The Benefits of Being Small: Duopolistic Competition with Market Segmentation*

CHAIM FERSHTMAN  
Department of Economics, Tel Aviv University, Tel Aviv, 69978, Israel

and

EITAN MULLER  
The Leon Recanati, Graduate School of Business Administration, Tel Aviv University, Tel Aviv, 69978, Israel

Abstract. Consider a duopolistic market in which consumers are not necessarily aware of the firms' existence. The market is characterized by the existence of four segments: a duopolistic segment which consists of consumers who are aware of both firms, a segment of consumers who are unaware of either firm and two captive market segments. We assume that by advertising, firms control the proportion of consumers who are aware of their existence. The relative sizes of the four segments affect the equilibrium of the duopolistic pricing game. We show that being large may be disadvantageous, and that even if gaining awareness is costless firms may wish to remain small.

Key words: awareness, segmentation, duopoly.

I. Introduction

The standard assumption in the economic analysis of markets is that although consumers might have incomplete information regarding prices, products' quality or other attributes, they are always aware of the existence of the different firms operating in the market. This assumption simplifies the analysis, but assuming that consumers are always aware of all the different firms and brands in the market is too restrictive. On the other hand analyzing consumer awareness gives rise to a number of conceptual difficulties. In particular, it is of importance to find the specific relationship between the awareness profile and the demand facing the firm.

The issue of awareness has been investigated in the marketing literature and culminated in the concept of the 'choice set'. The consumer, when comparing alternatives such as brands, does not take into account the set of all the brands in the market but a much smaller set, that is known as the choice set or evoked set of that consumer (see Kotler, 1988, pp. 195–97). Making the consumer aware of a product in our setting, therefore, involves more than just cognition; it is the comprehension of the attributes of brand and insertion of this brand into the consumer's evoked set.
When the firm is the sole producer in the market and gaining awareness is costless, it is clear that the firm will want the highest possible awareness such that all its potential clients are aware of its existence.

The main claim of this paper is that the above intuitive result cannot be extended to duopolistic markets. That is, there is a tradeoff between size and profitability: larger size may result in lower profits even though increasing size is costless.

The general belief of the existence of managerial diseconomies of scale is the main driving force of the literature claiming that "small is beautiful". The claim that a larger firm is harder to manage can be traced back to Kaldor (1934), while in his seminal paper Coase (1937) argued that the optimal size of a firm is determined by the comparative cost of internal transaction versus the cost of market transaction. There are several sources of such organization diseconomies. Calvo and Wellisz (1978), and Rosen (1982) discuss the cost of monitoring and its effect on effort; Williamson (1967) and Geanakoplos and Milgrom (1985) consider the limits to the capacity of managers to process information. Rotemberg and Saloner (1991) show that a narrow strategy enables the firm to motivate its employees to search for valuable innovation. McAfee and McMillan (1990) discuss the distorted incentives given by principals in a multi-layer hierarchy. Milgrom (1988) and Milgrom and Roberts (1988) consider the workers' incentives to influence their superiors' decisions and the inefficiencies associated with the need to impose mechanisms that offset such a behavior.

While the above literature concentrates on the internal reasons for the advantages of being small, we emphasize in our paper the fact that external and strategic aspects can lead to the conclusion that small is beautiful.

The issue of consumer awareness has been discussed by Fudenberg and Tirole (1984). They consider a two-stage game such that in the first stage the incumbent firm determines its awareness level and in the second stage, when entry occurs, both firms share the remaining consumers at a fixed investment in awareness. In discussing the incentives the incumbent firm has to invest in awareness, Fudenberg and Tirole pointed out that a high level of awareness might be disadvantageous for the incumbent firm as it reduces its incentives to respond aggressively to its competitor. This result contradicts previous results in which it was argued that the incumbent firm over-invests in order to deter entry (see, for example, Spence, 1977 and Dixit, 1979). It is important to note, however, that, in this setting, if gaining awareness is costless, the optimal choice of the incumbent is still full awareness (100%) since by choosing this level entry becomes impossible. To achieve an interior solution, Fudenberg and Tirole assumed that the cost of getting full awareness is infinity.

