Contestability: Developments since the Book

William J. Baumol; Robert D. Willig

Oxford Economic Papers, New Series, Vol. 38, Supplement: Strategic Behaviour and Industrial Competition. (Nov., 1986), pp. 9-36.

Stable URL:
http://links.jstor.org/sici?sici=0030-7653%28198611%292%3A38%3C9%3ACDSTB%3E2.0.CO%3B2-F

Oxford Economic Papers is currently published by Oxford University Press.

CONTESTABILITY: DEVELOPMENTS SINCE THE BOOK

By WILLIAM J. BAUMOL* *and* ROBERT D. WILLIG†

THE ideas encompassed in the theory of contestable markets have entered both the academic literature and the discussions of practitioners with surprising speed. At hearings before regulatory agencies in antitrust cases, terms such as "economies of scope," "sustainable prices" and "contestable markets" are interjected casually as though long usage has made their connotations familiar to everyone. This is in part attributable to the steady stream of papers, published and about to be published by authors offering careful theoretical extensions and critiques, empirical and interpretative applications, as well as heated reactions to contestability theory.

The central focus of this paper is a summary of substantive developments relating to contestability since the appearance of our book. We will discuss four subjects: theoretical advances provided by others, econometric studies using the multiproduct cost concepts accompanying contestability analysis, market simulation studies testing various attributes of contestability, and policy developments, covering both policy studies in the economic literature and some remarkable decisions in the regulatory and antitrust arenas, that rely explicitly on the theory of contestability.

However, before reaching these central areas in order to deal with any remaining misunderstandings we shall begin with a brief restatement of some of the things our work does not and was not intended to imply and the things we instead hope it accomplishes. Specifically, we will deny emphatically that it offers carte blanche to mindless deregulation and dismantling of antitrust safeguards. On the contrary, so far as policy is concerned, contestability theory provides guidance in ascertaining where intervention is warranted socially, and it provides a more widely applicable benchmark to guide regulatory agencies and the courts in those arenas where intervention is called for by considerations of economic welfare.

Contestability and "libertarian" ideology

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. In a market that approximates perfect contestability, it is true,

* Princeton and New York Universities.

† Princeton University. The authors want to thank the Information Science and Technology Division and the Regulation and Policy Analysis Program of the National Science Foundation and the C. V. Starr Center for Applied Economics for their generous support of this research. We also wish to acknowledge with gratitude the able assistance of Daniel Schult and the useful comments of Marius Schwartz.

we believe matters can be left to take care of themselves. Small numbers of large firms, vertical and even horizontal mergers and other arrangements which have traditionally been objects of suspicion of monopolistic taint and worse, are rendered harmless and perhaps even beneficent by the presence of contestability. But that observation is no whitewash and establishes no presumption, one way or the other, about the desirability of public sector intervention in any particular market of reality. For before anyone can legitimately use the analysis to infer that virtue reigns in some economic sector and that interference is therefore unwarranted, that person must first provide evidence that the arena in question is, in fact, highly contestable. The economy of reality is composed of sectors which vary widely in the degree to which they approximate the attributes of contestability. Thus, the conclusion that *perfectly* contestable markets require no intervention claims little more than the possibility (which remains to be proven, case by case) that *some* markets in reality may automatically perform in a very acceptable manner despite the small number of firms that inhabit them.

Thus, it is simply incorrect to associate our writings on contestability with an all-pervasive laissez-faire position on the role of regulation and antitrust (for such a characterization of our position see, e.g., Shepherd (1984)). We disagree vehemently with such a view of the world. On the other hand, we reject with equal conviction the position of those who hold that mere large size of a firm means that it *must* serve the economy badly, that high concentration ratios are sufficient to justify governmental restrictions upon the structure or conduct of an industry or that a horizontal merger merits automatic condemnation if it increases concentration ratios substantially. For it is true that a high degree of contestability, where and if it does happen to hold sway, should remove the undesirable consequences which might otherwise stem from these phenomena. Thus, contestability theory supports neither extreme interventionists nor extreme noninterventionists. We believe that antitrust and regulation have valuable roles to play, and that contestability theory can help to identify and sharpen those roles and thereby benefit the public.

