The Theory of Contestable Markets

Stephen Martin
Department of Economics
Purdue University
smartin@purdue.edu

July 2000
## Contents

1 Contestable Markets ................................. 5
   1.1 Introduction .................................... 5
   1.2 Principal results ................................ 6
   1.3 Critical assumptions .............................. 9
      1.3.1 Entrants act as price takers ......... 10
      1.3.2 Exit lag .................................. 12
      1.3.3 Long-term contracts .................. 14
      1.3.4 Absence of sunk costs ................. 15
      1.3.5 No transactions costs in financial markets .... 19
      1.3.6 No product differentiation .......... 20
      1.3.7 Recapitulation ......................... 24
   1.4 Goals ........................................ 24
      1.4.1 Generalization of perfect competition .... 24
      1.4.2 Endogenization of industry structure ...... 26
      1.4.3 Guidelines for appropriateness of intervention .... 29
      1.4.4 Contestability and the deregulated US airline industry 30
      1.4.5 Contestability theory as a guide to areas for intervention 36
      1.4.6 Guidelines for conduct of intervention, if appropriate ... 38
   1.5 Empirical tests of contestability theory ........... 40
      1.5.1 Experimental evidence ..................... 40
      1.5.2 Cross-section studies ...................... 41
   1.6 Conclusion .................................... 43

References ............................................. 45
Chapter 1

Contestable Markets

“But he hasn’t got anything on,” a little child said.

Hans Christian Anderson, The Emperor’s New Clothes

1.1 Introduction

The literature on contestable markets emerged from a research program that claimed two principal achievements in advancing economic knowledge, and two important policy contributions.

The theory of contestable markets was advanced as a generalization of the theory of perfectly competitive markets, and a generalization that (in contrast with the previous literature) endogenizes the determination of industry structure. Thus (Baumol, 1982, p. 2)\(^1\)

\[\text{in the limiting case of perfect contestability, oligopolistic structure and behavior are freed entirely from their previous dependence on the conjectural variations of incumbents and, instead, these are generally determined uniquely...by the pressures of potential competition...} \]

Further, the theory of contestable markets was presented as suggesting an improved set of guidelines for determining when government intervention

in the market is called for, and for the conduct of such activity when it is undertaken.\textsuperscript{2}

We begin with a brief statement of the formal results of the theory of contestable markets. We then explore the assumptions that seem necessary to produce these results. This is followed by a review of the progress that the theory of contestable markets has made toward its goals, and of empirical work related to contestability theory.

\section{1.2 Principal results}

The formal structure of the theory of contestable markets is disarmingly simple, particularly for the case of single-product firms.\textsuperscript{3,4} The essential definitions are:

\begin{enumerate}[label=(D\arabic*)]
\item an \textit{industry configuration} is a vector \((m, y^1, y^2, \ldots, y^m, p)\).
\begin{itemize}
\item \(m\) is the number of firms. \(y^i\) is the output of firm \(i (i = 1, 2, \ldots, m)\).
\item \(p\) is the price that clears the market: \(Q(p) = y^1 + y^2 + \ldots + y^m\).
\end{itemize}
\item A configuration is \textit{feasible} if production is sufficient to meet demand, and no firm is losing money.
\item A configuration is \textit{sustainable} if it is feasible and no potential entrant can cut price and make a profit supplying a quantity less than or equal to the quantity demanded at the lower price.
\item A \textit{perfectly contestable market} is a market in which sustainability is a necessary condition for equilibrium.
\item A configuration is a \textit{long-run competitive equilibrium} if it is feasible and there is no output level at which any firm could earn an economic profit at the prevailing price.
\end{enumerate}

\textsuperscript{4}The discussion that follows can be extended to multiproduct firms by interpreting \(y^1\) as a vector of outputs, \(p\) as a vector of prices, and \(Q(p)\) as a vector of demand functions. See Waterson (1987).
The main results follow almost immediately from the definitions. They are as follows.

(R1) A long-run competitive equilibrium is sustainable.

By (D5), in long-run competitive equilibrium there is no output level that earns an economic profit at the prevailing price, and the existing configuration is feasible. This satisfies (D3), the definition of a sustainable equilibrium.

(R2) A sustainable configuration is not necessarily a long-run competitive equilibrium.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure1.png}
\caption{Sustainability versus long-run competitive equilibrium}
\end{figure}

It suffices to provide an example of a sustainable configuration that is not a long-run competitive equilibrium. Let the demand curve be \( p = 7 - Q \). Let the cost function be \( c(q) = 4 + 2q \). Then \((m, y^1, p) = (1, 4, 3)\) is a sustainable equilibrium. If one firm sells 4 units of output, average cost is 3. 3 is also the price that clears the market. In this configuration, output equals
demand; price equals average cost; the firm is not losing money. Thus the
configuration is feasible. But at price \( p = 3 \), a firm producing more than 4
units of output would earn an economic profit, since average cost falls below
3 as output rises above 4 (see figure 1.1).

Thus \((m, y^1, p) = (1, 4, 3)\) is not a long-run competitive equilibrium. It
is perfectly correct that in this example a firm producing more than 4 units
of output could not sell it at a price equal to 3. To qualify as a long-run
competitive equilibrium, however, a firm that takes prices as given must be
unable to earn a profit at any output level.

(R3) If a configuration is sustainable

(a) each firm earns zero profit
(b) price is not less than marginal cost.

Proof of (a): if an incumbent firm were earning a positive profit, an
entrant could duplicate the incumbent’s output, cut price slightly, and still
earn a profit. But this would violate the definition (D3) of sustainability.
Hence each incumbent must earn zero profit in a sustainable equilibrium.

Proof of (b): if price were less than marginal cost and a firm were earn-
ing zero profit, it could reduce output slightly and collect a positive profit,
vio\(lating\) (a).

(R4) If at least two firms supply a good in sustainable configuration, then
price equals marginal cost for all firms.

Result (R3) establishes that price is not less than marginal cost. If price
were strictly greater than marginal cost, an entrant could supply an output
slightly larger than that of some one of the incumbent firms, at the market
price or perhaps a little less, and earn a greater profit than the incumbent
firm. But in sustainable equilibrium incumbents earn zero profit, so that
the profits of the entrant would be strictly positive. But this violates the
definition of sustainability.

(R5) For a single-product sustainable configuration with at least two firms,
each firm operates where returns to scale are constant.\(^5\)

\(^5\)When marginal cost is less than average cost, average cost falls as output rises: in-
creasing returns to scale. When marginal cost is greater than average cost, average cost
rises as output rises: decreasing returns to scale. The conventional neoclassical index
of the nature of returns to scale, the function coefficient, is the ratio of average cost to
marginal cost, and takes the value 1 when returns to scale are constant.
1.3. CRITICAL ASSUMPTIONS

By result (R3), each firm earns zero profit, which means price equals average cost. By (R4), price equals marginal cost. When marginal cost equals average cost, returns to scale are constant.

(R6) A sustainable configuration minimizes the cost of production.

If this were not so, it would be possible to produce the same total output at lower cost by rearranging output among firms. The same output would sell at the same price. With the same output produced at lower cost but selling at the same price, some firm would earn a positive profit. But if some incumbent firm could earn a positive profit by producing a different output and selling it at the market price, an entrant could earn a positive profit by producing and selling the same amount. This violates the assumption that the original allocation of output among firms is sustainable.

It may be remarked that (R6) is immediate for sustainable equilibria with one firm. For the single-product case, (R6) is simply a way of restating (R5) if there are at least two operating firms. If there are at least two firms producing a single product in sustainable equilibrium, the firms are producing where returns to scale are constant, and cost is minimized no matter what the distribution of output among firms.

1.3 Critical assumptions

As Baumol et al. point out (1983, pp. 495–6), the results of the theory of contestable markets are of a strictly static and equilibrium nature. In a formal sense, the theorems provided by the theory of contestable markets are devoid of dynamic considerations. Yet the definitions from which these theorems are drawn have certain implications for the kind of out-of-equilibrium behavior that is consistent with the possibility that contestability provides an interesting characterization of equilibrium.

It is the possibility of rapid entry and exit that ensures the optimality properties of sustainable configurations. Thus (Baumol et al., 1982, p. 5)

We define a perfectly contestable market as one that is accessible to potential entrants and has the following two properties: First, the potential entrants can, without restriction, serve the same market demands and use the same productive techniques
as those available to the incumbent firms. . . . Second, the potential entrants evaluate the profitability of entry at the incumbent firms’ pre-entry prices.

and (Baumol, 1982, pp. 3-4)

A contestable market is one into which entry is absolutely free, and exit is absolutely costless. . . . the entrant suffers no disadvantage in terms of production technique or perceived quality relative to the incumbent, and that potential entrants find it appropriate to evaluate the profitability of entry in terms of the incumbent firms’ pre-entry prices. . . .

The crucial feature of a contestable market is its vulnerability to hit-and-run entry.

Much of the literature dealing with contestable markets has focused on the circumstances under which it is plausible for incumbents to believe that entrants could realistically engage in rapid and reversible (hit-and-run) entry. If incumbents do not find such behavior credible, potential entry does not constrain the actions of incumbents.

