Theories of competition and market performance
van Witteloostuijn, A.

Published in:
De Economist

Publication date:
1992

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright, please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 13. Oct. 2019
THEORIES OF COMPETITION AND MARKET PERFORMANCE

MULTIMARKET COMPETITION AND THE SOURCE OF POTENTIAL ENTRY

BY

ARJEN VAN WITTELOOSTUIIJN*

1 STATIC VERSUS DYNAMIC EFFICIENCY

The theory of industrial organization (henceforth IO) studies business policy and market performance under specific competitive conditions. The key contribution of IO is the theory of competition. The structure-conduct-performance paradigm describes the (reciprocal) causalities between competitive conditions (structure), business policy (conduct) and market welfare (performance). The welfare debate focuses on the trade-off between static and dynamic efficiency of partial market results.\(^1\) Static efficiency of market behavior refers to (minimum) average cost pricing, whereas dynamic efficiency is concerned with the fact that, for instance, ‘new products may be introduced, new qualities of existing products may be developed, new methods of production may be ventured, new forms of industrial organization, financing, or tackling risk may be developed’ (Kirzner 1985, p. 30).

The central proposition is Schumpeter's well-known argument that perfect competition undermines the firms' incentives to introduce dynamic economies of market behavior. The point is that 'the fundamental impulse that sets and keeps the capitalist engine in motion comes from the new consumers' goods, the new methods of production of transportation, the new markets, the new forms of industrial organization that capitalist enterprise creates. ... A system — any system, economic or other — that at every point of time fully utilizes its possibilities to the best advantage may yet in the long run be inferior to a system that does so at no given point of time, because the latter's failure to do so may

* Current address: University of Limburg, Faculty of Economics and Business Administration, P.O. Box 616, 6200 MD Maastricht, The Netherlands, Phone: 043-888816. The author is grateful for the comments of Professor S.K. Kuipers and Professor J. Pen. Of course, the usual disclaimer applies.

1 The efficiency terminology is often used in welfare-theoretic arguments in the theory of industrial organization. It is, for example, predominant in Kamien and Schwartz' (1982) excellent survey of the economics of innovation. However, the reader must be aware of the fact that efficiency in the theory of industrial organization has a meaning which is narrower than the one in welfare (particularly general equilibrium) economics (Tirole 1988, pp. 11–12).
be a condition for the level or speed of long-run performance. ... But in capitalist reality, as distinguished from its textbook picture, it is not that kind of [price] competition which counts but the competition among firms from the new technology, the new source of supply, the new type of organization (the largest-scale unit of control for instance) – competition which commands a decisive cost or quality advantage and which strikes not at the margins of the profits but at their foundations and their very lives. ... It is hardly necessary to point out that competition of the kind we now have in mind acts not only when in being but also when it is merely an ever-present threat. It disciplines before it attacks’ (Schumpeter 1943, pp. 83-85). This is the static-dynamic efficiency trade-off, which serves prominently in the literature on R&D and innovation. That is, ‘under monopoly, innovation occurs but at a slower pace than is socially optimal, whereas under perfect competition there is none at all. This of course leads to the consideration of the trade-off between perfect competition and its static efficiency properties and monopoly, which lacks static efficiency but allows for innovation’ (Kamien and Schwartz 1982, p. 191).

This paper serves a twofold purpose. First, the theoretical literature on the static-dynamic efficiency trade-off is reviewed and classified by distinguishing three types of competition: pure-contestability, non-contestability and quasi-contestability (Van Witteloostuijn 1990a). Second, a microfoundation of quasi-contestability as a welfare-theoretic yardstick for the evaluation of market performance is presented: a quest for favorable market conditions identifies sources of potential entry that impose a discipline on incumbent firms' conduct. Note that the paper’s terminology is standard in IO: the argument is framed in terms of game theory (Shapiro 1989, pp. 125-126). The frame of reference is a one (static) or two-staged (dynamic) noncooperative game: in the first stage (ex ante) firms may decide on sunk investment; in the second stage (ex post) rivals compete for market share. The equilibrium concepts are Nash for static and subgame perfection for dynamic games (Rasmusen 1989, pp. 32 and 85, respectively). The paper is organized as follows. Section 2 reviews and classifies theories of competition. Section 3 focuses on static efficiency by investigating the role of sources of potential competition in disciplining pricing policies of incumbent firms. Section 4 studies dynamic efficiency by analyzing entry-deterring investment strategies of incumbent firms in the face of a credible entry threat. Section 5 summarizes the argument.

2 THEORIES OF COMPETITION

2.1 Pure-contestability

2.1.1 Literature
The well-established theory of statically efficient rivalry is perfect competition. A large number of price-taking producers of a uniform product induces average
variable cost pricing if information on prices is perfect (Varian 1984, p. 82). Perfect competition resembles Cournot competition under conditions of low minimum efficient scale and free entry (Novshek and Sonnenschein 1987). Both models assume a large number of firms. This is not the case with elementary Bertrand competition (being associated with assumptions of cost and product homogeneity, and nonbinding capacity restrictions). Elementary Bertrand competition introduces statically efficient price-setting in market configurations with only two incumbent firms (Waterson 1984, p. 25): the crucial assumption is that buyers switch to the lower-priced supplier before the higher-priced rival is able to react.

To date the final generalization of perfect competition is perfect contestability: under conditions of costless exit and free entry even a monopoly operates statically (Pareto or Ramsey) efficient (Baumol 1982). Exit is costless and entry is free if sunk costs are zero (Shepherd 1984, p. 572). The essential feature of a perfectly contestable market is the absence of entry and exit barriers. A contestable market is vulnerable to hit-and-run entry. That is, ‘[e]ven a very transient opportunity need not be neglected by a potential entrant, for he can go in, and, before prices change, collect his gains and then depart without cost, should the climate grow hostile’ (Baumol 1982, p. 4). Perfect contestability describes Bertrand competition with potential rather than actual rivalry.

Schwartz (1986) nicely summarizes the implications of a pervasive entry threat. He argues that ‘the threat of new entry may be sufficient to discipline incumbent firms. In the extreme, benchmark case of perfect contestability, threat of entry ensures satisfactory performance regardless of the size distribution of incumbent firms and regardless of any oligopolistic interactions among them. … To prevent costless hit-and-run entry incumbents must set price where average cost intersects market demand, which maximizes welfare subject to a breakeven constraint’ (Schwartz 1986, pp. 37–38). So, perfect contestability implies static efficiency. Perfect contestability embodies perfect competition as a special case, since average variable cost pricing is the equilibrium strategy irrespective of market structure.

Immediately following Baumol’s (1982) address a barrage of questions arose regarding the merits and robustness of the concept of (perfect) contestability (Brock 1983; Spence 1983; Schwartz and Reynolds 1983; Weitzman 1983; and Shepherd 1984). The purport of the critique is that perfect contestability only represents an ‘odd special case’ (Shepherd 1984, p. 577). The argument is that the first-best results are generated on the basis of very special (restrictive) assumptions, in particular with regard to zero sunk cost (Shepherd 1984, p. 577) and entry and exit lags (Schwartz and Reynolds 1983, p. 488). Brock (1983) clarifies matters by asserting that ‘Dixit (1982), for example, has argued that the economic conditions that must be present for perfect contestability to exist … are extremely stringent: (i) all producers must have access to the same technology, (ii) this technology may have scale economics such as fixed costs, but must not involve any sunk cost, (iii) incumbents can change prices only
with a nonzero time lag, and (iv) consumers must respond to price differences with a shorter lag. Baumol et al. (1983) argue that iii and iv are not needed if entrants can write firm contracts with consumers for delivery over some fixed period length \( t' \) (Brock 1983, p. 1057). That is, the criticism is that perfect contestability relies on restrictive assumptions of firms' homogeneity, zero sunk cost and Bertrand conjectures.

2.1.2 Game

The four theories of statically efficient competition – perfect competition, large-number Cournot competition, elementary Bertrand rivalry and perfect contestability – can be denoted pure-contestability. Essentially, pure-contestability describes a static (one-shot) Bertrand game (Knieps and Vogelsang 1982): the assumption of zero sunk investment implies that the first stage of the game is empty. An example of the ex post game for market share is summarized in Table 1 (where \( P \) is price, \( i \) incumbent firm, \( e \) potential entrant and \( A \) average variable cost). A cell depicts the profit combination \( (\pi^i, \pi^e) \) per strategy pair.

It immediately follows that the combination of incumbent firms' average variable cost pricing \( (P^i = V) \) and potential entrants' non-entry \( (P^e \geq P^i) \) is the unique Nash equilibrium if incumbent firms (weakly) prefer zero-profit production to exit (Grossman 1981, p. 1159). This is true for both the sequential (incumbent firms announce price first) and the simultaneous (incumbent firms and potential entrants decide on price whilst being unaware of the rivals' pricing policy) game.