Objectives of contestability research

Contestability theory offers an analytic framework within which the fundamental features of demands and production technology determine the shape of industry structure and many of the characteristics of industry prices. The theory accomplishes this via a process of simplification; by stripping away through its assumptions all barriers to entry and exit, and the strategic behavior that goes along with them both in theory and in reality.

The model of perfect competition has also served an analogous simplifying role, and in so doing has provided the foundation for the most elegant and well-articulated portions of economic theory. However, because the model of perfect competition prespecifies industry structure by its very

construction, it cannot serve as a useful benchmark for the study of the determinants of industrial structure. Perfect competition is a special case of perfect contestability, and perfect contestability applies with equal force to circumstances where perfect competition is impossible because economies of scale are present. Because of this fact, and because contestability theory encompasses an endogenous determination mechanism from which any industry structure may emerge (depending on circumstances), we feel that it is an extension of the competitive model appropriate for use in the theory of industrial organization.

In addition to this broad purpose in theoretical analysis, we feel that research on contestable markets has two objectives related to policy. The first is the establishment of an improved set of guidelines for appropriate government intervention in the structure and conduct of firms and industries, that is, of the rules to be followed by the regulators and antitrust authorities in those cases in which their intervention is called for. The second objective for policy analysis is the determination of criteria distinguishing between those cases in which intervention by the public sector is warranted and those in which it is not.

With regard to the first policy objective, it is widely understood that the basic guideline for intervention should be replication of the consequences of effective market forces in those cases in which competition, actual or potential, is insufficient to do the job. That is, intervention should aim to induce or compel an industry in which monopoly power is present to perform as it would if effective competitive pressures were available. But, for such cases perfect competition is often unsuitable as a standard, particularly in circumstances common to regulatory and antitrust issues where scale economies and related attributes dictate the presence of only a small number of firms, at least some of them relatively large.

Here, as we will see from recent developments in regulatory practice that will be reviewed presently, contestability theory was, indeed, able to suggest answers where other bodies of theory did not. Moreover, it was able to eliminate curious gaps in the standard theory of policy. Perhaps the most noteworthy of these was the lack of a defensible criterion for regulatory ceilings on prices, though one would have thought that to be the first order of business for social control of firms with monopoly power. Yet, to our knowledge, neither regulators nor courts previously had available to them from the theoretical literature any defensible test to determine just when a particular price set by one of the multiproduct firms of reality was excessive. We will also see later how contestability theory determines how high is enough and will observe that regulatory agencies have begun to adopt the criterion that emerges, with explicit acknowledgement of its source.

In saying this, we do not mean to claim that the theory sheds light on the welfare economics of markets that are highly uncontestable. The theory does not even try to contribute to the analysis of behavior in such cases or to the evaluation of the social cost of any resulting departures from

optimality. All the analysis pretends to say about such cases is that if they entail behavior which is deemed socially undesirable and if government undertakes to change this behavior, then it is appropriate to require the individuals to act as they would have if the threat of potential competition had constrained them effectively. If, for example, firms charge prices and earn profits that are deemed excessive, it is appropriate to offer consumers the protection that the threat of barrier free entry would have provided them.

Finally, we feel that contestability theory has helped to show that in *some* cases where regulatory or antitrust intervention was previously common, it was in fact undesirable. Contestability theory, we think, does help to clarify which arenas are the proper candidates for deregulation and cessation of other forms of intervention. But the theory neither calls for anything like universal deregulation nor for its automatic extension without careful study of the pertinent facts, case by case.

Theoretical developments

There has been an outpouring of theoretical work on contestable markets and allied subjects, as there has been recently throughout the field of industrial organization (as this very volume indicates). Our brief overview of the recent developments cannot do justice to the full range of valuable contributions and is, unfortunately, bound to miss important items. This review is organized into four topics: (i) conditions for sustainability of prices set by a monopoly incumbent; (ii) theoretical applications of contestability; (iii) the characteristics of production cost that determine market structure in contestable markets; and (iv) the relationship between strategic firm behavior and contestability.