1.3.1 Entrants act as price takers

The results of contestability theory require not only that rapid entry and exit be possible, but that potential entrants make their decisions taking the market price as given. Thus definition (D3) defines sustainability in terms of entrant profitability given the number of incumbents, their output, and the price at which that output clears the market. Under the definition of sustainability, the entrant is not permitted to take account of the price reduction that its own output will produce when it assesses the profitability of entry. The entrant is not permitted to take account of possible reactions of incumbents. Hit-and-run entry is supposed to occur if the potential entrant could make a profit at the pre-entry price. If the potential entrant comes into the market only if it could make a profit at the expected post-entry price, hit-and-run entry is much less plausible. But if hit-and-run entry is implausible, there is no dynamic mechanism to enforce the static results of the theory of contestable markets.
1.3. CRITICAL ASSUMPTIONS

In their original discussion, Baumol et al. make a comparison between the entry assumptions of contestability theory and those of the standard model of perfect competition (1982, p. 5):

That is, although the potential entrants recognize that an expansion of industry output leads to lower prices-in accord with the market demand curves-the entrants nevertheless assume that if they undercut incumbents’ prices they can sell as much of the corresponding good as the quantity demanded by the market at their own prices. This is an extension of the axioms on entry behavior in the classical model of perfect competition that makes it possible to deal with the small-numbers case.

The assumption of price-taking behavior is plausible for a model that describes the behavior of many small buyers and sellers. It is not plausible for markets with a small number of firms (in the limit, one or two actual firms with one potential entrant). It is clear that Baumol et al. recognize this (1982, p. 11):

Bertrand-Nash [price-taking] expectations are not always fulfilled, and in some cases they are unlikely to be. For...competitive entry can impose significant losses upon incumbents and thereby force a change in their prices. However, if an entrant’s output is “small” relative to that of the industry, the magnitude of these required adjustments may also be “small”, and hence it may be justifiable for the entrant to ignore them.

On this rationale, the theory of contestable markets can no longer be said to apply where technology requires firms to be large relative to the market. The theory of contestable markets applies where efficient firms can be so small that they make decisions taking price as given. This, of course, is the usual size condition imposed for applicability of the theory of perfectly competitive markets. If this size condition must be imposed before the theory of contestable markets is to be valid, it is difficult to see how it can be claimed that contestability theory extends the theory of perfect competition to markets where economies of scale are important.

Thus (Baumol and Willig, 1986, p. 17)
the critical issue that remains is the determination of the circumstances under which the Bertrand-Nash [price-taking] assumption holds or at least is assumed by the participants to hold approximately.

Most economists would reject the claim that price-taking behavior is a reasonable assumption to make about oligopolistic markets (Friedman, 1982, pp. 527–8):

The entry or exit of one firm from such a market is likely to be followed by a large discrete change in the policies being pursued (prices, etc.) by the firms which are active before and after the change. Even if none changed its behavior, amounts demanded of each firm would be much different after the alteration in the number of firms. That is to say, for the competitive firm it is reasonable for it to suppose that the prices, profitability, etc. of a given market are independent of whether it is in the market. For an oligopolist, however, it is abundantly clear that these variables must depend, in part, on whether or not it is active.

1.3.2 Exit lag

Baumol, Panzar, and Willig point out that an entrant need not believe prices are fixed forever in order for the results of contestable market theory to hold. The entrant need only believe that prices will not change for the duration of its stay in the market (Baumol et al., 1983, p. 493; emphasis in original):

To produce its results, even the limiting case of perfect contestability does not require entry and exit to be instantaneous. Rather, it is sufficient that the process be rapid enough so that the entrant does not find his investment vulnerable to a retaliatory response by the incumbent. *The length of this time period is not exclusively a technological datum, but is also the result of business practice and opportunities* in the market in question.

The observation that the length of time before an entrant would expect incumbents to respond reflects business practice and opportunities would seem to conflict with the claim that performance in contestable markets is determined in a way that is independent of oligopolistic interactions.
A simple model due to Schwartz (1986) shows that entry for any period, however short, can be rendered unprofitable if the incumbent can respond with sufficient speed.\(^6\)

Consider a market supplied by a monopolist charging the monopoly price. If entry occurs at all, it entails an investment \(F\) of fixed cost, with instantaneous rental cost \(rF\). If entry occurs, it must occur for a period of length \(X\). An entrant can undercut the monopolist and capture virtually the entire monopoly profit. At a time \(T < X\), the incumbent can react and cut price. In the worst case, price would fall to a level that would cover only variable cost, so the entrant would lose \(rF\) per instant after entry. Given these losses, exit would occur at time \(X\).

The present discounted value of the entrant’s income stream over the whole period from entry to exit is

\[
V = \pi_m \int_{t=0}^{T} e^{-rt} dt - rF \int_{t=T}^{X} e^{-rt} dt. \tag{1.1}
\]

The first term, which is positive, becomes smaller as \(T\) becomes smaller. The second term, which is negative, becomes larger as \(T\) becomes smaller. Thus for \(T\) sufficiently small \(V\) is negative. With a negative expected present discounted return, entry would not occur, even with the temptation of monopoly profits.\(^7\) A reasonable conclusion seems to be (Bailey and Baumol, 1984, pp. 114–5)\(^8\)

\(^6\)See also Pashigian (1968) and Bhaskar (1989).

\(^7\)Carrying out the integration in (1.1), we obtain

\[
V = (1 - e^{-rT}) \frac{\pi_m}{F} - (e^{-rT} - e^{-rX})F
\]

and so

\[
\frac{\partial V}{\partial T} = (\pi_m + rF)e^{-rT} > 0
\]

\(V\) is positive for \(T = X\), negative for \(T = 0\), and the first derivative of \(V\) with respect to \(T\) is positive. Thus there is some minimum value of \(T\), say \(T^*\), that makes \(V = 0\). An explicit expression for \(T^*\) can be obtained by setting \(V = 0\) in (1.1). If the incumbent can respond in time \(T^*\), the return to entry will be less than or equal to zero, and a profit-maximizing potential entrant would stay out of the market.

\(^8\)They continue by arguing that the absence of sunk costs or the possibility for incumbents to arrange long-term contracts prior to entry may resuscitate contestability. These arguments are discussed shortly.
where incumbents can counterattack quickly, contestability will
prevail only if hit-and-run entry can be carried out even more
rapidly.

Oligopolistic interactions among incumbents would reduce pre-retaliation
profit below the monopoly level. This would make entry even less likely. It
is always possible that under some circumstances incumbents would prefer
to accommodate entry rather than contest it. The expectation of accommo-
dating behavior would make entry more likely. All these eventualities mean
that market performance is determined by oligopolistic interactions, not by
the force of potential competition alone.

1.3.3 Long-term contracts

Baumol et al. present another dynamic mechanism that might allow potential
competition to determine market performance, even where hit-and-run entry
without a price response by incumbents is implausible: long-term contracts,
signed before entry, that protect the entrant from price retaliation because
they cover a period at least as long as the time it takes incumbents to react

With contracting feasible, the fact that successful entry may re-
quire commitment of assets to a particular market for a nontrivial
interval of time need not diminish the viability of hit-and-run en-
try, nor imply the presence of costs which are sunk in any econom-
ically significant sense. Even if such contracting is not feasible,
it is still possible for regulation, costs of communicating price re-
visions, or other impediments to delay an incumbent’s effective
price response for a period of length \( T > 0 \). In either case, the
concept of hit-and-run entry can survive the technological impo-
sition of a minimum production-time requirement \( t^* \leq T \).

The possibility that an entrant could protect itself from retaliatory price
responses in this way seems implausible, particularly if efficient operation
requires entry at large scale relative to the market – which is precisely when
it was claimed that contestability theory offered a generalization of the theory
of competitive markets.

9 An argument related to those of Chadwick (1859) and Demsetz (1968).
1.3. CRITICAL ASSUMPTIONS

If scale economies are large, an entrant would have to inform and negotiate contracts with a large fraction of the market to insulate itself from a retaliatory price response. Transactions costs rooted in imperfect and impacted information make negotiations difficult. The transaction costs reflect the fact that consumers have no experience with the potential entrant’s product, and they have no way of knowing whether or not the claims the entrant makes about its ability to perform can be relied upon. Such transactions costs—an investment in information and reputation—are largely sunk in the market, irrecoverable upon exit.\(^\text{10}\)

Furthermore, in a market where potential entrants could negotiate contracts in advance of entry, the same option would be available to incumbents. Incumbents could defeat a potential entrant’s strategy by negotiating contracts promising to “undercut any legitimate price” (Shepherd, 1984, p. 576, fn. 12). The possibility that potential competition in the offering of contracts insures optimal market performance seems unlikely.\(^\text{11}\)

1.3.4 Absence of sunk costs

Suppose rapid entry and exit is possible, that incumbents will not alter price in the event of entry, and that potential entrants know this. Suppose also that to enter at all, a firm would have to make an irrecoverable investment of 1,000,000. The investment might reflect the acquisition of information about the market before the decision to enter is taken. The investment might reflect the purchase of physical assets that could not be resold or transferred to other markets in the event of exit. So long as the firm remains in the market, the assets (tangible or intangible) are productive, and the firm enjoys the value of the marginal product of the assets as part of its income stream. If the firm exits, the investment is left behind.

