Consumer Search, Rationing Rules, and the Consequence for Competition

Christopher S. Ruebeck*

Department of Economics, Lafayette College
Easton PA, USA 18042
ruebeckc@lafayette.edu
http://sites.lafayette.edu/ruebeckc/

Abstract. Firms' conjectures about demand are consequential in oligopoly games. Through agent-based modeling of consumers' search for products, we can study the rationing of demand between capacity-constrained firms offering homogeneous products and explore the robustness of analytically solvable models' results. After algorithmically formalizing short-run search behavior rather than assuming a long-run average, this study predicts stronger competition in a two-stage capacity-price game.

Keywords: contingent demand, agent-based modeling, random arrival.

1 Introduction

The residual demand curve, or contingent demand curve, refers to the firm's conjectured demand given its assumptions about consumers and other firms' behaviors. Rather than discussing firms' beliefs about each other's actions, this paper investigates our more basic assumptions about customers' arrivals when firms are capacity-constrained. That is, it focuses on the allocation of customers between firms, the "rationing rule." The literature has shown that duopoly outcomes hinge on these assumptions, and we will see that previous characterizations of aggregate consumer behavior may have unrecognized ramifications.

When Bertrand [3] critiqued the Cournot [5] equilibrium in quantities, he asserted that two firms are enough to create the efficient market outcome: if the strategic variable is price rather than quantity, firms will undercut each other until price equals marginal cost. To instead reinforce the Cournot "quantity" equilibrium as abstracting from a "capacity" decision, Kreps and Scheinkman [7] showed that a two-stage game can support price competition with a capacity choice that is the same as the one-stage Cournot quantities.

This paper's discussion follows Davidson and Deneckere's [6] demonstration that the Kreps and Scheinkman result depends on the assumption that rationing is efficient. They characterized efficient rationing as one extreme, with Beckmann's [2] implementation of Shubik's [11] alternative, proportional rationing, occupying the

other extreme—although we will see behavior below that is outside those proposed extremes. Their analytic and simulation results cover rationing rules in general, and show that only the efficient rationing rule can provide the Kreps and Scheinkman result, arguing in addition that firms will not choose the efficient rationing rule if the two-stage game is extended to three stages that include the choice of rationing rule.

The theory discussed thus far was developed for homogeneous goods, and it has been expanded to include asymmetric costs (Lepore [8]) and differentiated goods (Boccard and Wauthy [4]), but the results below are not about differentiation or variation in costs. Instead, I model search in which the consumer will stop once a low enough price is found at a firm with available capacity. Although this abstracts from the reasons for consumers' need to search, doing so provides a baseline for the future both in terms of algorithms and results.

2 Rationing Rules

With two firms producing homogeneous goods in a simultaneous stage game, we study Firm 2's best response to Firm 1's decision of price and capacity. Both firms may be capacity constrained, but our discussion focuses on Firm 2's unconstrained price choice as it considers Firm 1's constrained decision. Because Firm 1 cannot satisfy the entire market at its price p_1, Firm 2 is able to reach some customers with a higher price $p_2 > p_1$. The rationing rule specifying which customers buy from each firm will be a stochastic outcome in the simulations to follow, but has not been treated that way in theoretical work thus far.

The shape of the efficient rationing rule, shown in Figure 1A, is familiar from exercises 'shifting the demand curve' that begin in Principles of Economics. This rationing rule's implicit assumption about allocating higher-paying customers to the capacity-constrained firm has been known at least since Shubik's discussion of the efficient rationing rule [12] [11]. The demand curve $D(p)$ shifts left by the lower-priced firm's capacity k_1, as depicted by the figure's darker lines. Recognizing that Firm 2 can also be capacity-constrained, residual demand when $p_2 > p_1$ is

$$D(p_2|p_1) = \min(k_2, \max(0, D(p_2) - k_1)) \ . \tag{1}$$

The rationing of customers between the two firms is indicated by the figure's lightly dashed lines, showing that those units of the good with highest marginal benefit are not available to the higher-priced firm.