In our book (1982), we identified one set of conditions sufficient to assure the existence of sustainable prices for a multi-output natural monopolist, and those conditions also were shown to guarantee that the Ramsey optimal prices are sustainable. Among those conditions are the stipulations that the cost function exhibit overall economies of scale and the property of trans-ray convexity. We showed that the sustainability of Ramsey prices could tolerate some deviation from these properties, such as product-specific fixed costs of bounded magnitude. Faulhaber and Levinson (1981) developed the idea of anonymously equitable prices which offered no cross-subsidies to any possible groups of consumers, and connected the properties of such prices to sustainability. Sharkey ((1981) and (1982)) and Spulber (1984) presented new sets of conditions sufficient for the sustainability of monopoly prices, and showed that the Ramsey prices need not be among them. ten Raa ((1983) and (1984)) sharpened the connections among subsidy-free, market-clearing, and sustainable prices. Mirman, Tauman and Zang (1985) further clarified these connections, established novel sets of sufficient conditions for the existence of sustainable prices, and

identified conditions under which the Aumann-Shaply prices would be sustainable.

Brown and Heal provided the valuable result that where Ramsey prices are unsustainable, the government can solve the consequent policy conflict between open entry and optimal prices via a system of excise taxes and subsidies to incumbent and any entrants alike. Working in a general equilibrium context with independent demands for the outputs of the natural monopolist, they proved the existence of such a tax system, with zero net effect on government revenues, that permits Ramsey optimal consumers' prices and sustainable producers' prices.

Panzar and Postlewaite (1985) demonstrated the sustainability of Ramsey optimal non-linear prices where costs are linear, and provided counterexamples for other circumstances. In contrast to the usual results for contestable markets, Perry (1982) showed that a natural monopolist can at once earn positive profits and deter entry by commiting itself to the sale of some units of homogeneous output at prices lower than those that apply to marginal purchases.

Our second category of new theoretical work comprises applications of contestability to various issues. Quirmbach (1982) studied vertical integration by an upstream monopolist into a competitive downstream industry. While the welfare effects are, in general, ambiguous, Quirmbach showed that profitable integration would raise social welfare if the upstream market were perfectly contestable.

In their pathbreaking book, Helpman and Krugman (1985) provide a sketch of a new theory of international trade in which industries in the trading nations are perfectly contestable rather than perfectly competitive, as in conventional trade theory. This construction permits them to study the effects of increasing returns to scale in the production processes internal to firms, without the need to specify particular forms of oligopolistic and strategic behavior. Among their conclusions are the result that the sustainable factor-price-equalization equilibrium is a highly useful construct for trade analysis; that even in the presence of economies of scale comparative advantage shapes trade patterns with unequal factor rewards; and that there is a strong presumption that there will be gains from trade when national and international markets are contestable.

Recent work on properties of production costs has been stimulated by the significance they have for the character of industry structure in contestable markets. Weitzman (1983) showed that constant returns to scale must hold when demand can be met by the instantaneous accumulation of inventories created by production that is turned on and off repeatedly without additional costs. He concluded that the hit and run entry opportunities said to be a necessary part of contestability must then be inconsistent with increasing returns to scale. In reply, we pointed out that quantities of services cannot be accumulated in inventories, no matter how short their duration, and that contestability does not require the ability to start and stop

production costlessly—only the ability to sell without vulnerability to incumbents' responses for a time long enough to render all production costs economically reversible.

Teece (1982) has contributed illuminating work on the sources of economies of scope, both in theory and in specific applications. Gorman (1985) has recently provided theoretical results on economies of scope where there are complementarities or anticomplementarities in marginal costs and also product specific fixed costs that may be subadditive or superadditive. Bittlingmayer (1985) has provided a theoretical analysis that builds multiservice cost functions for airlines upon the foundation of the costs of operating aircraft. He finds economies of scope among routes arising from the hub and spoke architecture of the cost efficient configuration of routes. He goes on to study the characteristics of rate structures that are sustainable and that are Ramsey optimal.