It follows that incumbents, for which any such sunk investments lie in the past, can earn a combination of rent and economic profit with a present discounted value of up to 1,000,000 without inducing entry. Very early in their book, Baumol et al. emphasize the importance of sunk costs for market contestability (1982, p. 7).\(^\text{12}\)

\(^{10}\)Essentially the same arguments rule out long-term contracts as a device to combat predatory pricing.


\(^{12}\)Cairns and Mahabir (1988) present a sharply different analysis of the impact of sunk
Clearly, when entry requires the sinking of substantial costs, it will not be reversible because, by definition, the sunk costs are not recoverable. However, if efficient operation requires no sunk outlays, then entry can, by and large, be presumed to be reversible, and the market can be presumed to be contestable.

Various definitions of sunk cost appear in the contestability literature. Thus (Baumol et al., 1982, pp. 280-1)

Definition 10A1: Long-Run Fixed Cost

Long-run fixed cost is the magnitude $F(w)$ in the long-run total cost function

$$C_L(y, w) = \delta F(w) + V(y, w)$$

where

$$\delta = \begin{cases} 
0 & \text{if } y = 0 \\
1 & \text{if } y > 0 
\end{cases}$$

$$\lim_{y \to 0} V(y, w) = V(0, w) = 0.$$ 

$V()$ is nondecreasing in all arguments, and $y$ and $w$ are, respectively, the vectors of output quantities and input prices.

Definition 10A2: Let $C(y, w, s)$ represent the short-run cost function, applicable for the flow of production, that occurs $s$ units of time (years) in the future. Then, $K(w, s)$ are the costs sunk for at least $s$ years, if

$$C(y, w, s) = K(w, s) + G(y, w, s)$$

$$G(0, w, s) = 0.$$ 

Here, since in the long-run no costs are sunk,

$$\lim_{s \to \infty} K(w, s) = 0.$$ 

They argue that the most likely potential entrants are firms with sunk assets in related markets, and suggest that contestability theory best describes competition among multiproduct firms.
1.3. CRITICAL ASSUMPTIONS

What is called “costs sunk for at least s years” in Definition 10A2 looks very much like short-run fixed cost, where the short run lasts s years.\textsuperscript{13} Nor is it clear why some costs might not be sunk even in the long run.\textsuperscript{14}

A quite different definition of sunk costs, that permits costs to be sunk over the long run, is given by Baumol et al. (1983, p. 494):

\begin{quote}
\textit{suppose that a unit of capital purchased at a price of }\beta\text{ per unit could be sold or utilized elsewhere... for a unit salvage value of }\alpha \leq \beta\text{. Thus it is possible to parametrize continuously the degree of sunkness of capital from zero (}\alpha = \beta\text{) to absolute sunkness (}\alpha = 0\text{).}
\end{quote}

According to this definition, sunkness depends on the nature of resale markets for capital assets. But if this is what determines whether or not costs are sunk, the possibility that sunk costs are completely absent – as perfect contestability requires – seems extremely limited.

Consider first physical assets. Many businesses require highly specific physical assets that might be resold at a substantial loss upon exit or not at all. A firm that wished to leave a retail food market because of low or negative profitability might recover some of its investment in kitchen and other equipment by selling it in that market. If the market is so depressed that the firm has decided to exit, it will most likely sell at a loss. The exiting firm might get a better price by shipping the physical assets to another market, which would involve the expense of transportation. The more specific the physical assets, the greater the extent to which the investment in the assets is sunk.

Other businesses involves assets that are not specific – they can be used as inputs to produce many different goods. Trucks and delivery vans are examples. Resale markets for such assets, however, suffer from the “lemon problem” of Akerlof (1970) – a consequence of imperfect and impacted information.

\textsuperscript{13}Weitzman (1983) argues that without sunk costs there cannot be truly fixed costs. In the absence of sunk costs, an entrant could come into a market, operate very briefly at minimum efficient scale, and exit. Provided the good can be stored, any production that occurs will take place at minimum efficient scale. But then returns to scale are effectively constant.

\textsuperscript{14}Investment in research and development creates an asset – knowledge – that is, at least potentially, productive forever. Much of the value of such knowledge, if highly specific and tied to the operations of the firm, would be lost upon exit.
CHAPTER 1. CONTESTABLE MARKETS

Some used delivery vans of perfectly acceptable quality may be on the market because they are owned by firms that are closing down operations. Some delivery vans of unacceptable quality may be on the market because their owners are trying to sell problems to someone else. Sellers of used vans know with certainty whether or not they are selling low-quality vans, but buyers do not. Further, buyers of used vans are not able to rely on the representations of sellers as to quality. It is in the interest of all sellers to represent themselves as offering high-quality used vans, because high-quality used vans sell at a higher price than low-quality used vans. Because of uncertain information about the quality of used vans, high-quality used vans sell at a price less than their quality would justify, in an objective omniscient sense.  

Thus if we confine our attention to physical assets, sunk costs are absent when firms employ nonspecific physical assets, the quality of which can be easily ascertained by purchasers in the event of resale. Otherwise, some of the cost of physical assets are sunk, and the theory of perfectly contestable markets fails.

Intangible assets are involved in every entry decision. A potential entrant will invest in information about the target market before the decision to enter is made. This information is valuable to the firm so long as it remains in the market, but it cannot be resold.

A potential entrant invests the time and ability of its corporate executives in organizing the new operation. This investment creates an asset — the internal organization of the new operation — that is valuable to the firm so long as it remains in the market, but which is unlikely to be susceptible to resale at anything like its cost of production.

In markets where product differentiation is important, an entrant will invest resources in product-differentiating sales efforts. Upon exit, the firm might recover a fraction of this investment for goodwill, but much of it will be sunk (Stiglitz, 1987, p. 889): “An airline must advertise to obtain customers; it must solve complicated routing problems.” Such investments create assets for a going firm, but their cost is largely sunk.

In technologically progressive industries, an entrant needs to invest in research and development to maintain viability. But (Stiglitz, 1987, p. 889):

Most expenditures on R&D are, by their very nature, sunk costs.

---

15 This is essentially the same logic that suggests that entrant and fringe firms in a market will have a higher cost of capital than leading incumbent firms.
1.3. CRITICAL ASSUMPTIONS

The resources spent on a scientist to do research cannot be recovered. Once his time is spent, it is spent.

Investment in knowledge, too, is valuable to the going firm. That value will be largely lost upon exit.

In short, sunk costs are ubiquitous for real world firms. By implication, the theory of contestable markets, particularly in its perfect form, is largely inapplicable to the real world.

1.3.5 No transactions costs in financial markets

Suppose there are no sunk costs at all, but firms finance some investment on markets for financial capital. Transaction costs on financial markets suggest that entrants and fringe firms will have a higher cost of capital than incumbents (Martin, 1989b).

The logic behind this argument is similar to that which suggests that the presence of some “lemons” on a market for capital goods will drive down the price of all capital assets, whether lemons or not. Some potential entrants may be perfectly capable of setting up a successful enterprise in a market. Others will not – they are “lemons” in the queue of potential entrants, more likely than not to go bankrupt and fail to repay borrowed funds. The potential entrants know which is the lemon and which is not, but all potential entrants will claim that they are likely to successful. A borrower who can convince lenders that he is likely to succeed will receive a lower rate of interest on borrowed funds. Lenders are not able to distinguish between the two classes of entrants. As a result, potential entrants as a group pay a higher rate of interest on borrowed funds than incumbents.

But if this is the case, incumbent firms are able to engage in limit pricing, if they should find it profitable to do so. Hit-and-run entry is impossible if incumbents engage in limit pricing.

Whether or not limiting behavior is a preferred strategy depends on oligopolistic interactions among incumbents and between incumbents and entrants. The force of potential competition alone is therefore insufficient to determine market performance. If entrants finance investment on real-world capital markets – which operate under conditions of imperfect and impacted information – the theory of contestable markets fails.
1.3.6 No product differentiation

Baumol et al. argue (1982, pp. 329–32; 1986, pp. 355–9) that the theory of contestable markets applies to monopolistically competitive markets in which products are differentiated.16 This claim is based on an unusual interpretation of the theory of monopolistic competition.

Figure 1.2: Long-run equilibrium, monopolistic competition

Figure 1.2 is a standard illustration of long-run equilibrium in Chamberlinian monopolistic competition. Compare Baumol et al. (1982, figure 11F3). \(dd\) is the demand curve for a single variety in a product group if all firms hold price constant. \(DD\) is the demand curve for a single variety if all firms match price changes.