2.1.3 Performance

Pure-contestability gives static efficiency. For the evaluation of dynamic performance the condition that sunk costs are zero is particularly important. Shepherd (1984) notes that sunk (exit) 'costs include many categories besides physical capital, such as R&D, advertising to establish brand loyalty, and training to create special workers' skills. These intangible forms are often more fully 'sunk' than physical capital, which can be leased or resold' (Shepherd 1984, p. 580). For example, R&D outlays are an important type of sunk investment. Stiglitz (1987) argues that '[m]ost expenditures on R&D are, by their very nature, sunk costs. The resources spent on a scientist to do research cannot be recovered (Stiglitz 1987, p. 928). This means that pure-contestability does not take account of important dynamic economies of market performance:

<table>
<thead>
<tr>
<th>TABLE 1 - PURE-CONTESTABILITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potential entrant</td>
</tr>
<tr>
<td>Incumbent firm</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
although static efficiency is induced, dynamic efficiency lies beyond the scope of the analysis as sunk costs are assumed to be zero.

2.2 Non-contestability

2.2.1 Literature
Dynamic economies of market behavior are inextricably bound up with sunk costs. However, the literature argues that a positive sunk cost introduces an entry barrier (Baumol and Willig 1981; Grossman 1981; Baumol et al. 1982; Farrell 1986; Stiglitz 1987; Dasgupta and Stiglitz 1988; and Martin 1989). Sunk costs follow from the part of capital which is unrecoupable (that is, which has zero selling value) outside the market in which the products that are produced with the capital involved are sold. Broadly speaking, sunk costs follow from irreversible investment in specific capital (Caves and Porter 1977; and Martin 1989). However, it is precisely irreversible investment (or, commitment) which raises entry barriers (Dixit 1982). Stigler (1968) defines entry barriers as 'a cost of producing (at some or every rate of output) which must be borne by a firm which seeks to enter an industry but is not borne by firms already in the industry' (Stigler 1968, p. 67). Examples of entry barriers (Gilbert 1989) are absolute cost advantages (through the introduction of process innovations), scale economies (via the installment of overcapacity) and product differentiation (as a result of advertising campaigns).

This means that a positive sunk cost raises an entry barrier if potential entrants have not yet sunk the cost necessary to eliminate (technology, capacity, goodwill or whatever) disadvantages relative to the incumbent firms. That is, 'it is the sunk costs rather than fixed costs which may deter entry. This is because if one firm 'sinks' some costs, then its (ex post) average cost of production is lower than firms which have not yet sunk costs' (Grossman 1981, pp. 1170-1171). Entry barriers, in their turn, are a necessary condition for imperfect competition. Along the lines of Stigler (1968, p. 16) Waterson (1984, p. 56) argues that '[b]ehind protecting entry barriers monopolistic and oligopolistic rivalry may yield a positive profit.' (An exception is the elementary Bertrand model, which describe a price war that drives profits down to zero.)

2.2.2 Game
The theories of statically inefficient competition can be denoted as non-contestability. The crucial feature of non-contestability is the incumbent firms' sunk cost advantage that raises entry barriers. Broadly speaking, commitment through sunk investment transforms the static game (Table 1) into a dynamic (two-shot) game (Dixit 1982). Incumbent firms can exploit a first-mover advantage: in the first stage incumbent firms decide on sunk investment (S), whereas in the second stage potential entrants consider entry. It is essential that the investment of sunk costs proceeds sequentially: incumbent firms
decide on commitment before potential entrants sink entry costs. The incumbent firms' ex ante sunk investment \( S^i > 0 \) forces entrants to sink \( S^e \) ex post in order to match the incumbent firms' first-mover advantages. Figure 1 summarizes an exemplary two-staged commitment game for a duopoly, assuming free entry \( (S^e = 0) \) for \( S^i = 0 \), monopoly profit \( M \) and post-entry (duopoly) profit \( D \).

The sunk investment is a profitable entry-deterring device if \( M-S^i > D^i \), \( M-S^i > 0 \), \( D^e > 0 \) and \( D^e-S^e < 0 \): the combination of entry deterrence \( (S^i > 0) \) and non-entry \( (S^e = 0) \) is the unique perfect equilibrium. (Note that superiority of incumbent firms may imply that \( S^e > S^i \).) The literature (Gilbert 1989) focuses on the study of profitable commitments (determining the size and content of entry-deterring investment \( S^i \)) and credible threats of post-entry competition (fixing the height of post-entry profits \( D^i \) and \( D^e \)).

### 2.2.3 Performance

Entry barriers have a significant impact on market performance (Bain 1956). The implications of entry barriers for market performance cannot be established in a straightforward way. For example, Schmalensee (1988) points out that (cf. Gilbert 1989, pp. 528–530) 'formal models of imperfect competition rarely generate unambiguous welfare conclusions. In such models, feasible policy options usually involve movements toward but not to perfect competition, so that welfare analysis involves second-best comparisons among distorted equilibria. In particular, there is no guarantee that making markets 'more competitive' will generally enhance welfare, particularly if non-price rivalry is intensified' (Schmalensee 1988, p. 677).

The easy part is the effect on price, which follows directly from the definition of barriers to entry. In terms of Stigler (1968) entry barriers imply that potential entrants face a cost disadvantage: \( T^i < T^e \), where \( T \) denotes average total

![Figure 1 - Non-contestability](image-url)
costs. The margin (\( G \)) between price (\( P \)) and average total costs, \( G = P^i - T^i \), indicates the height of barriers to entry. The entry-deterring (or, limit) price \( P^L \) is \( T^e \). For \( G = 0 \) there is no entry barrier. For \( G > 0 \) (i.e., if an entry barrier is present) the profit (\( \pi \)) of an incumbent firm, \( \pi^i = Q^i - G \) (where \( Q \) is output), is positive (ignoring time and discounting) as \( P^L = T^e > T^i \). Price exceeds average total costs: \( P^i > T^i \). Hence, the essential condition of static efficiency is violated. The very definition of entry barriers implies that in a static and partial context entry barriers induce static inefficiency as established firms can appropriate incumbency rent (Gilbert 1989, p. 478).

However, this is not to say that entry barriers necessarily introduce a welfare loss. The fact that barriers to entry can be closely related to dynamic efficiency introduces a complication. The welfare implications depend critically on the source of the entry barrier. In the literature the normative analysis of entry barriers is generally carried out with the help of a welfare function that consists of a producers' and a consumers' surplus (Tirole 1988, pp. 11-12). Profits are usually regarded as the producers' surplus, whereas the consumers' surplus follows from a utility function. Here it suffices to note that (i) the producers' surplus increases with price and quantity and decreases with average (total) costs and (ii) the consumers' surplus generally increases with product quality, variety of products and quality of information and decreases with price and transaction costs.

The producers' surplus follows from the profit of incumbents and entrants. On the one hand, from the incumbent firms' point of view entry barriers clearly, ceteris paribus, raise welfare. Zero-entry barriers are associated with free entry (pure-contestability) and so with a tendency to profit dissipation (subsection 2.1), whereas barriers to entry are so defined that they yield a positive profit. The higher the barrier to entry, the higher the profit rate which the incumbent firms are facing. On the other hand, entry barriers generally imply a reverse welfare effect on potential entrants, since barriers to entry restrict entry (and profit) opportunities.

From the consumers' perspective the source of the entry barriers is crucial. Take one typical example: advertising and goodwill. Advertising has a good and a bad side. Nelson (1975) argues that 'there is much, indeed, in advertising that is informational: price advertisements, helpwanted advertisements, and pictures of dresses and furniture available in stores ... But surely ..., this is not the whole story. What about the frequent endorsement of a brand by announcers, actors, or celebrities - all of whom are paid for their efforts? It is advertisements of this character that have produced much of the skepticism about advertising's information role' (Nelson 1975, p. 213). So, advertising...

---

2 Conditional of course upon the assumption that potential competition dominates actual rivalry (Shepherd 1984).

3 This is, for example, immediately clear from the extreme case where the incumbent firms' entry-deterring strategies raise the rivals' costs (cf. Salop and Scheffman 1983).
may introduce countervailing effects: if advertising raises an effective entry barrier, the limit price is increased, but transaction costs are lowered and information is increased. If advertising has a positive influence on price, the increased price may raise the incumbent firms' profits, but lowers the consumers' surplus. The net effect of advertising on the consumers' surplus depends upon the quality of the information. So, the debate centers on the quality of the information that is embodied in advertisements (Comanor and Wilson 1979, pp. 472–473). Informative and noninformative advertisements have a different impact (Schmalensee 1978 and 1986).

In effect, as is the case with entry barriers through informative advertising, the trade-off between static and dynamic efficiency is generally at stake (Von Weizsäcker 1980). A clear example is reflected in the study of competition, R&D and innovation (Reinganum 1989). Entry barriers facilitate the profit rate of incumbent firms, but restrict the entry opportunities of potential entrants. The impact on consumers' utility critically depends on the source of the barrier to entry. So, the literature on pure-contestability (with free entry) and non-contestability (with impeded entry) illustrates the static-dynamic efficiency dilemma.

