)(p - c)) + \frac{n - 1}{f(\epsilon(p, q^*))}\gamma_n(p)D(p)(p - c). \tag{A21}
\]

Since \(p > p^*(c), D(p) + D_p(p)(p - c) < 0\) for all \(p > p^*(c)\), so the first term above has the same sign as \(\gamma_n(p) \in \gamma_n(p)\).

Case 1. If \(\bar{q}(p) > q^*(p)/n\) for prices in a neighborhood of \(\beta\). Clearly, \(\text{FOC}(p) \geq 0\) and \(\text{FOC}(p)D(p) \geq 0\) near \(\beta\) (reducing output is feasible) and \(\gamma_n(p) > \gamma_n(p)\) by definition, and since \(\gamma_n(\beta) = \gamma_n(\beta)\), it follows that \(\gamma_n(p) > \gamma_n(p)\) and \(\epsilon(p, \bar{q}) < \epsilon(p, q^*)\) near \(\beta\). Since \(f\) is decreasing, the second term in the expression for \([\text{FOC}(p)D(p)]'\) above is negative, and since \(\gamma_n(p) > \gamma_n(p)\), we have \([\text{FOC}(p)D(p)]' < 0\). \(D(p) > 0\), so \(\text{FOC}(p) > 0\) for all \(p < \beta\), which implies that \(\bar{q}(p)\) is at the boundary of the set of all allowable functions for all \(p < \beta\). (As \(p\) falls, \(\bar{q}(p)\) rises at the maximum rate until \(\bar{q}(p)\) equals \(k_i\), and then \(\bar{q}(p)\) stays at \(k_i\). Since \(\text{FOC}(p)\) never falls to zero, \(\bar{q}(p)\) stays at the boundary.) But then \(\gamma_n(p) > 1, \) which is a contradiction.

Case 2. If \(\bar{q}(p) < q^*(p)/n\) for prices in a neighborhood of \(\beta\). Clearly, \(\text{FOC}(p) \leq 0\) and \(\text{FOC}(p)D(p) \leq 0\) near \(\beta\) (since increasing output is feasible) and \(\gamma_n(p) < \gamma_n(p)\) by definition, and since \(\gamma_n(\beta) = \gamma_n(\beta)\), it follows that \(\gamma_n(p) < \gamma_n(p)\) and \(\epsilon(p, \bar{q}) > \epsilon(p, q^*)\) near \(\beta\). Since \(f\) is decreasing, the second term in the expression for \([\text{FOC}(p)D(p)]'\) above is positive, and since \(\gamma_n(p) < \gamma_n(p)\), we have \([\text{FOC}(p)D(p)]' > 0\). Hence \(\text{FOC}(p) < 0\) for all \(p < \beta\), which implies that \(\bar{q}(p)\) is at the boundary of the set of all allowable functions for all \(p < \beta\). (As \(p\) falls, \(\bar{q}(p)\) falls at the maximum rate until \(\bar{q}(p)\) equals zero, and then \(\bar{q}(p)\) stays at zero.) But then \(\gamma_n(\beta) > 0, (\beta)\), which is a contradiction.

So if \(D(p) + D_p(p)(p - c) < 0\), then a maximum exists and satisfies the first-order condition. In an earlier version of the article I also show that this optimum does not contain mass points when the strategy set is enlarged to allow them. Hence the unique maximum, and the oligopoly equilibrium price distribution, is given by equations (15) and (16). \(Q.E.D.\)

Proof of Proposition 5. Under each market structure, the upper bound of the price support is the same, \(\bar{p} = \bar{p}\). However, the solution for \(p\) is not the same. In the oligopoly case, \(p\) is defined implicitly by \(\gamma_n(p) = 0\), using (15). Writing \(\gamma_n(p) \in \gamma(p, n)\), and treating \(n\) as a continuous variable, implicit differentiation yields

\[
\frac{dp}{dn} = -\frac{dy(p, n)/dn}{dy(p, n)/dp}. \tag{A22}
\]

The denominator is negative (see the proof of Proposition 3 above). Differentiating both sides of (A10) with respect to \(n\) and substituting using (A10) yields

© RAND 1999.
Let \( \beta = \min \{ \beta \in \mathbb{P} : dy(p, n)/dn < 0, \forall p \in (\beta, \mathbb{P}) \} \). It is possible to show that \( dy(\mathbb{P}, n)/dn = 0 \) and

\[
\frac{d^2y(\mathbb{P}, n)}{dpdn} = \frac{2}{(2n - 1)(\mathbb{P} - c)^2} > 0,
\]

so by continuity, \( \beta < \mathbb{P} \). Also by continuity, either \( dy(\mathbb{P}, n)/dn = 0 \) or \( \beta = p \). If \( \beta > p \), then \( dy(\mathbb{P}, n)/dn = 0 \), and by (A23), \( d^2y(\beta, n)/dpdn > 0 \), since the first term on the right-hand side of (A23) is clearly positive and the second term is zero when \( dy(\mathbb{P}, n)/dn = 0 \). But by continuity, \( d^2y(\beta, n)/dpdn > 0 \) implies that for some \( p > \beta \), \( dy(p, n)/dn > 0 \), which is a contradiction. Therefore, \( \beta = p \), and \( dy(p, n)/dn < 0, \forall p \in (\beta, \mathbb{P}) \). So, by (A22), \( dp/dn < 0 \). Q.E.D.