It is customary to describe the situation depicted in Figure 1.2 as a long-run equilibrium. Each firm is maximizing profit: marginal cost equals marginal revenue along the demand curve \(dd\) (for simplicity, the marginal revenue and marginal cost curves are omitted from Figure 1.2). Since each firm is

---

16 We discuss horizontal product differentiation; see Lambertini (1992, 1996) for a discussion of contestability theory and vertical product differentiation.
maximizing profit, there is no incentive for any firm to alter its own behavior. At the same time, economic profit is zero (price equals average cost). There is no incentive for entry or exit. With no incentive for entry or exit, and no incentive for incumbents to alter behavior, the configuration illustrated in Figure 1.2 will persist until some other factor alters demand or cost conditions. This is usually thought of as an equilibrium. Baumol et al. disagree, if the market is contestable (1982, p. 332; footnote omitted; emphasis added):

However, such a position cannot be an equilibrium in a contestable market. An entrant can closely or exactly duplicate the product design of the firm depicted, and enter at a lower price. In a contestable market, with an arbitrarily short lag in incumbents’ price reactions, he would expect demand to behave in accord with curve dd during the period of the lag. Since dd is more elastic than the DD curve that is tangent to the A[verage]C[ost] curve, the dd curve necessarily cuts into and above the AC curve, as illustrated. Consequently, there exist (temporarily) profitable entry opportunities, and so “high-tangency equilibria” cannot be equilibria in contestable markets. It should be noted that Chamberlin’s discussion implies strongly that the markets of which he was thinking satisfied the free-entry requirements of contestability. Thus, it is of some significance for his analysis that his high-tangency solution turns out to be unsustainable.

But the distinguishing characteristic of monopolistic competition is product differentiation. Under monopolistic competition, it is impossible for an entrant to duplicate exactly the variety of an existing firm (Chamberlin, 1933, p. 56):

A general class of product is differentiated if any significant basis exists for distinguishing the goods (or services) of one seller from those of another. Such a basis may be real or fancied, so long as it is of any importance whatever to buyers, and leads to a preference for one variety over another. . . .

Differentiation may be based upon certain characteristics of the product itself, such as exclusive patented features; trademarks; trade names; peculiarities of the package or container, if any; or singularity in quality, design, color, or style. It may
also exist with respect to the conditions surrounding its sale. In retail trade...these conditions include such factors as the convenience of the seller’s location, the general tone or character of his establishment, efficiency, and all the personal links which attach his customers either to himself or to those employed by him. In so far as these and other intangible factors vary from seller to seller, the “product” in each case is different.

As product differentiation has usually been characterized by economists, it would be impossible for a hit-and-run entrant to duplicate the variety of any existing firm exactly. Close approximation might be possible, but it would not produce the results of the theory of contestable markets, as Bailey and Baumol acknowledge (1984, p. 117, fn. 10):

A problem can arise if products in a perfectly contestable industry are heterogeneous, each supplier offering his own special brands with their own special features. However, it can be shown...that if each variant is sold by at least two different suppliers, perfect contestability will lead to marginal cost pricing.

By definition, however, if products are differentiated it is impossible for any variant to be produced by two suppliers. Kellogg’s Corn Flakes and Martin’s Corn Flakes are not the same product, even if they are physically identical (Caves, 1971, p. 5):17

In the nature of differentiation, a successful (rent-yielding) product variety is protected from exact imitation by trade markets, high costs of physical imitation, or both.

Hit-and-run entry is not possible where products are physically identical but bear different brand names. Aspirin is the generic example. As Farrell (1986) points out, if varieties produced by entrants and incumbents are identical but buyers are uncertain about the quality of entrants’ products, buyers will be reluctant to patronize an entrant, all else equal. The result is an entry barrier that impedes hit-and-run entry.18

---

17 See also Friedman (1982, p. 501): “At the heart of differentiated products models are the assumptions that no two firms produce identical products and that firms can be grouped according to the type of product they make.”

18 See also Seabright (1990, pp. 20–35).
Where products are differentiated, an entrant would have to introduce a new variety of the product. It would operate on its own set of demand curves ($dd$ and $DD$). If Figure 1.2 describes the pre-entry equilibrium, then entry would push the $DD$ curves of entrant and incumbents below the corresponding average cost curves. As the $DD$ curves show the actual, rather than the expected, relationship between prices and quantities, the entrant would lose money. Profit-maximizing firms should not be expected to enter an industry in a configuration of the kind illustrated in Figure 1.2.

If incumbents producing a group of differentiated products are earning economic profits, entry by a firm producing a new variety will move the market toward the sort of equilibrium illustrated in Figure 1.2. In this case, however, it is actual entry and competition, not potential entry and competition, which brings about an improvement in market performance. It seems likely that this describes events in the US stock brokerage market following the 1975 relaxation of regulatory constraints on commissions (Bailey, 1986, p. 10):

Discount firms have thus entered and been financially successful, but even though their rates are less than half those of the full-service brokers, they have not in any sense taken over the market from the full-line firms. Both groups have prospered.

Even though discount brokerage houses charge substantially lower commission rates than full-service houses, they do not capture the entire market. Yet the results of contestability theory require incumbents to believe that they would lose the entire market to an entrant charging a slightly lower price. It is implausible to suggest that incumbents producing varieties of a differentiated product would hold such a belief.

Chamberlin’s discussion, quoted above, concludes that (1933, p. 57)

it is evident that virtually all products are differentiated, at least slightly, and that over a wide range of economic activity differentiation is of considerable importance.

Thus, it is of some significance that the theory of contestable markets does not apply where products are differentiated.
1.3.7 Recapitulation

The theory of perfectly contestable markets yields results that are strictly static and refer to long run market equilibrium. For these results to be produced:

1. either incumbents must believe that potential entrants make the decision to enter on the assumption that incumbents’ prices are fixed, or at least could not be changed before an entrant could (costlessly) exit;

2. or incumbents must believe that potential entrants could protect themselves from retaliation by signing long-term contracts before entry;

3. incumbents must believe that an entrant could capture the entire market with a slight price cut;

4. sunk costs must be completely absent;

5. the cost of financial capital must be the same for entrants and incumbents;

6. products must be absolutely standardized.

1.4 Goals

1.4.1 Generalization of perfect competition

The theory of contestable markets was advanced as a substantial generalization of the theory of perfect competition (Bailey, 1982, p. xix):

The notion of contestable markets offers a generalization of the notion of purely competitive markets, a generalization in which fewer assumptions need to be made to obtain the usual efficiency results. Using contestability theory, economists no longer need to assume that efficient outcomes occur only when there are large numbers of actively producing firms, each of whom bases its decisions on the belief that it is so small as not to affect price. What drives contestability theory is the possibility of costlessly reversible entry.
But the theory of competitive markets produces the usual efficiency results only in long-run equilibrium. Economic profits or losses are quite possible in the short run in competitive markets.

The theory of competitive markets has a well-established theory about passage from the short run to the long run. If in an initial long-run equilibrium there is an exogenous outward shift in the demand curve, incumbent firms earn economic profits in the short run; these profits attract entry; the industry supply curve shifts out, price falls, profit is reduced. This process continues until profits are eliminated. There is a similar story with exit if in the short run incumbent firms suffer losses. The theory of contestable markets, which is acknowledged to be strictly static in nature, lacks the short-run and dynamic implications of the theory of competitive markets. In this sense, the theory of contestable markets is less general, not more general, than the theory of competitive markets.

Contestability theory is particularly defended as providing a welfare standard when the technology mandates an oligopolistic market structure (Bailey and Baumol, 1984, p. 119):

while perfectly competitive and perfectly contestable markets are both ideals, the latter is more ideal than the former. After all, one must be tempered in one’s praise of the many-firm structure of perfect competition in those cases in which the availability of economies of scale and scope means that an oligopoly structure can (perhaps) achieve far lower costs and offer far lower prices to consumers.

Two remarks ought to be made. First, it seems clear by now that the assumption of costlessly reversible entry, by itself, is insufficient to produce the results of long-run equilibrium in a perfectly contestable market. Either entrants must make decisions on the assumption that price will not change after entry, or it must be possible for entrants to negotiate contracts, before entry, that insulate them from post-entry responses by incumbents. There must be no sunk costs, not even sunk costs associated with the process of collecting information about the target market. Incumbents must believe that entrants who resort to financial markets would be able to raise capital on the same terms as incumbents. Products cannot be differentiated. Where these assumptions fail—which seems likely to be almost everywhere in the economy—one must be tempered in one’s praise of the costless entry-and-exit requirement of perfect contestability.
Second, the bulk of empirical evidence is that economies of scale are not important in modern economies. Average cost curves in most industries appear to flatten out at relatively small market shares, and observed levels of market concentration exceed, often substantially, that which can be explained in terms of economies of scale.\textsuperscript{19}

Yet the theory of perfect contestability does generalize a received body of economic theory — Bertrand’s (1883) model of price-taking oligopoly with standardized products.\textsuperscript{20} Bertrand obtains the efficiency results of long-run competitive equilibrium under assumptions strikingly similar to those of the theory of contestable markets. In Bertrand’s model, products are standardized. Firms set price taking rivals’ prices as fixed. The entire market is assumed to shift from one supplier to another in response to tiny price differences. If there are at least two suppliers in the market, performance is optimal.

The theory of imperfectly contestable markets, on the other hand, is now acknowledged to be an extension of the mainstream structure-conduct-performance school of industrial economics (Baumol et al., 1983, p. 494):\textsuperscript{21}  
models which support the robustness of contestability analysis follow a relatively long tradition going back at least to the work of J. S. Bain. This tradition holds that increased ease of entry and exit improves the welfare performance of firms and industries. On this subject, the theory of contestable markets has only sought to contribute insights on the underpinnings of that judgment.

The tradition referred to also holds that difficulty of entry allows incumbent firms to exercise some market power, and that market performance depends on oligopolistic interactions as well as potential competition.