2.3 Quasi-contestability

2.3.1 Literature
Clark (1940) introduced the concept of workable competition in order to replace the unrealistic norm of perfect competition (Clark 1940, pp. 241–242) by a benchmark that constitutes second-best performance: 'One central point may be put abstractly. If there are, for example, five conditions, all of which are essential to perfect competition, and the first is lacking in a given case, then it no longer follows that we are necessarily better off for the presence of any one of the other four. In the absence of the first, it is a priori quite possible that the second and third may become positive detriments; and a workable satisfactory result may depend on achieving some degree of 'imperfection' in these other two factors. ... imperfect competition may be too strong as well as too weak; and ... workable competition needs to avoid both extremes' (Clark 1940, pp. 242–243). Clark presented a list of the ten most important criteria that facilitate workable competition (Clark 1940, pp. 243–244). Moreover, he pointed to '[m]odified, intermediate or hybrid competition. ... The most important cases involve formally free entry, but no exit without loss' (Clark 1940, p. 245).4

In 1958 Sosnick pointed out that eighteen authors have listed criteria of

---

4 It is interesting to note that Clark indicates that 'the more attention centers on the imperfections of active competition, the more important become the forces of potential competition' (Clark 1940, p. 246).
workability that can be grouped into norms for performance, conduct and structure. The selection of norms follows from an inquiry that 'would be concerned in the case of performance norms to decide what state of the various performance dimensions would imply maximum service to buyers' desires, and in the case of structure and conduct norms, to decide what state of their dimensions would imply maximum effects on performance' (Sosnick 1958, p. 395). The dynamics of competition are very important: 'Indeed, market incentives are able to designate and evoke aspects of desirable behavior not merely when enterprises are constrained by conditions like substitutes' availability and opposites' knowledge, but when producers have freedom to innovate and latitude when successful' (Sosnick 1958, p. 397).

To give an idea of the nature of the workability criteria listed in the literature, Table 2 presents Reid's (1987, p. 125) listing of Sosnick's norms. The fact of the matter is that the workability concept can be improved by providing a formal microfoundation to the informal intuition. This is useful, since the 'difficulty with this approach [workability] lies in its sheer eclecticism. The criteria developed seem frequently arbitrary and vague, and occasionally inconsistent' (Reid 1987, p. 115). The lack of analytical rigor goes hand in hand with unanswered questions: what is the relative importance of the criteria listed; in what ways may interdependencies between criteria interfere; what critical levels are associated with the criteria listed? Attempts to elaborate formally on the workable competition concept seem therefore to be justified. To focus attention, two essential criteria of workable competition, which capture many of Sosnick's (particularly performance) norms, have to be kept in mind. First (norm I), '[p]rofits should be at levels just sufficient to reward investment, efficiency and innovation' (Scherer 1980, p. 42). Second (norm II), '[o]pportunities for introducing technically superior new products and processes should be exploited' (Scherer 1980, p. 42).

Broadly speaking, formal models resembling workable competition can be grouped into two categories: patent race competition (Fudenberg et al. 1983; Fudenberg and Tirole 1985; and Mills 1988) and investment contestability (Cairns and Mahabir 1988; Calem 1988; and Van Witteloostuijn 1990b and 1990c). The common denominator in this literature is the argument that positive sunk costs can be compatible with a free entry condition. The difference is that patent race competition starts from the condition of free entry \textit{ex ante} (that is, before costs have been sunk), whereas investment contestability is based upon an assumption of free entry \textit{ex post} (that is, after costs have been sunk). By way of illustration two exemplary models of patent race competition and investment contestability are briefly discussed: Mills (1988) and Calem (1988), respectively.

Patent race competition describes tournament games of a winner-takes-all nature (Dasgupta 1986). By law, a patent effectively impedes potential entrants from entering the new technology or product: the winner of the patent is a protected monopoly \textit{ex post} \textit{(i.e.,} after the patent has been granted). However, the
TABLE 2 - STRUCTURE-CONDUCT-PERFORMANCE NORMS

Structure norms
(1) No dominance, and traders as large as economies of scale will permit
(2) Quality differentials which are moderate and sensitive to prices
(3) No impediments to mobility
(4) Reasonable availability of market information
(5) Some uncertainty about responses to price cutting
(6) Freedom from legal restraint
(7) Development of new markets and trade contracts

Conduct norms
(1) Independent rivalry, in pursuit of profit
(2) No shielding of inefficient rivals, suppliers, or customers
(3) No unfair, exclusionary, predatory, or coercive tactics
(4) No unreasonable discrimination
(5) No misleading sales promotion
(6) Rapid response by buyers to differentials in attributes of products

Performance norms
(1) Efficient production and distribution
(2) No excessive promotional expenses
(3) Profits sufficient to reward investment, efficiency, and innovation
(4) Output consistent with efficient resource allocation
(5) Prices that do not intensify cyclical problems
(6) Quality consistent with consumers' interest
(7) Appropriate exploitation of improved products and techniques
(8) Conservation requirements respected
(9) Sellers responsive to buyers' needs
(10) Entry as free as the industry sensibly permits
(11) Regard for national security requirements
(12) Avoidance of excessive political and economic power in few hands
(13) Regard for employees' welfare

Key point is that ex ante entry into the patent race may be free. Competition for a patent is a race to be the first: the timing of the introduction of the innovation is the rivals' strategic instrument. A firm enters the patent race by sinking R&D outlays ex ante: increasing R&D efforts speeds up the introduction of the innovation. R&D is subject to time diseconomies: earlier introduction dates require higher R&D budgets. The specifics of the timing game determine the extent of rent dissipation (Fudenberg and Tirole 1987, p. 182). Mills (1988) provides an example of a timing game that gives zero profit to the winner.

Mills describes two scenarios: one with costless threats and one with costly threats. The former illustrates workable competition, whereas the latter shows
that sunk costs can raise entry barriers. The essential assumptions of the setting with costless threats are ‘that there are two firms equally capable of making the investment and capturing the payoff. Assume that they have full information about each other and behave in a noncooperative fashion. Assume that both firms would receive negative payoffs if both invest, because the second investment would be redundant. Unless both firms invest simultaneously, the prospect of certain loss is enough to prevent a second investment. Neither firm invests once it learns that its rival has invested. ... [It is] assumed that firms must act sequentially’ (Mills 1988, p. 116). So, entry is free ex ante.

Sequential investment implies a positive entry and adjustment lag, where entry always can occur before the incumbent is able to adjust. Three assumptions complete the model. First, $t_m$ is the monopoly’s unique optimal introduction date of the sunk investment: that is, a protected monopoly would invest a sunk cost at date $t_m$, since then the maximum payoff can be captured. Second, $\pi_{t-1} < \pi_t$ for $t \leq t_m$: this is an assumption of time diseconomies. Third, a zero profit accrues to the firm which refrains from investment (or entry): the opportunity cost of entry is zero. The result of the timing game is that ‘[t]he investment described above can be viewed as a natural monopoly where potential monopoly rents are dissipated by rivalrous behavior. While this outcome results from the potential for “hit-and-run-entry” in contestable markets, in this model it results from pre-emptive timing. In a contestable market rent dissipation occurs because entry involves no unrecoverable costs. Here it occurs because entry can be timed strategically’ (Mills 1988, p. 117).

Investment contestability starts from the critique that perfect ‘contestability theory leaves unanswered questions about the identity of entrants and the source of their resources. ... Rather than as a hindrance to potential entry, sunk costs are viewed as being of key significance for the disciplining of oligopolists through potential entry’ (Cairns and Mahabir 1988, p. 269). This means that ‘[e]stablished firm entry is easier than new firm entry, an important factor being reserve or excess capacity in the short run which may be shifted to producing other types of product. ... Altering the product set requires latent fungibility as well as latent capacity. The latter may be provided by previously sunk expenditures for capital: the former requires discovering the latent fungibility, through R & D and other expenditures that are necessarily sunk, as well as the use of the goodwill of the firm’ (Cairns and Mahabir 1988, p. 273). Investment contestability describes potential competition from existing potential entrants which can benefit from sunk investments in their home market (Van Witteloostuijn 1990b and 1990c): entry is free ex post. Calem (1988) provides an example of investment contestability.

Calem (1988) offers a model of penetrable markets: ‘The potential entrant may be a firm producing an identical product but operating in a geographically distinct market, in which case ease of entry would derive from low transport costs; or it may be a firm producing a distinct but technologically related product, in which case ease of entry would derive from substitutability in produc-
tion. A market faced with entry as such may be deemed a penetrable market. It is plausible that a firm in a penetrable market would be unable to adjust its total output as rapidly as a potential entrant could transfer some of its products into the market. ... the rivalry between a firm in a penetrable market and a firm threatening to enter that market may be reciprocal. As these firms operate in related markets, each may be a potential entrant into the other's market' (Calem 1988, p. 171). In line with Shepherd (1984) and Green (1987), Calem (1988, pp. 172–173 and 180) argues that the penetrable market model applies particularly to (the threat of) foreign competition.