The final category of theoretical work, the relationships between contestability and strategic firm behavior, is most difficult to review succinctly because it is intertwined with the bulk of new research in industrial organization concerning strategy and oligopoly solution concepts. Perfect contestability is a theoretical benchmark that is by its very construction immune from considerations of strategic behavior by dint of its assumption of the absence of economically sunk costs and irreversible commitments necessary for entry. Nevertheless, as pointed out by our distinguished reviewers, Brock (1983) and Spence (1983), and by our constructive critic Schwartz (1985), it is worthwhile to investigate what forms of game and what models of dynamic strategic oligopoly yield outcomes consistent or inconsistent with contestability. Here, Knieps and Vogelsang (1982) and Brock and Scheinkman (1983) have contributed models of quantity sustainability in which outcomes are quite different from those in contestable markets because potential entrants take incumbents' quantities rather than prices as given. In view of the enlightening work of Kreps and Scheinkman (1983), it is now clear that quantity-taking behavior is a reflection of precommitments to quantity-determining capacity by firms that later simultaneously announce prices.

In contrast, contestable outcomes are yielded by Nash equilibrium in a price setting game without precommitments (Mirman, Tauman, and Zang (1983)), and in a game in which supply schedules comprise the strategy sets (Grossman (1981)). We showed that contestable outcomes are a necessary feature of equilibrium in the limit as sunk costs approach zero, using a model in which incumbents' prices are sticky for at least a short period of time (Baumol, Panzar, and Willig (1983)). Maskin and Tirole (1982) analyze a model of dynamic oligopoly involving firms with fixed costs and show that the solution approaches the contestable outcome as strategic advantages fall to zero. In contrast to these results, Dasgupta and Stiglitz (1985) provide models in which incumbents' strategic advantages over a single potential entrant, no matter how small those advantages may be,

permit the incumbents to earn economic profits without danger of entry. And Applebaum and Lim (1985) have contributed a model in which the degree of contestability is endogenous, determined by the incumbent's incentives to commit to quantity-determining capacity in view of demand uncertainty and the pattern of costs over time.

In short, there is a flood of exciting research activity in the area of strategic firm behavior, and one of the ways in which the various models are differentiated is the relationship between the games' solutions and the outcomes that would emerge in perfectly contestable markets. At present, that relationship seems to be highly sensitive to the fine grained structure of the models' game forms, so much so that one suspects that empirical reality embodies relationships more robust and stable than does oligopoly theory in its current tumultuous state.

Experimental market studies of contestability

The fruitful and growing body of research using experimental simulation of market behavior has produced a number of papers investigating contestability issues. This work uses human subjects (usually students), real money payoffs and experimental rules corresponding to preselected cost and other pertinent functions that permit direct calculation of the results predicted by theory and their comparison with those that are observed from the behavior of the experimental subjects.

Taking off from earlier studies on monopoly behavior, Coursey, Isaac and Smith (1984) undertook the first of the experimental studies of contestability. They sought to determine whether complete freedom of entry and exit in a market which could most cheaply be served by one firm but where at least two (potential) participants were available would yield prices closer to the competitive than the monopoly level (they call this the "weak contestable markets hypothesis") or to prices actually equal to the competitive level (the "strong contestable market hypothesis"). Participants were given money which they were required to spend in accord with a strictly decreasing and known marginal cost function, the amount spent depending on how much of the good they decided to "produce". The consumers' demand function was also given but was not known by the sellers (except by experience derived from repeated play) who kept for themselves any excess of sales revenues over costs. A participant made a sale if that person's price offer was not undercut by a rival. Both sellers were required to post their price and quantity offers privately, *and* at the same time. Any seller was permitted (but not required) to sell less than the amount the market demanded at the posted price. The authors concluded that "the . . . experiments strongly support the contestable markets hypothesis, namely, that to observe approximately competitive behavior by a single producing firm with substantially decreasing costs, it is sufficient that (a) sunk costs are zero and (b) there are two contesting firms acting non-

cooperatively in the sense that there is no explicit nonprice communication between them that leads to excessive restriction of supply" (Coursey, Isaac, Luke and Smith (1984) p. 69).