\subsection*{1.4.2 Endogenization of industry structure}

One of the primary claims of the theory of contestable markets is that, in contrast with the previous literature, it endogenizes industry structure. This claim seems doubtful on two counts.

\textsuperscript{19}Scherer (1974). Baumol et al. recognize that average cost curves are typically found to flatten out (1982, fn. 50).

\textsuperscript{20}The relation of contestability theory to Bertrand’s work is emphasized by Knieps and Vogelsang (1982).

\textsuperscript{21}See also Schwartz (1986, p. 38).
First, the literature before contestability did not treat market structure as exogenously given. Second, the sense in which perfectly contestable markets provides a theory of market structure is extremely limited.

Analysis of market structure before contestability theory

There is a large literature in which economists develop theoretical and empirical models of market structure. A common simplifying assumption in models of oligopoly is that firms are symmetric — that they produce the same output in equilibrium. In this case, the critical element of market structure is the number of firms in the industry. Long before contestability theory, economists modeled the long-run equilibrium number of symmetric firms in a market on the assumption that entry occurs in response to excess profits.\textsuperscript{22}

The literature on dynamic limit pricing\textsuperscript{23} analyzes the optimal pricing strategy for a dominant firm that faces the possibility of fringe entry, over time, if price exceeds a critical level. An essential result of dynamic limit pricing models is that a dominant firm without a cost advantage will set a high short-run price and gradually lower price to a level that no longer induces entry. A high short-run price generates short-run profits but future loss of market share.

Models of dynamic limit pricing yield predictions about the time path of the dominant firm’s market share and about the equilibrium shares of the fringe and the dominant firm. When firms are not symmetric, market share is a critical element of market structure.

There is a large empirical literature that endogenizes the determination of industry structure (Weiss, 1963a; Shepherd, 1964; Carter, 1967; Ornstein et al., 1973; Mueller and Hamm, 1974; Strickland and Weiss, 1976; Martin, 1979; Caves 1981b). Indeed, one of the few unchallenged empirical regularities in industrial economics is that, roughly 2 years after the appearance of a new Census of Manufactures, an econometric study will be published analyzing changes in US concentration. Nor is this literature limited to examining US data (Shepherd, 1966; Jenny and Weber, 1978).

This literature is much older than contestability theory. The exponents of contestability theory have admitted that with respect to the determination of industry structure, the empirical literature has contributions to make which

\textsuperscript{22}Howrey and Quandt (1968); Okuguchi (1972). See also Prescott (1973).
\textsuperscript{23}See Gaskins (1971) and Ireland (1972) for seminal contributions.
may well be more useful than those of the theory of contestable markets (Baumol and Willig, 1986, p. 15):\(^{24}\)

one suspects that empirical reality embodies relationships more robust and stable than does oligopoly theory in its current tumultuous state.

**Contestability as a theory of market structure**

What is it that contestability theory has to say about market structure? That cost of production is minimized in the long-run equilibrium of a perfectly contestable market. This is really a statement about industry performance. The theory of contestable markets yields few descriptive statements about industry structure.

In the case of single-product firms with at least two firms in the market the theory of contestable markets predicts that production will take place where returns to scale are constant. It is acknowledged that this is the leading empirical case (Baumol et al., 1982, p. 33):\(^{25}\)

The assumption that the \(\text{A[verage]C[ost]}\) curve is flat-bottomed is consistent... with the mass of empirical evidence accumulated over the last 25 or 30 years, beginning with the pioneering work of Joe S. Bain.

The most that contestability theory can do in this case is to impose bounds on the range over which the long-run equilibrium number of firms

\(^{24}\)This position is strikingly similar to that taken by Edward S. Mason roughly sixty years ago (1939, p. 62):

It would no doubt be extremely convenient if economists knew the shape of individual demand and cost curves and could proceed forthwith, by comparisons of price and marginal cost, to conclusions regarding the existing degree of monopoly power. The extent to which monopoly theorists, however, refrain from an empirical application of their formulae is rather striking. The alternative, if more pedestrian, route follows the direction of ascertainable facts and makes use only of empirically applicable concepts.

In this connection, the general discussion that follows Stiglitz (1987) in which some participants urge that theorists should not be expected to produce empirically testable hypotheses, is also of interest.

\(^{25}\)See also Chamberlin (1933, p. 57).
will range (Baumol et al., 1982, p. 34, pp. 146-50). Within these limits, the allocation of industry output among firms is indeterminate. In the leading empirical case, contestability theory, even in the case of perfect contestability, does not specify market structure.

What contestability theory can specify is industry performance in long-run equilibrium. If average cost curves flatten out in a single-product industry in a range that permits at least two firms, cost of production will be minimized and price will equal average cost. This is a statement about market performance, not market structure. At a fundamental level, the theory of contestable markets is a theory of market performance in the long run. It is not a theory of market structure.

1.4.3 Guidelines for appropriateness of intervention

Baumol and Willig state (1986, p. 22; emphasis in original)

Contestability theory follows the lead of Bain, Sylos-Labini and others in stressing that potential competitors, like currently active competitors, can effectively constrain market power, so that when the number of incumbents in a market is few or even where only one firm is present, sufficiently low barriers to entry may make antitrust and regulatory attention unnecessary.

They continue

Since this viewpoint thoroughly antedates contestability theory it is not surprising that it has appeared in a variety of official policies.

In principle, then, the contribution of contestability theory in terms of clarifying areas in which public intervention is appropriate seems to be mainly to reinforce the existing advice of mainstream industrial economics. In application, however, the role of contestability theory has been somewhat different. It is useful to begin by noting an early call for caution in the policy application of the theory of contestable markets (Dixit, 1982, p. 16):

As a positive theory of market structure, it needs careful handling. In most cases in practice, production requires some commitments that can only be liquidated gradually, consumers assimilate and respond to price changes with some delay, and firms
need some time to calculate and implement price changes. Perfect contestability is the judgment that the third lag is the longest. . . . The traditional presumption in industrial organization is the opposite, that is, that prices can be changed more quickly than sunk capacity . . . .

In practice, careful empirical work in each specific context will have to be undertaken before we can say whether an industry is contestable and sustainable, and decide whether and what regulatory attention it requires.

Industrial economists early recognized the need for careful study before it would be possible to conclude that policymakers could prudently treat any particular industry as if it were contestable. This call for cautious application contrasts with the sometimes casual way in which it has been asserted that the prescriptions of contestability theory can be applied.

1.4.4 Contestability and the deregulated US airline industry

The airline industry is a case in point.26 The early contestability literature used passenger airline travel as the flagship example of a contestable market. Thus in 1981 (Bailey, 1981, pp. 179-80):

The new policies are based on the theory that both trucking and aviation markets are, in the absence of regulatory intervention, naturally contestable. Capital is highly divisible in the trucking industry, and there is every reason to suppose that market mechanisms will work. . . . Even in nondense city-pair markets in aviation . . . potential competition should be able to act as a potent force. This is true because the major portion of airline capital costs, the aircraft, can readily be moved from one market to another.

and (Bailey and Panzar, 1981, pp. 128–9)

---

26 It has also been argued that the theory of contestable markets applies to the barge transport industry. For discussion, and a negative assessment, see Tye (1985).
1.4. GOALS

Thus, there is no reason, a priori, to expect that economies of scale should lead to substantial barriers to entry in the airline industry because airline capital costs, while substantial, are not sunk costs. ... the major portion (i.e., aircraft) can be “recovered” from any particular market at little or no cost. Such factor mobility makes for ease of potential entry and exit in such markets. ... Thus, despite substantial natural monopoly attributes, most airline markets are likely to be readily contested. This fact ensures that, even if actually operated by a single firm, their performance should approach the competitive norm. . . .

The paper just cited is one of the earliest studies of the contestability of the passenger airline industry. It notes a number of factors that might impede contestability of a deregulated industry

1. State and local governments often find it convenient to lease airport facilities to particular airlines on a long-term basis. The airlines so favored can control rivals' access to airport facilities.

2. Local authorities in some areas limit airport access or growth for reasons of noise or pollution control.

3. Authorities ration slots (takeoffs and landings per hour) at the most congested airports.

4. Incumbents that operate many connecting flights from a single airport offer a quality of service that an entrant into a single city-pair market cannot duplicate.

The authors report a regression analysis that suggests (Bailey and Panzar, 1981, p. 143)

that actual competition of trunks was... an effective check on the pricing policies of local service carriers at mileage bands under 400 miles, while potential competition was the check in mileage bands over 400 miles. Potential competition between locals or between locals and commuters was not an effective check on the pricing policies of the locals. ... the gap in perceived quality of service (jets versus commuter aircraft) meant that commuters were not perceived by locals as potential entrants of sufficient stature to cause them to lower prices.
A number of these findings seem to conflict with the predictions of contestability theory. It was actual competition by trunk airlines, not potential competition, that limited the exercise of market power by local service carriers flying 400 miles or less. Potential competition by locals did not limit the exercise of market power by locals. Because of product differentiation — differences in the quality of service — potential competition from commuter airlines did not limit the exercise of market power by local service airlines. The only aspect of the results consistent with contestability theory is that potential competition from trunk airlines appears to limit the exercise of market power by local service airlines flying more than 400 miles.