Calem's model describes Cournot rather than Bertrand competition: quantity is the strategic variable. It is essential that incumbent firms and potential entrants decide on output (productive capacity) simultaneously. That is, '[a] novel feature of our model is that the monopolist in a market, and the firm threatening entry into the market, both make strategic choices (choose their total outputs) during the pre-entry stage' (Calem 1988, p. 172). Entry is anticipated by installing excess capacity. Further assumptions are that two monopolists in two distinct markets face a joint production technology, whereas the entry lag and transfer (entry) costs are positive. Moreover, a firm is committed to productive capacity for one period (with a pre-entry and post-entry stage).

The result of the model is that '[i]n a penetrable market, as in a contestable market, a threat of entry can limit the exercise of monopoly power' (Calem 1988, p. 173). This result depends critically upon the features of the setting involved. To be precise, '[a] firm facing a threat of entry will employ an entry-deterring strategy in its home market only if its transfer costs are large enough to rule out its being a potential entrant into its rival's market. Moreover, in the asymmetric case, the firm facing a threat of entry into its market will expand its output to an entry-deterring level only if its rival's transfer costs are in a medium range (not if those costs are small)' (Calem 1988, p. 181). This means that market performance resembles workable competition if the entry threat is one-sided. The entrant's home market is safe against entry by expelled incumbents.

2.3.2 Game
The three groups of theories that combine the merits of pure-contestability (average cost pricing) and non-contestability (sunk investment) – workable competition, patent race rivalry and investment contestability – can be denoted as quasi-contestability. The key feature of quasi-contestability is that sunk costs are positive and entry is free (ex ante or ex post) such that profit is driven down to zero. The quasi-contestability game has two stages: the first-stage subgame determines (the height and/or timing of) sunk investment; the second-stage subgame focuses on competition for market share. Games with ex ante free entry (patent race rivalry) are different from games with ex post free entry (investment contestability). Patent race rivalry gives a degenerate second-
stage subgame: the winner of the patent can benefit from a protected monopoly position. Market performance follows from the outcome of the first-stage timing subgame: rent dissipation may occur if entry is free \textit{ex ante}, which induces zero-profit pre-emptive timing under specific conditions (for example, sequential investment, time diseconomies and zero opportunity cost of entry).

Investment contestability is associated with second-stage free entry if incumbent firms and potential entrants face (cost and product) parity after the first-stage subgame on sunk investment. The key difference with non-contestability games (Figure 1) is that the first-stage subgame is dictated by \textit{simultaneous} strategy formulation: incumbent firms and potential entrants decide on sunk investment \textit{without} being informed of the decision of the rivals. The nature of an illustrative two-staged game for a duopoly is depicted in Figure 2, where payoffs indicate the gross profits (that is, without taking into account the amortization of sunk costs) captured in the incumbent firm’s market. Firms decide on sunk investment \textit{ex ante}. For the sake of convenience, take the discrete case where firms face two investment opportunities: $S = 0$ or $S = S^*$. Four second-stage subgame equilibrium outcomes can be identified:

(A) Both rivals face parity: the second-stage subgame resembles pure-contestability (Table 1) for average total cost pricing, which takes the amortization of sunk costs into account.

(B) The incumbent firm has developed a first-mover advantage: the second-stage subgame resembles non-contestability (Figure 1) with outcome $(M, 0)$.

(C) The incumbent firm is outpaced by the entrant: second-stage entry (Figure 1) gives outcome $(D^i, D^e)$, where $D^i = 0$ and $D^e = M$ if the entrant is able to fully replace the incumbent.

(D) Neither firm has sunk any cost: the second-stage subgame reflects pure-contestability (Table 1).

Quasi-contestability describes cases of competition where the perfect

![Quasi-contestability Diagram](image-url)
equilibrium of the two-staged game is \( S^i = S^e = S^* \), average total cost pricing (zero profit) and non-entry (equilibrium (A)). For example, in Calem (1988) \( S^* \) is the cost sunk in productive capacity in the first stage: the fact that both firms sink costs in the first stage may limit monopoly power in the second stage. The quest for competitive conditions that give free entry and zero profit \textit{ex post}, notwithstanding positive sunk investment \textit{ex ante}, is the subject matter of sections 3 and 4; that is, the focus is on investment contestability. The logic of backward induction (Rasmusen 1989, p. 88) dictates that section 3 focuses on the second-stage subgame for market share (what conditions explain why the second-stage payoff matrix resembles pure-contestability?), whereas section 4 deals with the first-stage subgame on sunk investment (why are firms inclined to undertake first-stage sunk investment in the prospect of zero profit?). However, first the scene will be set in subsection 2.3.3.

2.3.3 Performance

Quasi-contestability combines the merits of pure and non-contestability. Average total cost pricing goes hand in hand with positive sunk investment. Table 3 summarizes Scherer’s (1980, p. 42) performance features of the three categories of competition. Criterion I implies that \( P = T = V \text{ if } S = 0 \), whereas norm II is associated with \( S > 0 \).

Of course, criterion II is weak: the condition that \( S > 0 \) is insufficient to induce dynamic efficiency. The quality of dynamic performance stands or falls with the specifics of the sunk investment \( S \): this is the well-known dilemma of evaluating the welfare features of entry barriers (subsection 2.2.3). Section 4 illustrates dynamic efficiency of quasi-contestability for the cases of process and product innovation.

A review of the literature indicates that the credibility of the threat of potential entry is the key issue. The answers to three questions are particularly relevant in determining the credibility of the entry threat and hence the nature of competition.

(1) Are incumbent firms and potential entrants at par in terms of \textit{sunk} technology and product features (a)symmetry of sunk cost)?

(2) Are potential entrants able to attract the incumbent firms’ customers before the latter are able to respond (structure of time lags)?

(3) Is the cost or profit foregone with entry \textit{zero} (scale of opportunity cost of entry)?

<table>
<thead>
<tr>
<th>Theory of competition</th>
<th>Static performance</th>
<th>Dynamic performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pure-contestability</td>
<td>( P = T = V )</td>
<td>( S = 0 )</td>
</tr>
<tr>
<td>Quasi-contestability</td>
<td>( P = T &gt; V )</td>
<td>( S &gt; 0 )</td>
</tr>
<tr>
<td>Non-contestability</td>
<td>( P &gt; T &gt; V )</td>
<td>( S &gt; 0 )</td>
</tr>
</tbody>
</table>
Quasi-contestability prevails if the answer to questions (1), (2) and (3) is yes. This is, for example, the case for the models of Calem (1988) and Mills (1988). However, this begs the question: what competitive environments facilitate quasi-contestability? By way of illustration, sections 3 and 4 describe a quest for favorable market conditions that impose a discipline on incumbent firms' conduct. Then, quasi-contestability goes with a credible threat of potential entry that forces incumbent firms to behave in the interest of customers. It appears that the sources of potential entry are of crucial importance.

3 STATIC EFFICIENCY: PRICING POLICIES AND SOURCES OF POTENTIAL ENTRY

3.1 Established Potential Entrants

Outcomes resembling quasi-contestability stand or fall with the (in)credibility of the entry threat. Perfect contestability theory *a priori* assumes a credible entry threat. That is, the entry threat is exogenous to the theory. This is also true for the literature on entry barriers and patent races. Broadly speaking, the assumption of an exogenous entry threat is common in IO: the models remain silent on the identity of potential entrants. For the sake of convenience, this section takes perfect contestability as its point of departure. Among the many critiques of the contestable market theory, one line stands out that is relevant in the current context. Shepherd (1984), Calem (1988) and Cairns and Mahabir (1988) argue that the existence and credibility of the entry threat in a perfectly contestable market needs careful examination. What is the identity of these entrants that are able to hit-and-run to benefit from short-lived entry opportunities only? If potential entrants are new firms that still need to build up capacity, their speed of response is unlikely to be as fast as hit-and-run requests. Besides, entry by new firms is unlikely to be associated with a zero sunk entry cost. So, they probably are existing firms which contest, for example, not by investing in the entry market, but rather by exporting goods (Shepherd 1984, p. 584; and Green 1987, p. 485). Established firms are engaged in potential entry into each other's market (Calem 1988). For example, import competition does indeed constitute a major threat against market shares of dominant firms (Scherer 1980, p. 241).