In this paper we consider a duopolistic game in which firms compete in prices and advertising and consumers are not necessarily aware of the firms' existence. Thus, in principle, there are four different market segments. Each firm has a captive market segment of consumers who are aware of only one firm; there is a duopolistic segment consisting of consumers who are aware of both firms; and
there is a segment of consumers who are not aware of either firm. The level of awareness for each firm is determined by advertising. Advertising by one of the firms changes the size of all four segments: the size of its own captive segment increases, the duopolistic segment increases, but the captive market of its competitor as well as the segment of consumers who are not aware of either firm decreases. Thus, in determining its advertising policy firms should take into account the effect of all these changes on the pricing game and on the final equilibrium profits. In particular, if a firm increases its advertising it makes its rival more aggressive in the pricing game as it reduces the size of its rival's captive market segment and increases the duopolistic segment. We show that even if gaining awareness is costless, firms might wish to remain 'small', i.e., to have less than 100% awareness.

It is important to relate our model to the growing recent literature on competition with switching cost (see, for example, Farrell and Shapiro, 1988; Klemperer, 1987: von Weizacker, 1984, and Green and Scotchmer, 1986). This literature relates to the costs that consumers must pay if they switch from one firm to another. This phenomenon, which is usually referred to as 'lock in', 'inertia', or 'switching cost', has two effects: firms have an incentive to exploit their captive market, but at the same time the competition to capture new buyers intensifies. If price discrimination is impossible, firms formulating their pricing strategy must make a compromise between these two conflicting effects. The difference between the switching cost model and ours can be highlighted via Figure 1.

In contrast to the four market segments of the awareness model, in the switching cost model there are three market segments: two captive segments and a segment of newcomers that need to decide from which firm to buy, and thus which of the two captive market segments to join. In the awareness model once a consumer joins the duopolistic segment he remains in it. Switching, in this model, from one captive segment to another is not feasible while in the switching cost model it is
feasible (but costly). Another important difference is that in the awareness model $Z_1, Z_2$ and $Z_3$ are determined by advertising and the final market share is determined by the size of $Z_1$ and $Z_2$ and the division of $Z_3$ between the two firms, which depends on the firms’ prices. In the switching cost model market shares are determined by the segmentation $Y_1$ and $Y_2$, their relative size being determined by prices.

The paper is organized as follows. In Section II we present the model and elaborate on the relationship between awareness levels and the segmentation of the market. In Section III we discuss the pricing game. In Section IV we show that even when awareness is costless firms might choose not to be at the maximum awareness level. We discuss a simple example in Section V.

II. Awareness and Duopolistic Markets

Consider a duopolistic industry in which firms sell a differentiated product to $N$ identical consumers. Let us assume that these consumers are not necessarily aware of the existence of the two firms. Thus, although we let consumers be identical in terms of their preferences and endowments, they can differ in terms of the information they have. Specifically, consumers might have different evoked sets, i.e., they are aware of different sets of firms. We further assume that consumers become aware of a specific firm by being exposed to its advertising.

The firms in our model are competing along two dimensions: awareness and prices. We denote the number of consumers who are aware of firm $i$ as $B_i$. Clearly, $B_i \leq N$. Consumers can now be categorized into four groups:

(i) $Z_1$: the number of consumers who are aware only of firm 1 (captive market of firm 1);

(ii) $Z_2$: the number of consumers who are aware only of firm 2 (captive market of firm 2);

(iii) $Z_3$: the number of consumers who are aware of both firms (duopolistic segment);

(iv) $Z_4$: the number of consumers unaware of either firm.

Clearly, $Z_1, \ldots, Z_4$ are mutually exclusive and exhaustive. A consumer can only belong to one of these groups. Before proceeding, we would like to establish the relationship between $B_i$ and $Z_i$. We assume that potential customers are exposed to advertising at random so that the probability of a customer being aware of firm 2 is independent of his awareness of firm 1.