Harrison (1985) reaches a rather weaker conclusion from these results and those that emerge from a replication of the experiments (Harrison and McKee (1985) (1987)). Using as test criterion an "index of monopoly effectiveness" $M = (\pi - \pi c)/(\pi m - \pi c)$, where π, πc and πm are, respectively, the total profit that emerges from an experiment, the competitive, and the monopoly profit levels. They note that four of the six experiments did attain the $M = 0$ value expected of perfect contestability after 10 periods (i.e., 10 replications of the experiment by a given set of players). However, over all replications M averaged 29 percent in the original set of experiments and 19 percent in the second set, which while not very far from a competitive result, " . . . is significantly positive." Thus, they conclude, "clearly, we can reject this strong form of the CMH (contestable markets hypothesis)" (p. 13) (the pages refer to the unpublished manuscript).

The Harrison McKee study also concludes from experiments using essentially the same procedure that " . . . the discipline of contestability does serve to mitigate the monopoly (significantly) when we compare it to an unregulated monopoly . . . (but) enlightened monopoly regulation vastly outperforms contestability in terms of reducing monopoly effectiveness" (Harrison, pp. 13–14). (Here the regulator's role is simulated by the offer of a subsidy payment equal to the consumers' surplus from any output exceeding the monopoly level, as calculated from the demand curve.)

In another experiment, Coursey, Isaac, Luke and Smith (1984) evaluate the effects of sunk costs imposed equally on "incumbents" and "entrants". The sunk cost takes the form of a $2 entry permit, valid for five periods. The incumbent is one of two sellers chosen randomly and permitted to operate as a monopolist for five periods. Thereafter, the other seller is permitted to enter and post prices that compete with the incumbent, who is required to purchase an entry permit in period 5 in order to continue operation in periods 6–10. The results, according to Harrison's reanalysis, exhibit no limit pricing behavior and indicate " . . . that the threat of *future* entry in period six imposed virtually no discipline on the incumbent in periods 1–5" (p. 15). He also notes that the sunk costs constitute a differential incremental risk to the entrant (and hence an entry barrier) only in period 6–10. Accordingly, he considers it noteworthy that "For all eighteen periods there is a slight and statistically insignificant effect of sunk costs reducing efficiency. However, there is a significant decline in efficiency in periods 6–10 due to sunk costs" (pp. 19–20).

Finally, Harrison tested the effects of imposition in the Bertrand–Nash (or, perhaps more accurately, the Stackelberg leader–follower) assumption that entrants take the incumbents' prices to be given—a premise widely associated with contestability. This was done by letting all participants post prices simultaneously in the first period, designating the winner as the next

periods' incumbent, who is then *required to post a price before anyone else does, and to announce it publicly*, with the following periods handled in the same way. Both two and three participant experiments were run. The author concludes, "... satisfaction of the Bertrand–Nash assumption is associated with a dramatic decline in (M)... Moreover, we find support ... for a *strong form* of the (contestable market hypothesis) that claims that observed prices will converge to *and attain* competitive predictions" p. 37, italics in original).

These results would appear to offer strong support for the conclusions of contestability theory but, unfortunately, they still leave a major gap with which future research will have to deal. As Marius Schwartz, in his latest paper criticizing contestability theory, rightly observes "Harrison's experiments basically do constitute a fair test of behavior under perfect contestability: the incumbent faces a price response lag and the entrant can hit and run costlessly. The real question, however, is not whether competitive results will emerge under perfectly contestable conditions but how often such conditions are likely to exist and what happens when they do not" (Schwartz (1985) p. 20, footnote omitted). Put another way, the critical issue that remains is the determination of the circumstances under which the Bertrand–Nash assumption holds or at least is assumed by the participants to hold approximately. It is noteworthy to have it confirmed that *imposition* of Bertrand–Nash behavior leads to the results predicted by the theory. However, that result still leaves work for future statistical or experimental investigation, which Harrison and his colleagues are pursuing.

Econometric studies: multiproduct cost functions and industry structure

In the past three years there has been an outpouring of empirical studies using the cost concepts derived from the contestable markets literature (see the survey of earlier work by Bailey and Friedlaender (1982)). They provide estimates of the shapes of ray averages cost curves, of economies of scope, of the M-locus (the locus of points of minimum ray average costs on each ray in product-quantity space). Many such studies use the functional forms for the multiproduct cost function that were recommended in our book. The empirical studies encompass a wide range of industries, including banking, rail, truck and water transportation, insurance, hospitals and a number of others. Here we can only recapitulate briefly a sample of these studies which, predictably, show that industries differ markedly in their scale and scope attributes and in their degree of approximation to contestability.