Following Cairns and Mahabir (1988) and Calem (1988) home market arguments can be used to identify credible potential entrants. The key point is that the foundation of the credibility of the entry threat is facilitated by focusing on a multimarket framework (Van Witteloostuijn and Van Wegberg 1992). A multimarket approach is implicit in many accounts of the contestable market (for example, Baumol et al. 1982, p. xxi). Making it explicit, however, breaks

---

5 Perfect contestability gets around this dilemma by assuming that capital is not sunk. That is, capital can be bought and sold in an outside capital market in such a way that the selling firms only lose a user cost. However, this only begs the question: what rationale underlies the existence of such a perfect capital market?
with one dominant assumption in the literature on entry and exit in IO: the exogeneity of the *opportunity cost of entry*. The common assumption is that the opportunity cost of entry is zero: potential entrants earn a zero profit if they do not enter. Hence, they enter only if entry profits are strictly positive. This can be the case because the alternative to entry is either zero production and investment or stable home market profit. Home market profit does not influence the entry decision because of the usual key assumption that entry leaves the entrant’s home market profit unchanged. However, in a multimarket context the assumption of zero opportunity cost of entry is not so obvious.

On the one hand, serving the entry market with overcapacity has a zero opportunity cost. However, overcapacity invites dumping, which is not consistent with quasi-contestability, since dumping strategies are associated with prices below average total cost. On the other hand, without overcapacity the opportunity cost of entry may exceed zero. Potential entrants facing binding capacity restrictions incur a profit foregone in the home market if entry is undertaken. So, two questions focus on the credibility of the entry threat in a multimarket framework: (1) can dumping strategies be avoided if firms face excess productive capacities? (subsection 3.2); and (2) can a credible entry threat exist if potential entrants have to take into account the implications of binding productive capacity restrictions? (subsection 3.3) This section starts from the assumption that incumbent firms and potential entrants face parity: that is, both parties have invested $S^i = S^e = S$ in the first stage, which has induced cost and product parity in the second stage. This means that the focus is on the second-stage Bertrand subgame of a two-staged potential rivalry game (Figure 2). The outcome of the first-stage subgame on investment is taken for granted. Section 4 deals with the first-stage subgame.

3.2 Overcapacity

Cairns and Mahabir (1988) suggest that the credibility of the entry threat ensues from firms with excess capacity in related markets. They ‘argue that firms in related industries have an advantage because of their own sunk costs. An advantage (over completely new firms or firms in unrelated industries) may arise if costs are sunk and (1) there is short-run excess capacity because of unpredictable demand fluctuations; (2) there are regular peak and off-peak periods and the firm is actively seeking a use for its off-peak excess capacity; (3) it is normal to have some capacity reserve that can be utilized if need be, perhaps at higher operating costs; (4) the firm has excess capacity, created in order to deter entry to its own market...; or (5) the firm is a member of a monopolistically competitive industry and, for that reason, has excess capacity’ (Cairns and Mahabir 1988, p. 273).

However, the entry threat from potential entrants with overcapacity is unable to discipline incumbent firms if (reciprocal) dumping constitutes a
COMPETITION AND MARKET PERFORMANCE

profitable strategy. If a firm has excess capacity, then this firm is, ceteris paribus, inclined to use this excess capacity so as to supply commodities in the rivals' market for any price above marginal entry cost: this is dumping. Actual entry is likely to diminish the incumbent firm's profit – if it does not drive the incumbent firm off the market altogether. The incumbent firm may retaliate by reciprocal entry into the entrant's home market (Bulow et al. 1985; and Calem 1988). Bulow et al. (1985) point to strategic interactions that can occur in a multimarket framework. In an example of two monopolists A and B, which are potential entrants into each other's markets, they argue that 'B's entry will change A's equilibrium output in the market where it is incumbent and therefore possibly alter its decision of whether to enter B's market. ... if B enters A's market then it may be profitable for A to retaliate. So the threat that deters B's expansion is a credible one' (Bulow et al. 1985, p. 505).

The point is that exit gives idle productive capacity. The expelled incumbent firm can increase profits (or decrease losses) by selling output in the entrant's market for any price above average marginal entry cost. As exit is associated with zero production, an expelled firm can profitably enter the entrant's market for any price above average variable cost. The threat of dumping by expelled incumbent firms does not undermine the credibility of the potential rivals' entry threat either if the potential entrants' home market is safe against entry by expelled incumbent firms or if the incumbents' response lag is large enough to invalidate the retaliatory dumping threat (subsection 3.3). The case may be different, however, particularly if both incumbent firms and potential entrants have sustainable overcapacity. For illustrative purposes, an argument can be put forward so as to defend the assumption that dumping does not occur. Particularly the literature on international economics has been concerned with the issue of (reciprocal) dumping (Brander and Krugman 1983; Pinto 1986; Calem 1988; and Venables 1990).

Assume that potential entrants take into account the threat of reciprocal dumping. That is, incumbent firms have the opportunity to dump into the potential entrants' home market. Therefore, incumbent firms and potential entrants play a game on excess capacity (dumping) and price, which may or may not yield an equilibrium that is associated with statically efficient outcomes and the absence of (anticipative) dumping. Take two symmetric representative firms i (incumbent firm) and e (potential entrant), respectively. Assume that both firms play a (Bertrand-) Nash game on dumping in both markets. Ignoring lag conditions, the payoff matrix is indicated in Table 4.

The static (one-period) dumping game clearly reflects a noncooperative Prisoners' Dilemma, since \( Y<Z<(0<X) \), which yields a Nash equilibrium in which both firms decide to dump \((Z,Z)\), although the solution where neither firm dumps \((0,0)\) is preferable. The dynamic (multiperiod) repeated dumping

\[ 6 \text{ A second comment is that excess capacities are not always sustainable in a market with free entry (Van Witteloostuijn 1990a, pp. 147-149).} \]
TABLE 4 - RECIPROCAL DUMPING

<table>
<thead>
<tr>
<th>Incumbent firm</th>
<th>Potential entrant</th>
<th>Zero dumping</th>
<th>Positive dumping</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Zero dumping</td>
<td>(0,0)</td>
<td>(Y,X)</td>
</tr>
<tr>
<td></td>
<td>Positive dumping</td>
<td>(X,Y)</td>
<td>(Z,Z)</td>
</tr>
</tbody>
</table>

game is more complicated (Friedman 1986). At least three (related) arguments can be used to indicate that the strategy pair that neither firm decides to dump, is a (Nash) equilibrium.

First, reputation considerations can call into question the equilibrium features of the case where both firms are dumping. Firm i can communicate to firm e that firm i’s dumping will be followed by reciprocal dumping, which leaves firm e worse off. This argument bears a family resemblance to the predation literature (Milgrom and Roberts 1982; and Roberts 1986). Second, the reputation effect wins cogency if firms have or behave as if they have an infinite planning horizon. In effect, the Folk Theorem shows that the preferred payoff combination can be a subgame Nash equilibrium in an infinitely repeated Prisoners’ Dilemma (Pinto 1986). The assumption of planning as if the horizon is infinite is plausible, since firms are generally uncertain of the date at which they will cease to operate. Fisher (1989) points out that ‘real-life incumbents do not face a well-defined finite set of potential entrants. Corporations in most contexts are assumed to have an infinite horizon and surely cannot believe that any particular fight will be the last’ (Fisher 1989, p. 123). Third, firms can (tacitly) agree upon not dumping. The dumping game is non-cooperative but gives (tacitly) cooperative outcomes. The case where neither firm dumps is clearly the cooperative outcome.

3.3 Binding Capacity

Binding productive capacity constraints can introduce a positive opportunity cost of entry. Binding capacity restrictions force the firm to give up sales in its home market in order to be able to serve an entry market. Cairns and Mahabir (1988) point out that ‘[t]he entrant shifts capacity from an original [market]..., expecting to return after a hit-and-run entry. But that shift reduces capacity in the original market, as compared with the original equilibrium there, forcing a disequilibrium in which prices must rise to clear the market, thereby creating profits for remaining firms. That market will then be invaded by yet another 7 Moreover, firm e has to take into account the observation that, if firm e repeatedly decides to dump, it is likely that firm i will start to expect that firm e will dump. This induces firm i’s best anticipative reply to dump as well, which leaves firm e worse off.

8 The Folk Theorem ‘clarifies the role of trigger strategies in supporting cooperative outcomes by means of noncooperative equilibrium strategy combinations’ (Friedman 1986, p. 104). For example, high discount factors facilitate cooperative outcomes.
entrant. But that leaves no room for the first invader to return to the original market' (Cairns and Mahabir 1988, p. 271).