Although the efficient rationing rule would appear to be an attractive characterization of a residual demand curve "in the long run," firms are typically making decisions and conjectures about each other's prices in the short run. The long run outcome occurs as a result of firms' shorter-term decisions, their long-run constraints, their anticipations of each other's decisions, and the evolution of the market, including both consumers' and firms' reactions to the shorter-term decisions. Game theoretic characterizations of strategic behavior in the short run are important in determining the long run, as the market's players not only react to each other but also anticipate each other's actions and reactions. Yet,
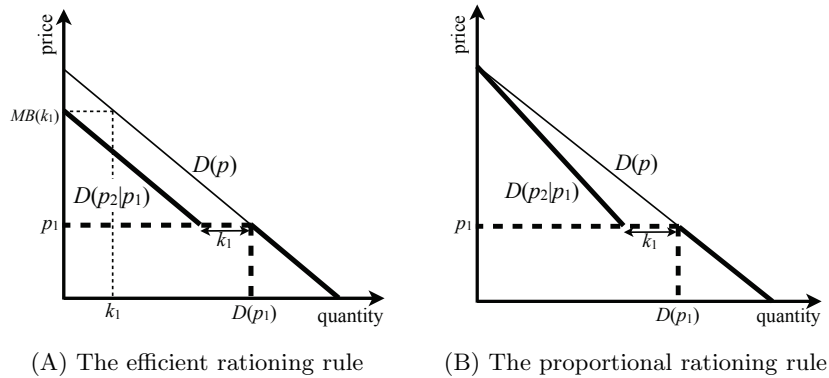
(A) The efficient rationing rule   (B) The proportional rationing rule

Fig. 1. Standard rationing rules $D(p_2|p_1)$ when Firm 1 chooses capacity k_1 and prices at p_1 in a homogeneous good market described by demand curve $D(p)$

in assuming the efficient rationing rule, the level of abstraction may penetrate further in some dimensions of the problem than others, leading to an unattractive mismatch between the short run one-shot game and long run outcomes. After discussing proportional rationing's residual demand curve, the discussion and simulations below will show that even the short run abstraction may be mismatched with the descriptive reasoning behind the rationing rule.

Both the proportional residual demand curve depicted in Figure 1B and the efficient demand curve in Figure 1A match market demand $D(p)$ for $p_2 < p_1$: they both take Bertrand's assumption that all customers buy from the lower-priced firm if they can. The simulation results below will not have the same characteristic. In the case of a "tie", $p_2 = p_1$, the literature makes two possible assumptions on the discontinuity. Davidson and Deneckere assume a discontinuity on both the right and the left of (a.k.a. above and below) p_1 with the two firms splitting demand evenly, up to their production capacities. Allen and Hellwig [1] assume a discontinuity only to the left (below) p_1, pricing at equality by extending the upper part of the proportional rationing curve. It may be more accurate to characterize the upper part of the proportional rationing curve as an extension of the outcome when prices are equal, a point to which I will return below.

The key feature of the proportional rationing rule is that all buyers with sufficient willingness to pay have a chance of arriving at both firms. That is, now all consumers with willingness to pay at or above $p_2 > p_1$ may buy from Firm 2. Thus in Figure 1B the market demand curve is "rotated" around its vertical intercept (the price equal to the most any customer will pay) rather than "shifting" in parallel as in Figure 1A. Describing the realization of this contingent demand curve between p_1 and the maximum willingness to pay is not as intuitive as it is for the efficient contingent demand curve, although the highest point is easy: the maximum of of $D(p)$ must also be the maximum of $D(p_2|p_1)$ because both firms have a chance at every customer.

To justify the $D(p_2|p_1)$ proportional demand curve at other prices $p2 > p1$, the literature describes some specifics on consumers' arrival processes and then connects these two points with a straight line, so that "the residual demand at any price is proportional to the overall market demand at that price" (Allen and Hellwig), or as Beckmann first states it, "When selling prices of both duopolists are equal, total demand is a linear function of price. When prices of the two sellers differ, buyers will try as far as possible to buy from the low-price seller. Those who fail to do so will be considered a random sample of all demanders willing to buy at the lower price."

The proportion adopted draws directly from the efficient rationing rule: the fraction of demand remaining for Firm 2 if Firm 1 receives demand for its entire capacity (up to Firm 2's capacity and only greater than zero if market demand is greater than Firm 1's capacity). Formally, the proportional contingent demand for Firm 2 when $p_2 > p_1$ is

$$D(p_2|p_1) = \min\left[k_2, \max\left[0, \frac{D(p_1) - k_1}{D(p_1)}D(p_2)\right]\right].$$

(2)

The key feature of this rationing rule is the constant fraction $\frac{D(p_1) - k1}{D(p_1)}$ that weights the portion of market demand Firm 2 receives at price p_2. First, this fraction is the demand firm 2 receives when prices are equal, and second this fraction of market demand $D(p_2)$ remains constant as Firm 2 increases price. The shape of the residual demand curve below will show that neither of these features need be true as we consider consumers' arrival.