I. Cost function estimates

Perhaps the most consistent message of the "new" multiproduct cost studies is that the results of previous work using Cobb–Douglas or CES functional forms cannot be trusted. The more general translog form used by most of the multiproduct literature contains, as special cases, the Cobb–

Douglas and CES forms. The appropriateness of these forms can be tested within the translog framework, and in all such tests reported, those forms are rejected in favor of the general translog form.[1]

A. *Banking and related activities*

Banking institutions seem a natural subject for multiproduct analysis. Gilligan, Smirlock and Marshall (1984) and Gilligan and Smirlock (1984) divide output into deposits and loans, using data on many banks provided by the Federal Reserve System. They find that there are large economies of scope "across the balance sheet." They suggest banking theory explicitly consider that fact, and that bank regulation intended to affect one output also be analyzed in multiproduct terms. They encounter decreasing ray average costs (RAC) for small banks, concluding that some small single-bank markets may indeed be subadditive. Overall scale economies are estimated to dissipate as bank size increases, and banks with deposits of $100 million or more manifest increasing RAC. Product specific diseconomies of scale also appear and the authors conclude that policy that precludes entry and encourages mergers cannot be justified by any cost savings.

Murray and White (1983) examine the investment side of British Columbian credit unions. They find significant economies of scope between mortgages and consumer lending and no economies of scope between all lending and investments. Unlike the U.S. studies, returns to scale are observed in most credit unions in the study. The authors conclude that regulation to "inhibit growth and diversification" hampers market efficiency.

Using the same output categories (mortgages, other lending, and investments) for Californian Savings and Loan data, Mester (1985) finds both overall and product specific constant returns to scale, suggesting that large assets offer no cost advantage. She also finds no advantage to a network of branches, given asset size, and absence of economies of scope between outputs. She rejects cost convexity, although finding pairwise transray convexity, and ray subadditivity along some rays.

Kellner and Mathewson (1983) provide a multiproduct analysis of the life insurance industry. They use a four-output industry model. Scale effects are rejected for the average firm while economies of scope appear in some

[1] Note that the translog form is itself unreasonable for cases involving zero output of one product since then the translog form implies that the total cost of the remaining products must be zero, no matter what their output levels. One method used to avoid this problem is recourse to a Box–Cox transformation of the outputs. (See Caves, Christensen, and Tretheway (1980) for a discussion of the translog form.) As an alternative, Friedlaender, Winston, and Kung Wang (1983) use a quadratic approximation to a hedonic function to deal with the problem. The quadratic form also offers a simple measure of economies of scale and scope as well as marginal cost.

product lines. In contrast to previous work, they find no evidence of "natural monopoly tendencies" in the cost structure.

B. *Transportation*

Bailey and Friedlaender (1982) discuss work done on the trucking and railroad industries. Network configurations are reported to offer large economies of scope in both industries. These bias upward the measures of scale effects in single product analyses. In the case of trucking, apparent scale economies disappear for large firms when multiproduct analysis is used. Thus, single product analysis of scale effects may improperly reflect scope effects.

Stimulated by the examination of the regulation of bus transport by the Interstate Commerce Commission (ICC), Tauchen, Fravel, and Gilbert (1983) examine the cost structure, taking regular travel, charter, local service and school busing as the outputs of the industry. Using data on 950 privately owned intercity bus firms in the United States, they find that economies of scale are exhausted at low output levels while economies of scope occur throughout. They point out that price and entry regulation is usually meant to deal with economies of scale which are minor here. Economies of scope can be enhanced by fostering of service cooperation among firms in terms of schedules, interlining of tickets and baggage and use of common terminals.

Friedlaender, Winston, and Kung Wang (1983) examine the American auto industry's cost surface for the individual firm. Output is divided into small cars, large cars and trucks. They find that the size of the firm has little to do with the returns to scope or to scale and that the cost function is not convex; some regions exhibit economies of scale and scope and others exhibit diseconomies. This demonstration of the variability of the industry's cost structures indicates the danger of using the neighborhood of a particular output vector as a basis for the empirical analysis of costs. Generalizations based on studies of aggregate output data are, thus, unreliable.