This subsection describes six scenarios where existing firms operating elsewhere imply a credible entry threat, because they (i) have fungible capital or (ii) face a negligible transportation cost disadvantage. That is, there are potential entrants which can easily switch productive capacity or transfer (part of) the commodities produced to the incumbents' market. Suppose that demand is localized so that only firms are mobile between markets (which is, for instance, plausible in the context of international competition) and assume the absence of anticipative (reciprocal) dumping (subsection 3.2). To illustrate the role of time lags, a number of \textit{ex post} periods is assumed (\textit{i.e.}, the second-stage subgame in Figure 2 is stretched over a number of periods). Incumbent firms announce prices and potential entrants decide on entry at the beginning of period $t=1$ on the basis of the result of R&D in period $t=0$. R&D requires sunk cost $S_0$. Prices remain fixed during a period\footnote{This assumption is in accordance with Hicks' (1939, p. 122). The identification of a period with the interval in time during which prices are constant is nowadays adopted in the popular 'temporary equilibrium' tradition (Grandmont 1977).} (incumbent firms' price adjustment delay $A = 1$). The response of incumbent firms in the case of exit takes one period (incumbent firms response lag upon entry $R = 1$). Discounting is ignored. The average total cost price is $P^* = T$. The life time of the innovation expires after period $t=F$: before $t=F+1$ the invested sunk costs must be recovered. Moreover, the prototype market argument is based upon the assumption that firms face (investment and production) cost parity ($S_i^0 = S_i^e$ and $V_i^0 = V_i^e$): the answer to question (1) in subsection 2.3.3 is affirmative. Potential entrants enter the market if they perceive any profit opportunity. The potential entrants' entry lag is assumed to be zero ($E = 0$). The combination of $E = 0$, $A = 1$ and $R > 0$ provides question (2) in subsection 2.3.3 with an affirmative answer. The outcome of the investment $S_0$ can be applied to both markets. However, with binding capacity restrictions a firm is not able to serve both markets simultaneously.

Home market arguments can be facilitated by a restriction to one-sided entry (threats). That is, the potential entrants' home market is safe against entry. This means that entrants can anticipate an easy return to the home market after hit-and-run entry. For example, Calem (1988) introduces the assumption of one-sided entry (threats) by arguing that the incumbent firms' entry cost is sufficiently high to keep it from entering the potential entrants' market or, alternatively, by supposing that there are legal or regulatory barriers which prevent incumbent firms from being potential entrants into the rivals' market. One-sided entry fixes a zero opportunity cost of entry (a definite yes to question (3) in subsection 2.3.3).

It is important to distinguish the case where $F=1$ from the one where $F>1$. For $F=1$, unlike $F>1$, the assumption of (potential) entrants' Bertrand conjectures is easily sustained.
EXAMPLE 1 (short-lived innovations). For \( F = 1 \) it is obvious that a potential entrant is able to adopt hit-and-run entry. Sunk cost must be recovered in period \( t = 1 \). If an incumbent firm is expelled from the market as a result of careless price setting in period \( t = 1 \) (\( P_i^t > P^* \)), then the incumbent firm fails to recover the invested sunk costs, because the incumbent can only re-enter after one period (\( R = 1 \) so that re-entry cannot take place before period \( t = 2 \) when the innovation is outdated):

\[
\pi_i^t = -S_0. \tag{1}
\]

Hence, the expelled incumbent firm faces an exit cost. The entrant can capture positive hit-and-run profits so that no exit costs are incurred:

\[
\pi_e = \left( P_i^e - V_i \right) \cdot Q_i^e - S_0 > 0, \tag{2}
\]

where \( Q \) is the quantity sold. Condition (2) gives entry if the potential entrant captures a zero profit in the home market. \( F = 1 \) corresponds, for instance, to the case where firms introduce incremental, short-lived innovations so that R&D expenditures must be amortized in a short (one) period of time. The incumbent firm is only able to retaliate through re-entry after the expiration date of the innovation (\( F = R = 1 \)) when the entrant can safely return to its home market.

For \( F > 1 \) it is likely that potential entrants anticipate the response of an expelled incumbent firm at the beginning of period \( t = 2 \). Hence, they only temporarily employ Bertrand conjectures. So, assume that potential entrants only employ Bertrand conjectures for the first period after entry. They correctly anticipate the response of an expelled firm at the beginning of \( t = 2 \) (or \( R + 1 \) in Example 4). The likely response of an expelled incumbent firm follows from the argument that an incumbent is inclined to re-enter after exit at the beginning of period \( t = 2 \) for any price in excess of its average variable cost (that is, in excess of the shut-down-price). A price \( P_i^t > V_i \) (where \( t > 1 \)) enables the recovery of (part of) the sunk cost invested. It is the sunk nature of investment expenditures that induces the incentive to an expelled incumbent firm to re-enter for any \( P_i^t > V_i \), because the opportunity cost of refraining from entry (\( \pi_N^t \)) exceeds the cost of re-entry (\( \pi_E^t \)):

\[
\pi_N^t = -S_0 < \pi_E^t = \left[ \sum_{t=2}^{F} \left( P_i^t - V_i \right) \cdot Q_i^t \right] - S_0 < 0, \tag{3}
\]

for any \( P_i^t > V_i \). There are now a number of examples of cases where incumbent firms face a credible threat of entry, despite their ability to respond with low-priced re-entry at the beginning of period \( t = 2 \). One example fits into the set of scenarios which assumes a protected home market condition.
EXAMPLE 2 (temporary cross-entry). Potential entrants deploy temporary cross-entry. Before the incumbent firm is able to respond, the entrant returns to its home market \((H)\). Suppose that the entrant is able to sell a quantity \(Q_i^H\) at price \(P_i^H\) in the period from \(t = 2\) to \(t = F\) after returning to its home market at the beginning of period \(t = 2\). If an incumbent firm sets \(P_i^t > P_i^*,\) then the potential entrant enters even for \(F > 1\) if entry gives a positive profit

\[
\pi^e = [(P_i^e - V_1) \cdot Q_i^e] + \left[ \sum_{t=2}^{F} (P_i^H - V_i) \cdot Q_i^H \right] - S_0 > 0, \tag{4}
\]

for any \(P_i^e > P_i^*\) (provided that \(P_i^H = P_i^*\)). Hence, if condition (4) holds, the entrant does not face any exit cost. Re-entry of the expelled firm at the beginning of period \(t = 2\) is associated with \(P_i^t = P_i^*\) (where \(t = 2, \ldots, F\)) so as to avoid further exit. So, the incumbent firm, again, faces exit costs:

\[
\pi^i = \left[ \sum_{t=2}^{F} (P_i^* - V_i) \cdot Q_i^t \right] - S_0 < 0. \tag{5}
\]

Condition (5) corresponds to the case where (internal) conditions in the potential entrant’s home market dictate price \(P_i^H = P_i^*\) in the period \(t = 1\), so that the opportunity cost of refraining from entry for any \(P_i^t > P_i^*\) is positive, conditional upon the assumption that the potential entrant is able to return to its home market at the beginning of period \(t = 2\) without suffering negative consequences (in the sense of losing the home market). So, \(P_i^H = P_i^*\) gives the threshold entry-deterring price level \(P_i^t = P_i^*\) where the potential entrant is indifferent between entry and non-entry.

Further examples follow from dropping the assumption of an easy return to a home market. Then, although an entrant is able to ‘hit’ at the beginning of period \(t = 1\), it is unable to ‘run’ at the beginning of period \(t = 2\). It may be that a return to the home market is impossible altogether or that a profitable return is only feasible after the sunk costs have been fully recovered at period \(t = F\). Examples 1 and 2 assume that the incumbent firm faces idle productive capacity during the period of exit. This need not be the case. For example, it can be that the expelled incumbent firm enters another (perhaps even the entrant’s) market. By shipping goods to an entry market, entrants without excess capacity withdraw supply from their home market. This may invite further entry into the entrant’s market. Hence, when the entrant returns from its hit-and-run, it is likely to discover that it has lost its home market (Cairns and Mahabir 1988). The next examples assimilate this argument by focusing on hit-and-stay entry. Note that one complication is ignored: collusion. Firms meeting in many markets may have incentives to collude (Van Witteloostuijn and Van Wegberg 1992). Broadly speaking, second-stage collusion tends to damage static efficiency (Jacquemin and Slade 1989). The examples below ignore collusion by
assuming that (i) collusion is prohibited by law or (ii) incentives to compete are dominant (that is, the payoff of cheating exceeds the profit of colluding).

EXAMPLE 3 (small R&D budgets). For \( F > 1 \) the potential entrant responds with entry to a \( P_f^i > P^*_i \) (provided that home market price \( P_{1t}^H = P^*_i \), which implies a zero opportunity cost of entry) if it is profitable to deter re-entry by the expelled incumbent firm by reducing \( P^e_i \) to the expelled incumbent firm's opportunity price at the beginning of period \( t = 2 \). Take, for example, the case where the incumbent firm receives a price \( V_2 \) outside its home market. Then, if

\[
[(P^e_i - V_1) \cdot Q^e_t] - S_0 > 0,
\]

then the potential entrant is able to capture positive profits without inducing re-entry. Condition (6) implies that a potential entrant is inclined to enter if the entry price permits the recovery of the sunk cost in one period. If condition (6) holds, the entrant faces no exit costs. Then, successful entry by the expelled firm requires \( P_2^e < V_2 \), which is not worthwhile, since condition (3) is violated. The expelled incumbent firm faces exit costs (equation (1)). Obviously, this case is extremely sensitive to the precise value of \( S_0 \) (i.e., the scale of investment). Low values of \( S_0 \) (that is, small R&D budgets) facilitate the likelihood that this case is valid. However, in general condition (6) is unlikely to hold. Consider, therefore, three further scenarios.