The authors conclude (Bailey and Panzar, 1981, pp. 145):

We cannot claim to have done an exhaustive empirical analysis of airline markets in transition. However, we do feel that the admittedly scanty evidence during the first year after deregulation is consistent with our theory that airline markets are basically contestable.

In a perfectly contestable natural monopoly market, actual entry is redundant. The mere threat of entry will discipline the market even if it is a natural monopoly. We have argued that long haul airline markets served by local service carriers most closely fit this theoretical ideal. The empirical evidence of late 1979 and early 1980 does, in fact, bear us out. Local service monopolists have been pricing more or less competitively on their long-haul routes.

Airlines are again presented as a contestable market in Contestable Markets and The Theory of Industry Structure (Baumol et al., 1982, p. 7):

A clear example is provided by small, and therefore naturally monopolistic, airline markets. . . . because airline equipment (virtually “capital on wings”) is so very freely mobile, entry into the market can be fully reversible. In principle, faced with a profitable opportunity in such a market, an entrant need merely fly his airplane into the airport, undercut the incumbent’s price, and fly the route profitably. Then, should the incumbent respond with a sufficient price reduction, the entrepreneur need only fly his airplane away. . . . Thus, it is highly plausible that air travel provides real examples of contestable markets.
1.4. GOALS

<table>
<thead>
<tr>
<th></th>
<th>Equivalent number of equal-sized firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1979</td>
<td>1.00 1.25 1.67</td>
</tr>
<tr>
<td>1980</td>
<td>1.69 1.77 2.02</td>
</tr>
</tbody>
</table>

Table 1.1: Changes in city-pair airline market concentration (1979.IV–1980.IV); For the 200 largest markets, the average monopolized city-pair market in 1979 had the equivalent of 1-69 equal-sized firms in 1980; the average city-pair market with the equivalent of 1.25 equal-sized firms in 1979 had 1.77 equal-sized firms in 1980, and so on. Source: Herfindahl indices reported in Bailey et al., 1983, p. 58.

Additional tests of the contestability of the airline industry continued to appear. Graham et al. (1983, pp. 118–38) report an econometric analysis of airfares as a function of market concentration – measured by a Herfindahl index – and other variables describing market characteristics. They find that fares rise with concentration until the Herfindahl index reaches 0.5, but do not rise appreciably with additional increases in concentration.

A Herfindahl index of 0.5 or more characterizes a market supplied by two or fewer firms of equal size (Adelman, 1969). Thus the results of Graham et al. indicate that airfares rise with market concentration until a market is as concentrated as a duopoly of equal-sized firms. Yet contestability theory predicts that with two or more suppliers price will equal marginal cost (in long-run equilibrium in a perfectly contestable market). Thus Graham et al. find an effect of concentration on airfares precisely where contestability theory predicts that no effect will be found. They find that the force of potential competition is not sufficient to eliminate the exercise of market power.

Bailey et al. (1983) report on airline contestability at about the same time as Graham et al. They examine changes in concentration in airline city-pair markets following deregulation as a source of information on the presence or absence of entry barriers around such markets.

Concentration fell for almost all distance and size classifications examined; illustrative results are shown in Table 1.1. Monopoly or near-monopoly markets moved to near-duopoly and duopoly levels in one year. The authors conclude (Bailey et al., 1983, pp. 58–9):

This finding makes the premise behind the innate contestability of airline markets quite believable. If highly concentrated city-
pair markets are subject to quite rapid deconcentration, then it is entirely reasonable to suppose that a carrier serving a particular market views potential entry . . . as a force tempering his ability to raise fares above cost.

Examining information on fares, however, leads them to a different conclusion. Significantly higher fares at the hub airport in Atlanta are consistent with the argument that control of feeder traffic into the airport allows the two major airlines there to exercise market power. Airlines discriminate in price between business and tourist travellers. Price wars suggest that actual competition, not potential competition, limits the exercise of market power.

Despite low entry barriers in city-pair markets, there is a positive relationship between market concentration and airfares.

Bailey et al. (1985) present a study of the determinants of airfares in 1981. Treating market concentration as exogenous, which is supported by the results of a Hausman test, they find (1985, p. 165)

the fare in a market with two equal-sized competitors (i.e., with a Herfindahl index of 0.5) was 6 percent lower than in a monopoly market. A highly competitive market with four equal-sized airlines (i.e., a Herfindahl index of 0.25) is estimated to have an average fare that is about 11 percent below the monopoly fare.

The characterization of a market with four equal-sized firms as “highly competitive” is one which can reasonably be questioned. In a market with four equal-sized firms, each firm would surely be aware of the other three, and each firm would be aware that its own profit depended on the strategies of the other three. Oligopolistic rather than competitive behavior might well be expected. Indeed, this is consistent with the result that market concentration has a significant positive effect on the level of airfares. This result, of course, is inconsistent with the predictions of the theory of contestable markets.

Call and Keeler (1985) report a statistical study of airfares with similar results. They also find a significant positive effect of market concentration, as measured by the Herfindahl index. They compare the performance of the deregulated US airline industry with the performance of the California airline market in the 1950s and 1960s, and suggest that both are better described by declining dominant firm models than by contestability theory.

\[27\] See Chapter 6, footnote 42, and the associated text.
1.4. GOALS

If contestability theory applies to passenger airlines, they suggest, it is most likely to do so in the long run.

Hurdle et al. (1989) examine the determinants of revenue per passenger mile for a sample of 867 city-pair markets for 1985. They find that the number of likely potential entrants is a significant factor explaining differences across markets. But they find that the number and size distribution of incumbents is a factor as well, unless there are neither economies of scale nor economies of scope.

By 1986, contestability advocates had abandoned the position that the airline industry is inherently contestable (Baumol and Willig, 1986, p. 24):

In the initial enthusiasm with which we described contestability analysis we agreed with this assessment, and more than once cited the airline industry as a case in point, using the metaphoric argument that investments in aircraft do not incur any sunk costs because they constitute “capital on wings.” Reconsideration has led us to adopt a more qualified opinion on this score. . . . trucks, barges, and even buses may be more highly contestable than passenger air transportation. Barges and trucks have business firms . . . as their primary customers, and that facilitates the provision of services via contracts on which potential entrants can effectively bid against incumbents. . . . trucks and buses do not face the heavy sunk costs involved in the construction of airports or the shortage of gates and landing slots at busy airports. . . .

By this time, it is acknowledged that several features of airline markets make them imperfectly contestable (Baumol and Willig, 1986, p. 24):

First, . . . there have been constraining shortages of facilities and services of air traffic control at several pivotal airports. . . . Second, technological advances, changes in relative prices of jet fuel and equipment, and changes in the desired configurations of route networks have significantly altered the types and mix of aircraft demanded by the industry. As a result, there have been shortages in the availability of aircraft demanded. . . . Third, newly certified airlines have been able to avoid the costly labor contracts that pervaded the industry before deregulation, so that their labor costs have been substantially lower than those facing the older established carriers.
To these industry characteristics may be added the fact that incumbent airlines can respond to competitors’ rate changes instantaneously (Stockton, 1988, p. 1; see also Evans and Kessides, 1994):

With fare regulation gone and airlines now possessing sophisticated computer systems that monitor their own and competitors’ fares, literally hundreds of thousands of fares and the availability of discounts can change with lightning speed—even hourly. If, say, Midway Airlines lowered its fare between Cleveland and Omaha by $10, United, American, and Northwest would learn of it almost immediately through their computer connections and could react quickly if customers began defecting.

Some airline executives claim pricing is so flexible that on some days the industry introduces as many as one million fare changes.

In such an industry, it is hardly reasonable to expect that entrants will make decisions in the belief that incumbents’ prices are fixed.

The role of actual rather than potential competition in determining market performance is now admitted (Baumol and Willig, 1986, p. 25):28

Several econometric studies have confirmed the imperfection of the contestability of the airline market. . . there is a significant positive correlation between profits and concentration in airline markets. Thus the threat of entry does not by itself suffice to keep profits to zero. . . when new entry does occur, established carriers do reduce their fares in response, something one would expect in a conventional oligopolistic market.

1.4.5 Contestability theory as a guide to areas for intervention

In principle, the message of contestability theory is that policy intervention in market processes is unnecessary, if entry and exit are free and easy. In this case, potential competition as well as actual competition will influence market performance.

---

28 For later studies that confirm the imperfect contestability of the airline industry, see Morrison and Winston (1987) and Bailey and Williams (1988).
1.4. GOALS

<table>
<thead>
<tr>
<th></th>
<th>Pre-merger market shares</th>
<th>Post-merger market shares</th>
</tr>
</thead>
<tbody>
<tr>
<td>National</td>
<td>0.27</td>
<td>0.51</td>
</tr>
<tr>
<td>Texas International</td>
<td>0.24</td>
<td>0.23</td>
</tr>
<tr>
<td>Delta</td>
<td>0.23</td>
<td>0.23</td>
</tr>
<tr>
<td>Continental</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>Eastern</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>Approximate Herfindahl index</td>
<td>0.22</td>
<td>0.35</td>
</tr>
<tr>
<td>Numbers equivalent = 1/(approximate Herfindal index)</td>
<td>4.6</td>
<td>2.9</td>
</tr>
</tbody>
</table>

Table 1.2: Shares refer to the Houston–New Orleans market for the 12 months ending 30 June 1978. Remaining firms supply 2 percent of the market. Their contribution will affect the $H$ index only in the fourth decimal place and is ignored here. Source: Bailey (1981, p. 181).