3   Implementing Random Arrival

We now turn to implementing consumer arrival and search rules. For the efficient rationing rule, consumers would be sorted according to the discussion of Figure 1A. They would not arrive in random order: all customers with willingness-to-pay (WTP) greater than the marginal benefit at k_1, $MB(k_1)$, are sent to the capacity constrained, lower-price firm. Somehow these customers rather than others get the good, perhaps by resale (Perry [9]), but it is not the point of this investigation to model such a process. Instead, we will have consumers that arrive in random order—one of the assumptions of the conventional proportional rationing rule, but just part of it. The other part is that the customers that arrive first are always allocated to the lower-priced firm. Relaxing this second part of the assumption is what drives the new results below.

3.1   Random Arrival, Random First Firm

Buyers arrive in random order at two capacity-constrained sellers. As a buyer arrives he chooses randomly between sellers; if the first seller's price is higher than the buyer's WTP, the buyer moves on to the other one, or does not buy at all if both prices are higher than his WTP. As each later buyer arrives, she can only consider sellers that have not yet reached their capacity constraints.

As in the analytic models described above, we take the perspective of Firm 2 given Firm 1's chosen price and capacity. Labeling one firm as "1" does not indicate that buyers in the simulation arrive first at either location; each buyer has an equal chance of arriving first at either Firm 1 or Firm 2.

Figure 2A presents an example of the simulation results. Each of the lines in the plot (some overlapping each other) shows a statistic of 1300 runs; the statistics are min, max, mean, and median quantity for Firm 2 at each price p_2. The plot overlays six 1300-run sets, capturing the stochastic residual demand curve facing Firm 2 with no capacity constraint while the other firm sets price $p_1 = 9$ and capacity $k_1 = 7$. Market demand is $q = 24 - p$. Each of the 23 buyers has unitary demand at price $D_{23} = 23, D_{22} = 22, \ldots, D_1 = 1$. The proportional rationing rule is also shown; It is piecewise-linear, connecting the outcomes for $p = 0, 8, 9$, and 24. (The line segment connecting $p = 8$ and 9 reflects the discrete nature of demand in the simulation.) The efficient rationing rule is not shown, but it would maintain the slope of the lower line from $p_2 = p_1 = 9$ all the way to the vertical axis. The curved line in the plot is the mean residual demand received by Firm 2. As foreshadowed above, it is to the right not only of the efficient rationing rule, but the proportional rationing rule as well. The median is the series of stair steps around that line.
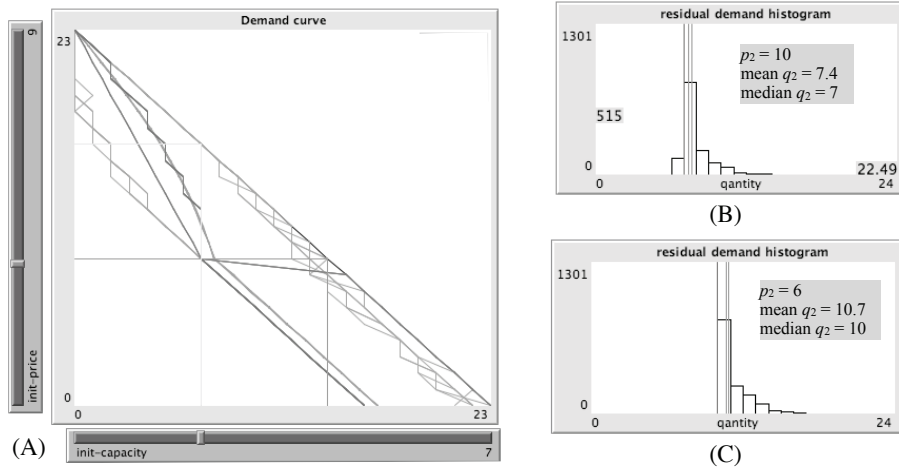

(A)   (B)   (C)

Fig. 2. Simulation results. Firm 2's residual demand curve with no capacity constraint and market demand $q = 23 - p$. Firm 1 has price $p_1 = 9$ and capacity $k_1 = 7$