The probability that a randomly selected customer is aware of firm $i$ is $B_i/N$. Given the independence assumption, the probability that a random consumer is exposed to both firms is $B_1B_2/N^2$. Similarly, the probability that a randomly selected consumer is unaware of either firm is $(1 - B_1/N)(1 - B_2/N)$. Thus, given that there is a large number of consumers, it is evident that
\[ Z_1 = B_1(N - B_2)/N; \quad Z_2 = (N - B_1)B_2/N; \]

\[ Z_3 = B_1B_2/N; \quad Z_4 = (N - B_1)(N - B_2)/N. \]

Notice that \( Z_i \) is not directly under the control of one particular firm. Each firm can only affect the level of awareness of its own product, \( B_i \). The different segments are determined jointly by the level of awareness to both firms.

When firm 1 advertises and increase the level of awareness of its product, \( B_1 \), rise, affecting the distribution of consumers among the four segments. An increase in \( B_1 \) causes an increase in \( Z_1 \) and \( Z_3 \) and a decrease in \( Z_2 \) and \( Z_4 \). We can describe this process as follows. Consumers who were in \( Z_4 \), by being exposed to the new advertising, become aware of firm 1 and thus become part of \( Z_1 \). Consumers that were captive consumers of firm 2 now become aware of firm 1 and shift to the duopolistic segment \( Z_3 \). Intuitively, these shifts look favorable for firm 1 since its exposure increases and in particular the size of its captive market increases.

Since the consumers in the first group are captive customers of firm 1, this firm has complete monopolistic power with respect to these customers. Let \( p_1 \) be the price charged by the first firm. We assume that each customer in the first segment buys \( a(p_1) \) units of the product, where \( a(p_1) \) is a standard downward sloping demand function. The total demand of this segment is \( Z_1a(p_1) = (B_1(N - B_2)/N)a(p_1). \) Similarly for firm 2, given its price \( p_2 \), its captive customers buy \( Z_2a(p_2) = (B_2(N - B_1)/N)a(p_2). \) The consumers in the duopolistic segment are aware of both firms. Given the firms' prices, we assume that the demand facing each firm is \( (B_1B_2/N)b_i(p_1, p_2). \) Thus, given \( (B_1, B_2) \) and the firms' prices, the demand facing each firm is:

\[ D_i(B_1, B_2, p_1, p_2) = \frac{B_i(N - B_i)}{N}a(p_i) + \frac{B_iB_j}{N}b_i(p_1, p_2) \quad (1) \]

We model a two-stage game in which, in the first stage, the firms set the level of awareness by advertising, and in the second stage there is a pricing game. We start our analysis by discussing the second stage of the game.

### III. The Pricing Game

Given awareness levels of \( (B_1, B_2) \) and assuming for convenience zero production cost, the firm's profit function is:

\[ \Pi_i(B_1, B_2, p_1, p_2) = p_i\frac{B_i(N - B_i)}{n}a(p_i) + p_i\frac{B_iB_j}{N}b_i(p_1, p_2) \quad (2) \]

The firm can thus be thought of as having two types of markets, a captive segment and a duopolistic segment. If it is possible to identify the consumers of each group and to price discriminate, the firm will charge its monopolistic price
\( p_1^m = \text{Argmax}_{p_1} a(p_1) \) in its captive segment and \( p_2^m(p_2) = \text{Argmax}_{p_2} p(b(p_2, p_2)) \) in its duopolistic segment. However, we assume that such a strategy is not feasible as firms cannot identify the type of the consumer they are facing, or alternatively, they cannot price discriminate.

We assume concavity of the revenue functions \( p_1 a(p_1) \) and \( p_2 b(p_2, p_2) \), which guarantees the concavity of the profit function (2). The first-order conditions are therefore:

\[
(N - B_j)[a(p_i) + p_i a'(p_i)] + B_j[b(p_i, p_i) + p_i b'(p_i, p_i)/\partial p_i] = 0. \tag{3}
\]

Equation (3) can be regarded as the \( i \)'th firm's reaction function and it is a linear combination of the monopolistic reaction function (i.e., its f.o.c.) and the reaction function of a regular duopolist. Observe that \( B_j \) does not affect the reaction function of firm \( i \) in the second stage pricing game. Thus unlike most of the literature (see, for example, Brander and Spencer (1983), Bulow, Geanakoplos and Klemperer (1985), Shapiro (1989), Fershtman and Judd (1982), Vickers (1985), and Sklivas (1987), among many others), the investment in awareness in the first stage of the game does not serve as a mean of precommitment for the second stage of the game. Rather, firm \( i \), by choosing \( B_j \) affects the reaction functions of firm \( j \) and thus affects firm \( j \)'s behavior in the duopolistic pricing game.