Caves, Christensen, *et al.* used the multiproduct framework to analyze cost attributes other than economies of scope and scale. In Caves, Christensen and Herriges (1984) they construct models of consumer reactions to peak load electricity rate structures. They also examine productivity growth in U.S. railroads employing multiproduct estimates (Caves, Christensen and Swanson (1981)). And they analyze the effects of deregulation on airline productivity—distinguishing between local service and trunk lines (Caves, Christensen and Tretheway (1983)).

C. *Telecommunications*

Much work has been done on the telecommunications industry. Its output is readily divided at least into local and long distance service, and so

constitutes a classic multiproduct case. Evans and Heckman (1984) perform an empirical test of the subadditivity of the industry's costs. To avoid the need for global cost information Evans and Heckman derive a local test for subadditivity by restricting the measurement of subadditivity to a subregion of the observed data. For each year of data they calculate the cost savings offered by a single firm relative to the minimum two-firm cost. Finding all such savings to be positive, they conclude that cost was subadditive during the period 1958–1977.

On the other side, Charnes, Cooper and Sueyoshi (1985) use a goal programing/constrained regression method (derived in operations research studies) and a multiproduct framework, data set and functional form identical with those of Evans and Heckman to obtain opposite results. Basically, they drop the standard economic assumption that the firm always operates on the efficient frontier and seek to derive an envelope from observed costs. While Evans and Heckman found the savings were positive, Charnes, Cooper and Sueyoshi found the maximal savings to be negative. They question the validity of standard econometric procedures, but accept the multiproduct nature of the issue and the test of subadditivity. They stress the importance of use of methods derived from several disciplines, particularly when large issues of policy depend on the outcome.

D. *Other industries*

Cowing and Holtmann (1983) divide hospital care into five care types to analyze the efficiency of general-care hospitals in New York state. They reject the use of a single aggregated measure of output and find the economies of scope "may be at least as important as scale effects in designing more efficient hospital service." They urge policy makers to take scope effects into account when deciding on mergers and encourage future health care research to employ multioutput specifications.

Scott (1982) examines the effects of concentration on manufacturing firms. He investigates the effects of multiproduct competition, or "contact", between firms upon excess profits and resource mobility. He constructs a measure of multimarket contact and then shows that an increase in contact in markets with high sales concentration permits larger profits. However, when sales concentration is low, multiproduct contact can drive profits toward costs.

Vertical integration theory suggests that the presence of economies of scope in an industry need have no categorical implications for the scope of individual firms. Teece (1980) examines this issue and concludes that so long as the economies of scope derive from proprietary information or a specialized indivisible physical asset, a multiproduct enterprise is most efficient. He takes the energy industry as an example of this, where petroleum search and removal techniques can be applied to coal.

Mayo (1984) examines data to pursue Teece's ideas. He uses a cost

relationship with coal and petroleum as the outputs. Economies of scope, presumably attributable to similarity of extraction techniques, are found. He also shows that ray average costs decline and, near the axes, declining incremental costs and transray convexity yield evidence of output specific subadditivity. This suggests that divestiture of large laterally integrated petroleum firms may increase costs.

Water supplies are generally run as monopolies in the United States, and Kim and Clark (1983) analyze this industry's cost structure with supplies to residential and nonresidential customers as outputs. Using gallons of water per day as the metric, they find overall constancy in returns to scale. But nonresidential output offers product specific economies while diseconomies appear in residential supply. By analyzing ray average costs for many output mixes, they plot the M-locus for the industry and find it to be concave to the origin. This yields a wide range for the optimal number of firms in a market. In addition, by looking at an "average" ray (79 percent residential), they find that most firms produce near the essentially flat portion of the average cost curve. They conclude that market concentration is generally unnecessary for scale economies and may hinder potential competition in water supply.

II. Tests of the contestability of industries

There have been a number of industry studies seeking to test whether the area of the economy under investigation can legitmately be classed as "contestable". Several, but not all, have yielded results that are somewhat surprising. For example, as we will see below, some studies of the airline industry conclude that the industry is less close to the model of perfect contestability than has sometimes been suggested. On the other hand, one very recent study indicates that the aluminum industry is closer to contestability than might have been expected. Froeb and Geweke (1984) describe the U.S. aluminum industry as an area in which the long term threat of entry drives profits downward. Allowing for short run effects of entry barriers, they test the long run behavior of firms from 1949 to 1972 and find it consistent with that one would expect in an imperfectly contestable market.