In practice, the prescriptions of contestability theory seem often to have been applied in a free and easy manner, without the kind of detailed analysis necessary before it would be safe to conclude that a particular market could be treated as “workably contestable.”

The airline industry was early and easily asserted to be contestable. Real-world policy decisions were influenced by this position. Bailey (1981, p. 181) comments on a merger case that passed before the Civil Aeronautics Board in 1978, long before definitive evidence on the contestability of airlines was in hand. The prospective merger, as shown in table 11.2, involved market shares of the size that traditionally evoke policy concern.

Thus (Bailey, 1981, p. 181)

The share of the two leading firms was therefore 51 percent and would be almost 75 percent after a combination of Texas International and National. This number was greater than comparable figures in mergers declared unlawful by the Supreme Court. The C[ivil]A[eronautics]B[oard] countered by arguing that concentration ratios were not instructive in this case since with the passage of the Airline Deregulation Act of 1978...there was now relative ease of entry, even for small carriers, into such markets.

Subsequent research has demonstrated that the premise of this decision
was invalid. Potential competition is *not* sufficient to produce optimal performance in airline markets. The concentration ratios involved in this merger are well within the range in which – according to the results of Graham et al. (1983) – increases in concentration raise airfares and worsen market performance.\(^{29}\) The emphasis that the CAB placed on the force of potential competition in making this decision was misplaced.

Contestability advocates have cautioned against the blanket use of contestability theory to justify regulatory inaction (Baumol and Willig, 1986, p. 9):

> Contestability theory does not, and was not intended to, lend support to those who believe (or almost seem to believe) that the unrestrained market automatically solves all economic problems and that virtually all regulation and antitrust activity constitutes a pointless and costly source of economic inefficiency.

Yet the prescriptions of contestability theory continue to be offered in advance of the kind of economic analysis necessary to establish whether or not particular markets are workably contestable, much as was the case with the airline industry.\(^{30}\) It is difficult to resist the conclusion that, with reference to its role as a guide for policy intervention, the reach of contestability theory has exceeded its grasp.

### 1.4.6 Guidelines for conduct of intervention, if appropriate

Baumol and Willig suggest that (1986, p. 27)

> the viewpoint of contestability may make its main contribution... as a guide for regulation, rather than as an argument for its elimination.

\(^{29}\)For other evidence that airline mergers tend to result in higher fares, see Werden et al. (1989) and Borenstein (1989).

\(^{30}\)Baumol and Willig (1986, p. 23), discussing a merger case before the Federal Trade Commission, seem to suggest that barriers to entry in the automobile aftermarket are sufficiently low that a large post-merger firm could not exercise market power. This may be true or may not (product differentiation in particular seems likely to be a factor), but it cannot be said to be so established that policy decisions could safely treat the automobile aftermarket as workably contestable in any sense. See Crandall (1968).
1.4. GOALS

Contestability theory has served as a poor guide to areas in which regulation of antitrust activity is appropriate. There is little evidence that potential competition determines market performance in sectors of the economy that ordinarily attract the attention of regulatory and antitrust authorities, despite claims to the contrary. But contestability theory may offer insights into the conduct of public policy that sets the rules according to which firms compete.31

A simple way to summarize the insights that contestability offers to policymakers is to say that public authorities should make markets as contestable as possible, given the constraints imposed by other (and possibly noneconomic) goals.

Public authorities should not limit entry or exit. Limitations on either entry or exit reduce the force of potential competition.

Prices should not be set by arbitrary formulas, which tend to limit the impact of the interaction of demand and supply on price and to make prices unnecessarily rigid.

Nor should public authorities approve restraints on trade that tend to exclude competitors. The traditional antitrust hostility toward such practices as tying, full-line forcing, and resale price maintenance are consistent with the prescriptions of contestability theory. More often than not, the New Learning turns out to produce the same advice as the Old Learning.

Where production requires a large investment in sunk assets, public authorities should ensure access of all competitors to the use of the assets on equal terms (airports are an example).

If some segments of an industry are workably competitive, then those segments should be freed from regulatory activity.

In many ways, the most valuable policy contribution of the theory of contestable markets has been to emphasize, in a dramatic and successful way, that public policy toward business ought to be carried out for the benefit of the consumer, not of the firms that supply a market (or the stakeholders in those firms).

31 See Bailey (1981); Baumol and Willig (1983); and Bailey and Baumol (1984); and Evans and Kessides (1994).
1.5 Empirical tests of contestability theory

Tests of the contestability of the passenger airline industry, which are typical of industry-level studies of contestability, have been reviewed above. Two other types of empirical tests deserve mention.

1.5.1 Experimental evidence

Various studies have carried out experimental tests of markets that either conform closely to the assumptions of perfectly contestable markets or depart from those assumptions in controlled ways. Experimental subjects are given the roles of sellers or buyers, and make price, output, and purchase decisions. The experiments simulate the operation of the market for several periods. It is possible to examine how closely, if at all, performance conforms to the predictions of contestability theory.

Coursey et al. (1984a) examine an experimental market which, by construction, satisfies the assumptions of perfect contestability. The market is endowed with a technology that satisfies the conventional definition of natural monopoly: the average cost curve falls throughout the range of market demand. In four experiments, there is a single supplier. In six experiments, there are two suppliers, each able to supply the entire market demand. Sellers enter prices and quantities that will be offered at those prices; buyers select the suppliers from whom they will purchase.

In four of the six duopoly experiments, price and output moved to competitive levels. In the remaining two duopoly experiments, price remained between the monopoly and competitive levels, although appearing to tend toward the competitive level.

These results support the theory of perfect contestability. If there exist at least two firms, each capable of supplying the entire market, and the entire market moves from one supplier to another in response to small price differences, performance approaches the competitive extreme, even if average cost declines throughout the relevant output range.

---

32 See Baumol and Willig (1986) for references to empirical studies of the contestability of other industries.

33 In addition to the experiments discussed below, see Harrison and McKee (1985).

34 Contestability theory identifies subadditivity as the defining characteristic of natural monopoly; see Baumol et al. (1982).
1.5. EMPIRICAL TESTS OF CONTESTABILITY THEORY

Coursey et al. (1984b) extend the original work by adding sunk costs to the experimental design of the duopoly markets. One player, the incumbent, is compelled to purchase a permit that allows it to supply the market for ten periods. After the tenth period, the incumbent must purchase an additional permit, valid for five periods, if it is to continue to operate. The other player can purchase such a permit, if it wishes, in or after the sixth period. The price of the permit is an entry cost that is sunk for five periods.

They ran 12 trials of this experimental duopoly with sunk cost. In six of the trials, the tendency was for both firms to operate and for price to approach the competitive level. Four trials exhibited unstable pricing, in the sense that low prices would induce one supplier to exit, the survivor would raise prices, re-entry would occur, and price would fall. One trial showed an episode of limit pricing: five periods during which one player exited the market, with the remaining firm raising prices to the top of the competitive range. Two trials exhibited periods of tacit collusion: two firms in the market, with prices rising toward and falling from the monopoly level.

Taken as a group, the experiments seem far from supporting the hypothesis that markets that are almost perfectly contestable behave nearly as predicted by the theory of perfectly contestable markets. Contestability theory predicts that potential competition determines market performance. In 11 of the 12 trials reported, it was actual competition, or the lack of it, that determined market performance.

Only the experiment that produced an episode of limit pricing conforms closely to what one would expect if potential competition alone determines market performance. If actual competition must be examined to explain market performance, then an important goal of contestability — to free the analysis of market performance from dependence on analysis of oligopolistic interactions — fails.

1.5.2 Cross-section studies

In a study of entry into 266 US industries, Kessides (1986) finds evidence that the need to invest in advertising is a sunk cost of entry.\(^{35}\) This is plausible, as advertising creates an asset — goodwill or reputation — that is subject to extremely limited resale upon exit. At the same time, net entry is shown

\(^{35}\)The industries are defined at the four-digit Standard Industrial Classification (SIC) level. The study examines net entry — the number of entrants minus the number of exits — between 1972 and 1977 for industries where net entry was positive.
to rise with the industry advertising-sales ratio. This is consistent with the hypothesis that advertising is a device that entrants can employ to inform customers and carve out a niche in the marketplace.

Kessides also reports that net entry is less, all else equal, in more profitable industries. He interprets this result as suggesting that incumbents are more likely to react to entry in an aggressive way the greater the profits that would be lost under a more competitive regime.

Net entry is less, the greater the initial level of market concentration. The greater is initial concentration, all else equal, the more certain a single incumbent will be that it will benefit from its own efforts to deter entry. In an unconcentrated market, the benefits of entry-deterring activity accrue mainly to rivals.

Net entry, in short, is less, the greater the profits to defend and the fewer the firms that benefit from defending them. If such markets were perfectly contestable, incumbents would act in the belief that entrants treated price as given, and incumbents would set price at sustainable optimal levels. Market concentration would have no effect on net entry. Kessides’ evidence is against this proposition. Actual entry is affected by elements of market structure in ways that suggest that potential competition will have limited effect, in general, in tempering the exercise of market power.