EXAMPLE 4 (large response lag). Suppose that the argument by Baumol et al. (1982) that 'incumbents are restrained by law and other impediments from undertaking retaliatory moves' (Baumol et al. 1982, p. 350), i.e., post-entry price reductions or low-priced re-entry, holds in such a way that expelled incumbent firms suffer from a response lag \( 1 < R < F \). Then, condition (6) transforms into

\[
[(P^e_i - V_1) \cdot Q^e_t] + \left[ \sum_{t=2}^{R} (P^*_i - V_1) \cdot Q^e_t \right] - S_0 > 0,
\]

since, for \( t = 2, \ldots, R \), the entrant only fears hit-and-run entry by other potential entrants, so that \( P^e_i = P^*_i > V_1 \) deters entry for \( t = 2, \ldots, R \). Again, the expelled firm, unlike the entrant, faces exit costs (equation (1)).

From condition (7) it follows that high values of \( R \) (that is, large response lags of expelled incumbent firms) make it more likely that this case is valid. In effect, this scenario shows close resemblance to Example 2 (for \( P^H_i = P^*_i \)), although the entrant is unable to return to its home market as, for example, other entrants have invaded its home market (before period \( t = F \)). Moreover, for \( R \to F \) this case converges to Example 1.

EXAMPLE 5 (alternative use). Assume, for the sake of simplicity, that \( F = 2 \).
Suppose that an expelled firm is able to benefit from an alternative use of its R&D output. Then, of course, the R&D expenditures \( I_0 \) are not fully unrecoupable. If sunk exit cost is denoted by \( S_0 \), then \( S_0 < I_0 \); so far, it has been tacitly assumed that \( S_0 = I_0 \). Suppose that

\[
\pi_F^U = \pi_1^U + \pi_2^U < I_0,
\]

where \( \pi_F^U \) denotes the (negative) profits from alternative use, composed of profits during period \( t = 1 \) (\( \pi_1^U \)) and \( t = 2 \) (\( \pi_2^U \)). Then, for \( t = 2 \) there is a \( P_2^e > V_2 \) such that re-entry is deterred. This entry-deterring price \( P_2^L \), and associated quantity \( Q_2^L \), follows from

\[
(P_2^L - V_2) \cdot Q_2^L = \pi_2^U.
\]

Hence, the potential entrant now decides to enter if the following condition holds:

\[
[(P_1^e - V_1) \cdot Q_1^e] + \pi_2^U - I_0 > 0.
\]

If condition (10) holds, the entrant does not face an exit barrier. The expelled firm is only able to enter successfully by offering a \( P_2^e < P_2^L \), which is not worthwhile, since then

\[
(P_2^e - V_2) \cdot Q_2^e < \pi_2^U.
\]

The incumbent’s exit costs follow from assumption (8):

\[
\pi_i = -S_0^i = \pi_F^U - I_0 < 0.
\]

So, the likelihood of a credible fear of hit-and-stay entry is facilitated by high benefits from the alternative use of R&D output.

**EXAMPLE 6 (reciprocal entry).** Suppose that an expelled incumbent firm can take the place of the entrant in the latter’s home market.\(^\text{10}\) If competitive conditions in the potential entrants’ home market dictate price \( P_t^H = P_t^* \) for \( t = 1, \ldots, F \), the entrant can deter re-entry by the expelled incumbent after period \( t = 2 \) (\( R = 1 \)) by setting the average total cost price from period \( t = 2 \) to period \( t = F \). Hence, entry is profitable if

\[
[(P_1^e - V_1) \cdot Q_1^e] + \left[ \sum_{t=2}^{F} (P_t^e - V_t) \cdot Q_t^e \right] - S_0 > 0,
\]

\(^{10}\) Alternatively, it can be assumed that the expelled incumbent firm is able to enter a third market. Porter (1980, pp. 84–85) calls competition over a third market ‘cross-parry.’
which holds for any $P_i^H = P_i^* < P_i^* = P_i^H$. Hence, if expelled incumbent firms are able to undertake reciprocal entry into the potential entrants’ home market against $P_i^H$, only $P_i^* = P_i^*$ deters entry when $P_i^H = P_i^*$ (for $t = 2, \ldots, F$). The expelled incumbent firm faces a zero exit cost as capital can be used profitably in another market. This example points out that incumbent firms’ low exit cost (as a result of the opportunities of reciprocal entry) increases the credibility of the hit-and-stay entry strategy of potential entrants.  

4 DYNAMIC EFFICIENCY: ENTRY-DETERRING STRATEGIES AND SUNK INNOVATION

Examples 1 to 6 predict static efficiency even for cases with positive sunk costs. However, the question of dynamic efficiency is bypassed by assuming that $S_i = S_i^*$ and ignoring the nature of the investment. This section complements section 3 by focusing on the investment issue. Section 3 described scenarios with static efficiency: this is the outcome of the second-stage subgame of a two-staged game (Figure 2). The first stage focuses on investment; the second stage deals with price (potential) competition. The outcome for investment strategy can be illustrated with an easy-to-understand first-stage subgame. This section takes the second-stage outcomes as given while concentrating on the firms’ investment strategies in the first stage. For the sake of convenience, time indices are suppressed.

Suppose that incumbent firms face prototype potential entrants which will capture a zero profit if they refrain from entry ($P^H = P^* = T$). Incumbent firms and prototype potential entrants have access to identical production ($V$) and investment ($S$) technologies. An evaluation of dynamic performance requires that the nature of the sunk investment is taken into account (subsection 2.2.3). By way of illustration the focus of investment is assumed to be innovation. That is, investment $S$ gives a process or product innovation. $S^*$ (in combination with $P^*$) maximizes buyers’ utility either through cost and price reduction (process innovation: Van Witteloostuijn 1990b) or product improvement (product innovation: Van Witteloostuijn 1990c). Hence, the choice of $S^*$ gives an innovation that is in the interest of the buyers through decreased price (process innovation) or increased quality (product innovation). Firms face two

11 The argument in Examples 5 and 6 resembles Eaton and Lipsey’s (1980) proposition that ‘[t]o make an entry-deterring threat ... the sitting monopolist must threaten that in the event of entry, he will stay in the market ‘long enough that the entrant’s present value at the time of entry will be nonpositive. ‘Long enough’ is $A^*$ periods. ... $A^*$ can be interpreted as the monopolist’s minimum commitment to the market or as the minimum barrier to his exit. It is in this sense that barriers to exit are barriers to entry’ (Eaton and Lipsey 1980, p. 728). Examples 5 and 6 mirror this result by pointing out that the credibility of the incumbent’s re-entry threat decreases if the incumbent’s exit barrier diminishes.

12 It remains to be seen what dynamic performance features quasi-contestability achieves with other types of sunk investment than R&D (Subsection 2.2.3).
investment options: \( S \neq S^* \) and \( S = S^* \). For the sake of comparison, take the discrete choice of \( S = 0 \) or \( S = S^* \).

Consider a representative incumbent firm and potential entrant which both face nonbinding productive capacity. Both have to deal with the investment decision in view of the payoff matrix that indicates the profit combinations depicted in Table 5 (calculated over both markets; that is, assuming a reciprocal entry threat). Equilibrium (A) gives pure-contestability, since sunk cost is zero (Table 1); equilibria (B) and (C) indicate non-contestability, because sunk cost is asymmetric (Figure 1); equilibrium (D) reflects quasi-contestability as both rivals are at par in terms of sunk investment (Figure 2). (Note that the case where (i) \( S' \neq S^* \), (ii) \( S' \neq S^* \) and (iii) \( S^e \neq S^* \) is redundant: then, one rival outpaces the other, which indicates non-contestability.)

The game is different for \( E=0 \) and \( E>0 \), respectively. First, quasi-contestability is a weak and nonunique equilibrium in pure strategies if \( E = 0 \). Exit may be costless for \( S = 0 \neq S^* \). Zero sunk cost gives a zero exit cost. If \( E = 0 \), both non-contestability (\( S' = S^* \) and \( S^e = 0 \) or vice versa: equilibria (B) and (C)) and quasi-contestability (\( S' = S^e = S^* \): equilibrium (D)) are equilibrium outcomes in pure strategies. For example, both monopoly outcomes are sustainable, since neither the expelled firm nor the monopolist has an incentive to change strategy: both innovation and noninnovation give a zero payoff. The non-contestable and quasi-contestable equilibria are weak and imply coordination difficulties (Rasmusen 1989, pp. 35–37): which rival (if any) will monopolize the sunk investment? A coordination mechanism is needed to select an equilibrium (for instance, through ex ante communication). However, quasi-contestability is a unique equilibrium outcome in mixed strategies. Even the smallest probability of being a monopolist (\( f \), which is the likelihood that the rival firm decides not to invest \( S^* \)) induces decision to invest optimally: \((1-f) \cdot E + f \cdot M = f \cdot M > 0 \) for \( E = 0 \) and any \( f>0 \).