Consider first the plot's horizontal extremes. Across the six 1300-run sets, Firm 2's contingent demand has multiple maxima and minima q_2 (horizontal axis) for some p_2 (vertical axis) values. Because the 1300 simulated runs are not always enough to establish the true maxima and minima for all customer arrival outcomes, the distribution's tail is (i) thinner in the minima for higher p_2 in the upper left of the plot and (ii) thinner in the maxima for lower p_2 in the lower

right of the plot). Why? (i) The minimum q_2 at any p_2 is 7 units to the left of the market demand curve (but not less than 0), reflecting the capacity constraint of Firm 1. Thus the minimum contingent demand for Firm 2 is less likely to occur when its price p_2 is high because it is less likely that all high-value buyers arrive first at Firm 2 when there are few possible buyers. (ii) The theoretical maximum at each price p_2 is the entire demand curve: by chance all the customers may arrive first at Firm 2. It is more difficult to reach that maximum in a 1300-run sample at lower prices p_2, when there are more potential customers.

Returning to the central part of the distribution, the mean and median lines, we see that—unlike the assumptions behind Figure 1B—the residual demand curve's distribution is not always centered on the interval between its minimum and maximum. Figure 2B and Figure 2C illustrate the skewed distributions for two of Firm 2's prices, $p_2 = 6$ and 10. With these overall differences in hand, now consider, in the three areas of the graph, the differences in these simulations as compared to the theory pictured in Figure 1.

In the lower part of the curve, the median is actually equal to the minimum for all prices $p_2 < p_1$. The mean is pulled away from the median by the upper tail of the skewed distribution. The explanation for the observed behavior in the extremes applies here. Note, too, in the lower part of the curve that demand is smaller than the usual homogeneous goods assumption: The lower-priced firm does not capture the entire market. This is a consequence of maintaining the random arrival assumption for all prices in the simulation, not just for $p_2 > p_1$ as usually assumed (following Bertrand) in the analytical work described above.

At equal prices $p_2 = p_1$, the outcome is continuous in p_2, a point to consider again below. Finally, we saw that in the upper part of the curve $p_2 > p_1$, the mean residual demand curve no longer changes linearly with price: the two firms share the market with varying proportions rather than the usual constant proportion assumption. Note that the traditional proportional residual demand curve (Davidson and Deneckere) provides Firm 2 with a "better" result than that provided by the efficient residual demand curve because it provides a larger share of the market. The results in Figure 2 are even better for the higher-priced firm! Firm 2's proportion of the market grows as p_2 increases, rather than remaining a constant proportion, because Firm 1's capacity constraint improves the chance that more of the higher-value customers may happen to arrive first at Firm 1 instead of Firm 2.

3.2   Random Arrival, to the Lowest Price If Constraint Unmet

Next consider a specification of behavior that lies midway between the proportional rationing rule and the results in Figure 2. In this case, consumers will continue to search if they do not receive a large enough consumer surplus, $WTP - p$. This specification could be viewed as redundant because we have already specified both that consumers have a willingness-to-pay and that they do not know or have any priors on the firms' prices. It may make more sense to instead specify those priors or updating of priors; I will address this further in the closing section.

In Figure 3, the residual demand curve changes smoothly from the varying-proporition in Figure 2 to the proportional demand curve. When $p_2 > p_1$, the choice of some consumers to continue searching pulls the demand curve back towards the proportional prediction. When $p_2 < p_1$, the choice to continue searching pushes the demand curve out towards the usual homogeneous goods assumption. At $p_2 = p_1$, we have a discontinuity for any search threshold that is not zero (the case of Figure 2).
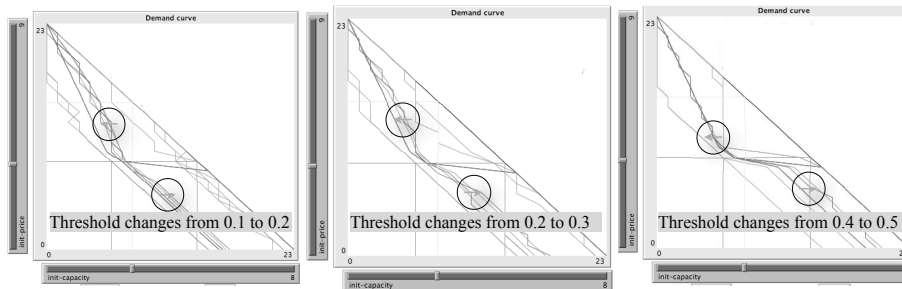