Let \( \phi_i(p_2, B_2) = \text{Argmax}_{p_2} \Pi_i(B_1, B_2, p_1, p_2) \) be the reaction function of firm 1 as a function of the awareness of firm 2. When \( B_2 = 0 \) the duopolistic segment disappears and the firm faces only its captive customers. The size of this captive segment depends on \( B_1 \) and the optimal price is the monopolist price, i.e., \( \phi_1(p_2, 0) = p_1^m \).

For \( B_2 = N \), the captive segment disappears. The firm faces a duopolistic market and its reaction function, \( \phi_i(p_2, N) \) is identical to that of a regular duopolistic firm. As \( B_2 \) increases we can see from Equation (3) that the emphasis of firm 1 shifts from its captive market segment to its duopolistic segment, and thus an increase in \( B_2 \) pushes the reaction function of firm 1 towards the duopolistic reaction function \( \phi_1(p_2, N) \).

The possibility that a firm can control the reaction function of its competitor introduces a new and interesting dimension to the strategic interaction. In determining the degree of awareness of its product, the firm not only affects the different segments it faces but also affects the strategic behavior of its competitor in the second stage of the game.

Let \( \Pi_i' = \delta \Pi_i / \partial p_i \). Assuming that \( \Pi_1[1] \Pi_2^{22} > \Pi_1[2] \Pi_2^{21} \) guarantees that there is a unique intersection point of the two reaction functions \( \phi_i(p_2, B_2) \) and \( \phi_2(p_1, B_1) \). This intersection point defines the equilibrium prices. For a given \( B_1 \) and \( B_2 \) we let \( p_1^e = h_1(B_1, B_2) \) and \( p_2^e = h_2(B_1, B_2) \) be the unique equilibrium prices for the awareness levels \((B_1, B_2)\). When \((B_1, B_2) = (N, N)\) there is a standard duopolistic competition in which no firm has a captive segment and the equilibrium is identical.
to the standard Nash equilibrium of a price competition in a differentiated good duopoly.

Lemma 1: (i) The equilibrium prices \((p_1^*, p_2^*)\) decrease as the levels of awareness \(B_1\) or \(B_2\) increase.

(ii) \(B_2 > B_1\) implies that \(p_2^* > p_1^*\), i.e., the large firm charges a higher price at equilibrium.

Proof: The proof is straightforward, using a simple reaction function analysis (see also the survey on oligopoly behavior by Shapiro (1988)).

IV. Investing in Awareness

Since the equilibrium prices are determined by the levels of awareness, we are now able to express the firms’ profit as a function of \(B_1\) and \(B_2\). Let \(D_i^*(B_1, B_2)\) be the quantity sold by firm \(i\) as a function of the firm’s awareness level, i.e., \(D_i^*(B_1, B_2) = D_i(B_1, B_2, h_i(B_1, B_2), h_2(B_1, B_2))\).

\[
R_i(B_1, B_2) = \Pi_i(B_1, B_2, h_i(B_1, B_2), h_2(B_1, B_2)) = h_i(B_1, B_2)D_i^*(B_1, B_2) = h_i(B_1, B_2)\left[\frac{B_i(N - B_j)}{N}a(h_i(B_1, B_2)) + \frac{B_iB_j}{N}b_i(h_1(B_1, B_2), h_2(B_1, B_2))\right] \quad \text{(4)}
\]

The first stage of the game is now well defined. The players choose the awareness levels \((B_1, B_2)\) and their payoffs are specified by Equation (4). Clearly the existence of an equilibrium in such a game is not guaranteed. Although the concavity of \(R_i(B_1, B_2)\) with respect to \(B_i\) is sufficient for existence, it is not guaranteed by our previous assumptions regarding the demand and the cost functions.