Secondly, if \( E>0 \), quasi-contestability (\( S_i = S_e < S^* > 0 \) and zero profit to both firms) constitutes a unique Nash equilibrium which follows from dominating strategies. That is, both firms are inclined to invest optimally, notwithstanding the zero profit. The reason is simple: the investment strategy \( S^* \) yields at least as much payoff as when efficient investment is abandoned, while negative profits (i.e., exit) are avoided. The quasi-contestable equilibrium is strong, since the associated strategies are best replies no matter what the rival is doing. Moreover, this equilibrium is also the maximin solution. \( E \) may ex-

<table>
<thead>
<tr>
<th>TABLE 5 - SUNK INVESTMENT AND EXIT COST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potential entrant</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>Incumbent firm</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
ceed zero for two reasons. First, for $S = 0$ the expelled firm may face an exit cost (which may be arbitrarily close to 0) which only entails the burden indicating that firms (weakly) prefer zero-profit production to exit. This is the assumption that also explains the uniqueness of the pure-contestability equilibrium (subsection 2.1.2). Second, $E > 0$ if $S^* > 0$. If one firm invests optimally while the other does not ($S^i = S^*$ and $S^e = S^*$ or vice versa: equilibrium (B) or (C)), then the innovating firm is able to push aside the noninnovating rival so that the expelled firm fails to amortize its sunk costs.

5 APPRAISAL

The examples of the second-stage subgame on price in section 3 predict static efficiency even for cases with positive sunk (and exit) cost. The first-stage subgame on investment in section 4 indicates conditions which induce the introduction of dynamic efficiency for the case of innovation. This combination describes cases of quasi-contestability. Quasi-contestability (just as workable competition) is an umbrella notion that covers all scenarios which generate favorable performance (Maks 1986; Van Witteloostuijn and Maks 1988; and Van Witteloostuijn 1990a). So, as workable competition, quasi-contestability is defined with regard to its performance. Quasi-contestability is associated with market conditions that induce incumbent firms' behavior which gives in to both static and dynamic efficiency considerations. Quasi-contestability offers a formal microfoundation to a workability concept by introducing sunk investment in a contestability framework. This means, first, that a case of intermediate competition is framed and, second, that contestable market theory is transformed into a dynamic benchmark.

Existing models of perfect contestability deal with the case where sunk investment and exit cost cannot occur (Schwartz and Reynolds 1983; and Shepherd 1984). The endogenization of innovations (i.e., R&D activity) requires the introduction of sunk costs, since R&D cost can generally not be (fully) recouped outside the market (Stiglitz 1987). Hence, in order to introduce firms' sunk investment in a free entry framework, firms can undertake investment that is associated with a positive sunk cost. The sunk cost can give a positive exit cost if exit occurs before the sunk cost is fully amortized. However, the pervasive fear of entry is retained as incumbent firms anticipate full-scale (i.e., replacing) entry through the response of (a) alert potential entrant(s) from prototype markets which will take advantage of any profit opportunity by entering the incumbents' market, notwithstanding the sunk nature of investment.

The key point is that sunk capital can give an exit cost. This contradicts pure-contestability. Pure-contestability is compatible with investment, if the investment cost not yet amortized can be recouped at any exit date. That is, pure-contestability assumes that firms are always able to benefit from cost-effective selling or alternative use of the capital outside the market, so that exit cost is
zero by assumption irrespective of the date of exit. This is not the case with quasi-contestability. If exit occurs before the sunk investment is fully amortized, selling or alternative use of the sunk capital is not cost effective. That is, (part of) the investment cost is not fully recoupable outside the market. This means that (part of) the investment cost has to be recovered in the quasi-contestable market.

By way of illustration the conditions that underlie quasi-contestability are stated explicitly for the case where the entry threat ensues from existing potential entrants which either are inclined to refrain from dumping strategies (subsection 3.2) or face binding capacity restrictions in a home market (subsection 3.3). The key assumption is that incumbent firms and potential entrants are at par (in terms of technology, product and sunk exit cost). Moreover, to simplify matters, the discount factor is ignored. The credible fear of hit-and-run entry can only be sustained if the innovating potential entrants are able to recover precisely the investment cost in a home market. The reason is straightforward. Suppose that potential entrants face the opportunity to capture a (gross) profit ($\Pi$)

$$\Pi^H = \sum_{t=1}^{T} (P_t^e - V_t) \cdot Q_t^e$$

in a home market. Hence, (gross) profit $\Pi^H$ represents the opportunity cost of entry.

Assume, first, that all potential entrants capture a $\Pi^H > S_0$. This implies that potential entrants will only decide to enter when post-entry (gross) profits $\Pi^e$ exceed $\Pi^H$. Hence, incumbent firms are able to set an entry-deterring price ($P_t^L$) without provoking entry, where $P_t^L$ and associated $Q_t^L$ follow from

$$\sum_{t=1}^{T} (P_t^L - V_t) \cdot Q_t^L = \Pi^H,$$

provided that firms face sunk cost parity. Incumbent firms then capture positive (net) profits, since $\Pi^H > S_0$ (recall the assumption that $S_0^H = S_0^e$). Suppose, second, that there are full-scale potential entrants facing $\Pi^H < S_0$. An analogous argument indicates that the incumbent firms' limit price is now associated with negative (net) profits, which undermines the very incentive to innovate in the first place. So, the usual (tacit) assumption that potential entrants earn a zero net profit outside the market (section 2) is very important. However, this assumption only reflects a special case. In effect, it is the opportunity cost of entry that matters.

It may be illuminating to compare the assumptions of the theories of pure, quasi and non-contestability. The theories of pure and quasi-contestability, unlike that of non-contestability, are based upon the assumption of free entry. Pure-contestability, unlike the frameworks of quasi and non-contestability, is
TABLE 6 – COMPARISON OF ASSUMPTIONS

<table>
<thead>
<tr>
<th>Theory of competition</th>
<th>Assumptions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Free entry</td>
</tr>
<tr>
<td></td>
<td>Incumbents</td>
</tr>
<tr>
<td>Pure-contestability</td>
<td>Yes</td>
</tr>
<tr>
<td>Quasi-contestability</td>
<td>Yes</td>
</tr>
<tr>
<td>Non-contestability</td>
<td>No</td>
</tr>
</tbody>
</table>

associated with costless exit for incumbent firms and potential entrants. In a non-contestable market potential entrants have to overcome entry barriers, whereas both incumbent firms and potential entrants may face exit barriers. In a quasi-contestable market (prototype) potential entrants face free entry, while both incumbent firms and potential entrants may have to incur exit costs. The point is that in a non-contestable market, unlike in a quasi-contestable market, all potential entrants face barriers to entry that are higher than the exit barriers (which may even be absent). As Shepherd (1984) argues: ‘If entry barriers are higher, then exit barriers do not matter’ (Shepherd 1984, p. 578). In Table 6 the assumptions are summarized.

The distinction of the three types of markets is worthwhile, since they are associated with different implications for market behavior. In pure and quasi-contestable markets, unlike non-contestable markets, a credible entry threat forces incumbent firms to satisfy a zero-profit condition. Existing pure-contestable market models, however, unlike the quasi and non-contestability frameworks, fail to take account of the case where firms’ (innovative or strategic entry-deterring) sunk investment behavior can occur. The implications are summed up in Table 3.

Quasi-contestability theory is a quest for market conditions that generate incentives to (incumbent and potential) suppliers to adopt careful pricing on the one hand and to innovate and satisfy buyers’ desires on the other. Quasi-contestability offers a solution to the static-dynamic efficiency dilemma by introducing positive sunk and exit costs in a free entry framework. The force of rivalry from prototype potential entrants is crucial. The credible entry threat disciplines the incumbent firms’ behavior. This means that the source of potential entry and the multimarket perspective are essential. Quasi-contestability offers formal microfoundations for a workable competition concept. In doing so, the theory can be a welfare-theoretic yardstick for the evaluation of market rivalry.
REFERENCES


COMPLEXITY AND MARKET PERFORMANCE


Witteloostuijn, A. van, Rationality, Competition and Evolution: Entry (Deterrence) in Dynamic Barrier Market Theory, Maastricht, 1990a.


Summary

THEORIES OF COMPETITION AND MARKET PERFORMANCE

An important issue in the theory of industrial organization involves the question of market performance. This paper deals with the static-dynamic efficiency trade-off. Theories of competition are reviewed and classified. The concept of workable competition offers a verbal listing of conditions facilitating both efficiency dimensions. A crucial feature of workable competition is the force of potential rivalry. The study of sources of potential entry and market games identifies cases of intermediate competition which can serve as a welfare-theoretic yardstick for the evaluation of market performance. The nature of multimarket competition proves to be essential.