**Proof of Proposition 6.** The demand function is given by \( D(p, e) = eD(p) \), where \( D(p) = a - bp \). From (A3) we have

\[
\frac{de(p, q)}{dp} = \frac{q(p)}{D(p)}
\]

which implies

\[
\frac{dp(e, q)}{de} = \frac{D(p(e))}{qn(p(e))}
\]

So, using a change of variables, \( Q(p) \) can be written as

\[
Q(p) = \int_p^\epsilon q(p) dp = \int_0^{\epsilon(p \bar{e})} D(p(e)) de,
\]

where \( \epsilon(p, q) = \epsilon = 0 \).

From (4) and (10), and using the assumption that demand is linear, \( D(p) = a - bp \), we have

\[
\rho'(e) = c + \frac{1}{1 - F(e)}
\]

\[
\rho''(e) = -\frac{1}{b} + \frac{\lambda}{1 - F(e)}
\]

which implies \( D(p(e)) = D(p'(e))/2 \) and, using (A27), that \( Q(p''(e)) = Q(p'(e))/2 \). So for each realization of demand, the monopolist sells exactly one-half of the competitive equilibrium industry output.

The difference between the variance of the monopolist’s price distribution and variance of the competitive equilibrium market price distribution is given by

\[
\text{var}[p'] - \text{var}[p''] = \int_p^\epsilon p^2 \frac{q^2(p)}{Q^2(\mathbb{P})} dp - \left( \int_p^\epsilon q(p) dp \right)^2 - \int_p^\epsilon p^2 \frac{q''(p)}{Q^2(\mathbb{P})} dp - \left( \int_0^\epsilon \rho''(e) \frac{D(p''(e))}{Q''(\mathbb{P})} de \right)^2 - \int_0^\epsilon \rho''(e) \frac{D(p''(e))}{Q''(\mathbb{P})} de
\]

where \( m \) denotes monopoly and \( c \) denotes perfect competition. By a change of variables, this difference becomes

\[
\text{var}[p'] - \text{var}[p''] = \int_0^\epsilon \rho''(e) \frac{D(p''(e))}{Q''(\mathbb{P})} de - \left( \int_0^\epsilon \rho''(e) \frac{D(p''(e))}{Q''(\mathbb{P})} de \right)^2 - \int_0^\epsilon \rho''(e) \frac{D(p''(e))}{Q''(\mathbb{P})} de
\]

where \( \mathbb{P} = \epsilon(\mathbb{P}, q) \) is the same for both monopoly and perfect competition. Since demand is linear,
\[ D(p^*(e)/Q^*(p)) = D(p(e)/Q(p)), \]

so (A31) can be rewritten as

\[ \text{var}[p^*] - \text{var}[p^*(e)] = \int_0^\infty (p^*(e) - p^*(e))(p^*(e) + p^*(e)) \frac{D(p^*(e))}{Q^*(p)} \, dp^*(e) \]

\[ - \left[ \int_0^\infty (p^*(e) - p^*(e)) \frac{D(p^*(e))}{Q^*(p)} \, dp^*(e) \right] \int_0^\infty (p^*(e) + p^*(e)) \frac{D(p^*(e))}{Q^*(p)} \, dp^*(e). \]

Since \( p^*(e) + p^*(e) \) is increasing in \( e \), it follows that if \( p^*(e) - p^*(e) \) is increasing in \( e \), then

\[ \text{var}[p^*] - \text{var}[p^*(e)] > 0. \]

From (4) and (10), \( p^*(e) - p^*(e) = D(p^*(e)/Q^*(e)), \) so

\[ \frac{dp^*(e)}{de} - \frac{dp^*(e)}{de} = \frac{[D_p(p^*(e))^2 - D(p^*(e))D_{pp}(p^*(e))] \, dp^*(e)}{(D_e(p^*(e)))^2} . \]

which is clearly positive when demand is linear, because \( D_{pp} = 0 \). More generally, (A33) is positive when \( (D_e)^2 - DD_{pp} > 0 \), for all \( p \), that is, a monopolist will not pass on all of its cost increases. Q.E.D.

References


© RAND 1999.