Using Equation (4) we are now able to discuss the optimal awareness level of the firm. Notice that in defining the profit function \(R_i(B_1, B_2)\) we assume that gaining awareness is costless. This assumption is not made only for the sake of simplicity. It is made to emphasize the fact that firms choose partial awareness levels not because creating awareness is costly, but for strategic reasons.

Differentiating Equation (4) with respect to \(B_i\), using the envelope theorem, and evaluating at \(B_i = N\) yield:

\[
dR_i/dB_i = bh_i(B_1, B_2)[1 - \epsilon_i\delta_i] \quad \text{(5)}
\]

where \(\epsilon_i = (\partial b_i/\partial p_j)(p_j/b_i)\) and \(\delta_i = -\left(\partial h_i/\partial B_i\right)(B_i/h_i)\).

Thus, \(\epsilon_i\) is the cross-elasticity of demand and \(\delta_i\) is the percent change in the equilibrium price of firm \(j\) for a percent change of the awareness level of firm \(i\).

Thus, if these two elasticities are large enough (i.e., their product is larger than one) the marginal profit at \(B_i = N\) is negative (see Equation (5)). This can be summarized in the following proposition:
Proposition 1: Full awareness (i.e., $B_1 = N$) may not be the optimal strategy of a firm in a duopolistic market even if creating awareness is costless.

Intuitively, under the assumption of costless awareness one would expect $\partial R_1(B_1, B_2)/\partial B_1$ to be positive. The level of awareness is regarded as something of an asset to the firm and an increase in such an asset should induce higher profits. Analyzing Equation (4), however, yields that intuition might be misleading. An increase in $B_1$ benefits the first firm by providing it with a larger potential market. The captive consumer segment is larger and so is the duopolistic segment. But the increase in $B_1$ causes opposite changes in the market segmentation faced by the second firm. Its captive segment decreases while its duopolistic segment increases. Using the second firm's reaction function (i.e., Equation (3)), yields that an increase of $B_1$ implies that the second firm will put more emphasis on the duopolistic segment, and in order to gain a larger share of this segment the second firm lowers its price. Firm 1 responds by lowering its price and consequently the intersection of the two reaction functions occurs at lower prices. Finally these lower prices may offset the increase in the firm's captive and duopolistic segments, causing a reduction in the firm’s profits.

Note that the reduced form profit function is not necessarily increasing in the firm's own investment. This is true irrespective of the cost function since it is the revenue function of the firm that is non-monotonic in the firm’s investment level. This is in contrast to some of the assumptions in industrial organization literature (see, for example, Spence (1979)).

V. An Example

In this section we present an example that demonstrates our main claim. Let $a(p_i)$ and $b_i(p_1, p_2)$ be defined as follows:

$$a(p_i) = 2 - kp_i$$  \(\text{(6)}\)

$$b_i(p_1, p_2) = (1 - kp_i + (p_i/p_i)^\gamma)/2$$  \(\text{(7)}\)

Clearly, when $p_1 = p_2 = p$ we have:

$$b_1(p_1, p_2) + b_2(p_1, p_2) = a(p).$$

The function $a(p_i)$ is a standard linear demand function. In defining $b_i(p_1, p_2)$ we deviate from the linear case and let the cross-effect term of the demand be $(p_i/p_i)^\gamma$. The parameter $\gamma$ plays a crucial role in our analysis, as will be shortly demonstrated. It is possible to use the same procedure, mutatis mutandis, to show that in the linear case, i.e., when the cross-term effect is $\gamma(p_1 - p_1)$, the equilibrium is $B_1 = B_2 = N$.