Fig. 3. Continue search when consumer surplus is not a large enough fraction of WTP

4   Conclusions and Further Work

Allen and Hellwig state that consumers go to firms on a "first-come first-served basis" and that, "Firm j with the higher price p_j meets the residual demand of those consumers—if any—who were unable to buy at the lower price." The results above focused directly on a first-come first-served rule, with the result that some consumers with $p_i <$ WTP $< p_j$ may instead arrive too late to be served by the lower-priced firm and thus not buy. The resulting contingent demand curve provides a varying proportion of market demand at price p_j.

This approach recognizes that the searching consumers may know very little about the prices available before they arrive at the firms, and this assumption is unusual in search models. The limited literature that considers this type of search includes Rothschild [10] and Telser [13], but their analyses still require rather strong assumptions to arrive at a model amenable to analytic methods. Telser's investigation finishes with the conclusion (emphasis added), "If the searcher is ignorant of the distribution, then acceptance of the first choice drawn at random from the distribution *confers a lower average cost* than more sophisticated procedures for a wide range of distributions. In most cases these experiments show that it simply does not pay to discover and patronize the lower price sellers . . . [and] we face the problem of explaining how a seller wishing to stress lower prices than his rivals can attract customers."

Thus, one direction to continue this work is to investigate the effect of this demand curve on equilibrium outcomes, moving from a one-shot static analysis to a dynamic interaction. Davidson and Deneckere show that any contingent

demand curve different from the efficient demand curve must provide more market power (a larger quantity effect) to the higher-priced firm when the other firm is capacity-constrained. Thus firms in the first stage have greater incentive to avoid being capacity-constrained and choose capacities greater than the Cournot quantity, and leading to an equilibrium price support in the second-stage game that is closer to the Bertrand outcome of marginal cost pricing.

The results of these simulations show that the competitive pressure is even stronger when we take seriously the assumptions underlying uninformed consumer serach. Setting a price above one's competitor leads to less loss of demand, so that the market power that remains for the higher-priced firm is even larger than in the case of proportional rationing. At the same time, undercutting is less profitable: the lower-priced unconstrained firm does not receive all demand as assumed generally in the literature on homogeneous products. Thus we can expect that further analysis in a dynamic setting will provide outcomes that are even more competitive than found in existing analytical models.

References

1. Allen, B., Hellwig, M.: Bertrand-Edgeworth duopoly with proportional residual demand. International Economic Review 34(1), 39–60 (1993)
2. Beckmann, M.: Edgeworth-Bertrand duopoly revistied. In: Henn, R. (ed.) Operations Research Verfahren III, Meisenhein. Verlag Anton Hain, Meisenhein (1965)
3. Bertrand, J.: Théorie mathématique de la richesse sociale (review). Journal des Savants 48, 499–508 (1883)
4. Boccard, N., Wauthy, X.: Equilibrium vertical differentiation in a Bertrand model with capacity precommitment. International Journal of Industrial Organization 28(3), 288–297 (2010)
5. Cournot, A.: Researches into the Mathematical Principles of Wealth. Macmillan, London (1838)
6. Davidson, C., Deneckere, R.: Long-run competition in capacity, short-run competition in price, and the Cournot model. The RAND Journal of Economics 17(3), 404–415 (1986)
7. Kreps, D., Scheinkman, J.: Precommitment and bertrand competition yield Cournot outcomes. The Bell Journal of Economics 14(2), 326–337 (1983)
8. Lepore, J.: Consumer rationing and the Cournot outcome. B.E. Journal of Theoretical Economics, Topics 9(1), Article 28 (2009)
9. Perry, M.: Sustainable positive profit multiple-price strategies in contestable markets. Journal of Economic Theory 32(2), 246–265 (1984)
10. Rothschild, M.: Searching for the lowest price when the distribution of prices is unknown. The Journal of Political Economy 82(4), 689–711 (1974)
11. Shubik, M.: A comparison of treatments of a duopoly problem. Econometrica 23(4), 417–431 (1955)
12. Shubik, M.: Strategy and Market Structure: Competition, Oligopoly, and the Theory of Games. John Wiley & Sons, New York (1959)
13. Telser, L.: Searching for the lowest price. American Economic Review 63(2), 40–49 (1973)