Kessides (1988) examines the implications of sunk cost for market structure. He finds concentration to be greater, all else equal, the greater is minimum efficient scale, and lower, all else equal, the more firms in the industry are able to use rental or used capital equipment. Rental and used capital markets reduce the sunkenness of capital investments, so these results are consistent with the argument that it is the sunkenness of assets that creates barriers to entry.

If a market is perfectly contestable, its structure will be efficient in long run equilibrium. It follows that variance in market shares ought to be less, all else equal, the more contestable a market. Kessides also finds that the variance of market shares is less the more firms are able to use rental or used capital equipment. This is consistent with the hypotheses that market structure is endogenous, that market forces will tend to produce an efficient market structure, and that this tendency will be greater the less important are sunk costs.

\footnote{The study uses a sample of 400 1982 US four-digit SIC industries.}
1.6 Conclusion

The theory of perfectly contestable markets was presented as a generalization of the theory of perfectly competitive markets. It is at least as sensible to argue that it generalized Bertrand’s model of price-setting oligopoly with standardized products.

The theory of perfectly contestable markets was presented as a theory of industry structure. It was much more a theory of industry performance. The theory of perfectly contestable markets predicted only that market structure would be efficient. It had much less to say about the determinants of market structure than the extensive antecedent literature on this subject.

The theory of perfectly contestable markets was presented as suggesting areas appropriate for regulatory intervention. It was used as a slogan to defend a policy of deregulation à outrance. The US airline industry was deregulated in no small measure following arguments that it was naturally contestable. It is now generally admitted that this judgment was incorrect, and the deregulated US airline industry is well on its way to evolving into a concentrated oligopoly.

The theory of contestable markets was presented as providing guidelines for the conduct of regulation, if regulation was called for. Here the advice of contestability theory – to allow freedom of entry and exit, to permit price flexibility, and to ensure equal access of competitors – seems likely to offer improvements over the rigid regulatory practices of the past.

Experimental evidence suggests that perfectly contestable markets will behave as predicted by the theory of perfectly contestable markets, but that the performance of imperfectly contestable markets depends on actual rather than potential competition. Cross-section evidence confirms the importance of sunk costs in influencing market structure and performance. Such studies also suggest that entrants into concentrated markets are subject to strategic responses by incumbents, responses of a kind that are ruled out by assumption in perfectly contestable markets.

The theory of contestable markets aspired to be all things to all people. It was an uprising in the theory of industry structure, yet consistent with the mainstream structure–conduct–performance school of industrial economics, going back to the work of Joe S. Bain. It extended the results of long-run competitive equilibrium to markets in which the technology requires efficient firms to be large (relative to the market), but if there are at least two firms then each firm operates where returns to scale are constant. It
freed equilibrium from arbitrary assumptions about oligopolistic interactions among incumbents, but its results hold only if incumbents entertain a very particular set of beliefs about the way potential entrants behave.

Baumol et al. said of their work (1986, p. 361)

It provides a static (partial) equilibrium theory of industry structure, conduct, and performance more generally applicable than what was available before.

This view of the theory of contestable markets is not, by and large, shared by students of industrial economics.
References


1.6. CONCLUSION


Chadwick, Edwin “Results of different principles of legislation and administration in Europe of competition for the field, as compared with competition within the field, of service,” Journal of the Royal Statistical Society 22 September 1859, pp. 381–420.


Pashigian, B. Peter “Limit price and the market share of the leading firm,” 

Prescott, Edward C. “Market structure and monopoly profits: a dynamic 

Scherer, F. M. “Research and development resource allocation under ri-

— *Industrial Market Structure and Economic Performance*. Chicago: Rand 
McNally, 1970.

Schwartz, Marius “The nature and scope of contestability theory,” *Oxford 

Schwartz, Marius and Reynolds, Robert J. “Contestable markets: an upris-
ing in the theory of industry structure: comment,” *American Economic 

Seabright, Paul “Can small entry barriers have large effects on competi-
tion?,” Department of Applied Economics, Cambridge University, Eco-

Shepherd, William G. “Trends of concentration in American manufacturing 
12.

— “Changes in British industrial concentration,” *Oxford Economic Papers* 

— “‘Contestability’ vs. competition,” *American Economic Review* 74(4) 

— “On the core concepts of industrial economics,” in Henk W. de Jong and 
William G. Shepherd, editors *Mainstreams of Industrial Organization*. 

— “Entry barriers, contestability, and predatory pricing,” *Revue d’Economie 
1.6. CONCLUSION


Index

Bertrand model
theory of contestable markets, 26, 43

concentration
seller
Herfindahl index, 33

conjectural variations
Bertrand-Nash, 11, 12

contestable markets, theory of
assumptions, 9
Bertrand model, 26
empirical tests of, 40
generalization of structure-conduct-performance paradigm, 26
goals, 24
guide to areas for public intervention, 36
guide to conduct of public intervention, 38
guide to suitability of public intervention, 29
hit-and-run entry, 9
long-term contracts, 14
market performance, 29
market structure, 26
oligopolistic interactions, 12
passenger airline industry, 30
price-taking entrants, 10
product differentiation, 20
results, 6

sunk costs, 15
transaction costs in financial markets, 19

Herfindahl index
seller concentration, 33

limit pricing
dynamic, 27
in experimental markets, 41

monopolistic competition
contestable markets, 20
long-run equilibrium, 20

Names
Adelman, Morris A., 33, 45
Akerlof, George A., 17, 45
Anderson, Hans Christian, 5
Arrow, Kenneth J., 48
Bailey, Elizabeth E., 5, 13, 22–25, 30–34, 36, 37, 39, 45, 46
Bain, Joe S., 26, 28, 29, 43
Baumol, William J., 5, 6, 9–17, 20–22, 25, 26, 28, 29, 32, 35, 36, 38–40, 44–46
Bertrand, Joseph, 26, 46
Bhaskar, V., 13, 46
Borenstein, Severin, 38, 47
Brock, William A., 15, 47
Cairns, Robert D., 15, 47
INDEX

Call, Gregory D., 34, 47
Carter, Anne P., 27, 47
Caves, Richard E., 22, 27, 47
Chadwick, Edwin, 14, 47
Chamberlin, Edward H., 21, 23, 28, 47
Chevaillier, Margaret, 46
Coursey, Don, 40, 47
Crandall, Robert, 38, 47
Daughety, Andrew F., 46, 47
Demsetz, Harold, 14, 48
Dixit, Avinash, 29, 48
Evans, William N., 36, 39, 48
Farrell, Joseph, 22
Finsinger, Jorg, 45, 46
Friedman, James W., 12, 22, 46, 48
Gaskins, Darius W. Jr., 27, 48
Graham, David R., 33, 38, 45, 48
Hamm, Larry G., 27, 49
Harrison, Glenn W., 40, 48
Howrey, E. P., 27, 48
Hurdle, Gloria J., 35, 48
Intriligator, Michael D., 48, 49
Ireland, Norman J., 27, 48
Isaac, R. Mark, 47
Jenny, Frederic, 27, 48
Johnson, Richard L., 48, 51
Jong, Henk W. de, 50
Joskow, Andrew S., 48, 51
Kaplan, Daniel P., 45, 48
Keeler, Theodore E., 34, 47
Kessides, Ioannis N., 36, 39, 41, 42, 48, 49
Knieps, Gunter, 26, 49
Lambertini, Luca, 20, 49
Luke, Margaret, 47
Magnan de Bornier, Jean, 46
Mahabir, Dhanayshar, 15, 47
Martin, Stephen, 19, 27, 49
Mason, Edward S., 28, 49
Mathewson, G. Frank, 46
McKee, M., 40, 48
Morrison, Steven A., 36, 49
Mueller, Willard F., 27, 49
Okuguchi, Koji, 27, 49
Ornstein, Stanley L., 27, 49
Panzar, John C., 12, 30–32, 46
Pashigian, B. Peter, 13, 50
Prescott, Edward C., 27, 50
Quandt, Richard E., 27, 48
Reynolds, Robert J., 15, 50
Scherer, F. M., 26, 50
Schwartz, Marius, 13, 15, 26, 50
Seabright, Paul, 22, 50
Shepherd, William G., 6, 15, 27, 50
Sibley, David S., 45, 48
Smith, Vernon L., 47
Spence, A. Michael, 6, 51
Stiglitz, Joseph E., 18, 28, 46, 51
Stockton, William, 36, 51
Strickland, Allyn D., 27, 51
Sylos-Labini, Paolo, 29
Tye, William B., 30, 51
Vogelsang, Ingo, 26, 49
Waterson, Michael, 6
Weber, Andre-Paul, 27, 48
Weiss, Leonard W., 27, 51
Weitzman, Martin L., 17, 51
Werden, Gregory J., 38, 48, 51
Weston, J. Fred, 49
Williams, Jeffrey R., 36, 46
Williams, Michael A., 48
Willig, Robert D., 5, 6, 11, 12, 28, 29, 35, 36, 38–40, 46
Winston, Clifford, 36, 49

oligopoly models
   limit price, 19, 27, 41

product differentiation
   and theory of contestable markets, 20

sunk cost
   defined, 16