If the firms could price discriminate, they would choose a price of $p^{''} = 1/k$ for the monopoly segment of the market and a price of $p^* = (2 - \gamma)/2k$ for the com-
mon., duopoly part of the market. To ensure nonnegativity of the duopoly price we restrict the parameter \( \gamma \) as follows: \( 0 \leq \gamma \leq 2 \). Equation (3), i.e., the first-order condition for the pricing stage of the game now becomes:

\[
\frac{\partial \Pi_i}{\partial p_i} = 2(N - B_i)(1 - kp_i) + (B_i/2)(1 - 2kp_i + (1 - \gamma)(p_i/p_j)^\gamma) = 0.
\]  

(8)

Let \( \eta \) be the price elasticity of demand for the duopolistic segment, i.e., \( \eta = \left( \frac{\partial b_i}{\partial p_i} \right)(p_i/b_i) \).

Define \( \gamma^* \) and \( \gamma^{**} \) as follows. Let \( \gamma^* \) be such that \( \gamma > \gamma^* \) if and only if \( \eta > 1 \) and let \( \gamma^{**} \) be such that for all \( \gamma < \gamma^{**} \) the following inequality holds:

\( \Pi_i^{1/4} \Pi_j^{1/2} > \Pi_i^{1/2} \Pi_j^{1/2} \), where \( \Pi_i' = \frac{\partial \Pi_i}{\partial p_i} \), and \( \Pi_i \) is defined in Equation (2). Note that the condition \( \Pi_i^{1/4} \Pi_j^{1/2} > \Pi_i^{1/2} \Pi_j^{1/2} \) is the standard condition that guarantees the uniqueness of the equilibrium in the pricing game.

Claim i: The two-stage game with Equations (6) and (7) defining the demand functions is such that for all \( \gamma^* < \gamma < \gamma^{**} \), \( (B_1, B_2) = (N, N) \) is not a subgame perfect equilibrium of the game.

Proof: We show in Appendix A (available from the authors) that if \( \gamma^* < \gamma < \gamma^{**} \) when \( B_1 = B_2 = N, \, \partial R/\partial B_i < 0 \). Thus, under our assumptions, \( B_1 = B_2 = N \) is not an equilibrium.

The intuition of this claim is as follows: \( \gamma \) measures the elasticity of the cross-effect term in the duopolistic segment. If this elasticity is large enough, so that the elasticity of demand of the duopolistic segment is large (larger than unitary elasticity), then any increase in firm 1’s awareness \( (B_1) \) will cause the duopolistic market to expand. This will induce the rival (firm 2) to lower its own price, since its price is a convex combination of its monopolistic and oligopolistic prices. Since the weight of the duopolistic market is now larger, firm 2 lowers its price. Since the cross-elasticity is large, this reduction in the rival’s price causes a reduction in profit in the duopolistic segment of firm 1 that is not covered by the increase in profits in its captive segment.

Thus, indeed, firm 1, by increasing its awareness level, applies too much pressure on its rival. Its reaction, coupled with the fact that the cross-elasticity is large, causes this increase in awareness to be unprofitable.

Claim ii: The two-stage game, with Equations (6) and (7) defining the demand function is such that for \( \gamma = 1 \) there exists a unique equilibrium with \( B_i < N \).

VI. Concluding Remarks

The main objective of this paper is to demonstrate that it is not always optimal to be a ‘big’ firm. The intuitive explanation of our result of ‘live and let live’ does not have any cooperative flavor to it. Each firm wishes its rival to have a captive market segment so it will be less aggressive. However, the analysis was carried
out on the assumption that the size of the firm does not affect its feasible strategy set. This is a straightforward assumption when we interpret size as the degree of awareness. Size, however, can describe other characteristics of the organization, some of which can be advantageous to the firm and some of them not. One can imagine a situation in which the size of a firm affects its feasible strategy space or the speed at which it can process information and react to new information. The relationship between a firm's size and its performance in different market structures is of major importance, as it combines the analysis of internal organization and industry equilibrium.

Notes
* We would like to thank Paul Klemperer and an anonymous referee for valuable comments.
† Proof is available from the authors upon request.

References
Greene, J. and S. Scotchmer (1986), 'Bertrand Competition with a Distribution of Switch Costs and Reservation Prices', mimeo, Harvard University.
McAfee, R., Preston and John McMillan (1990), 'Organization Diseconomies of Scale', mimeo.