On the other hand, the results of some of the studies we are now discussing are less surprising. Extensive work by Davies (1986) for the Canadian Transport Commission on the liner shipping industry concludes that contestability theory may offer the most appropriate model to describe that industry since the transferability and resaleability of capital characterize oceanic shipping. The theory "accounts for the condition of entry to the industry, the cost characteristics and structure of the industry, the existence of loyalty ties (to cartels), the reported absence of supernormal profit and the pricing structure operative in the industry. No other body of theory can explain the existence of all these different phenomena in such a coherent

and elegant manner." These conclusions are supported in an independent study of ocean transport by Peter Cassidy (1982).

Contestability and public policy

As has already been said, contestability theory aspires to offer two types of guidance to policy makers: first, it undertakes to provide criteria to distinguish the cases in which government intervention is desirable from those in which it is not and, second, it seeks to offer tools to the regulator that will increase the public welfare benefits of this intervention. There have recently been significant applications of contestability theory on each of these fronts. On both issues we will examine some of the recent discussions in the academic literature and devote the bulk of our attention to pertinent developments in recent regulatory and antitrust decisions.

A. *Contestability as a guide to arenas inappropriate for intervention*

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. Indeed, their costs and the inefficiencies they cause may then offer little or no offsetting benefit.

Since this viewpoint thoroughly antedates contestability theory it is not surprising that it has appeared in a variety of official policies. For example, the 1982 Merger Guidelines of the U.S. Department of Justice defines market power as the ability to profit by raising (and maintaining) price above the competitive level, not referring to market share in that definition. It includes in its market definition firms whose entry would be attracted should an elevation of price be attempted. It asserts, moreover, that mergers will go unchallenged by the Department if they affect only markets subject to potential competition that is sufficiently strong. It proposes to take concentration into account only where potential competition is inadequate.

Similarly, the Federal Communications Commission has shown its awareness of the role of potential competiton. In a recent Notice of Inquiry (FCC (1985)) it adopts a market definition for the analysis of long distance telecommunications similar to that of the Department of Justice. It asserts that " . . . satellite carriers readily can *shift* their capacity from one area to another and terrestrial carriers can readily *expand* their service areas through new facilities, interconnection, or resale." "There may be numerous large *potential* entrants in common carrier or private systems that further check AT & T's market power to some degree." And, "The absence of entry does not show that an existing firm possess market power; existing

firms may be charging competitive prices because of competition among themselves or the *threat* of *potential* entry, and thereby make entry unattractive." Thus, the FCC implies that a large market share need not confer any monopoly power.

Contestability theory makes direct appearances in a recent decision of the Federal Trade Commission (FTC (1985)).[2] The case involved the acquisition in 1982 by the Echlin Manufacturing Corporation of the automotive aftermarket divisions of the Borg Warner Corporation, including its carburetor kit activities which overlapped with some activities of Echlin. The FTC staff challenged the acquisition under Section 7 of the Clayton Act and Section 5 of the Federal Trade Commission Act on the grounds that the merger would add significantly to the concentration of this market.

However, testimony before the Administrative Law Judge indicated that the field was not beset by any significant entry barriers and that potential entry was sufficient to deprive the merged firm of any monopoly power over the assembly and sale of carburetor kits. The Commission agreed, and permitted the acquisition. In its decision the Commission asserted (explicitly citing contestability theory and other writings on potential competition):

> "An attempt to exercise market power in an industry without entry barriers would cause new competitors to enter the market. This additional supply would drive prices back to the competitive level. Indeed, the threat of new entry can be as potent a procompetitive force as its realization. As the Supreme Court has recognized, the presence of potential entrants on the fringe of a market can prevent the exercise of market power by the incumbent firms even if the potential entrants never actually enter the market. Thus, in the absence of barriers to entry, incumbent firms cannot exercise market power, regardless of the concentration in the nominal 'market,' and indeed even if that market has been 'monopolized' by a single firm" (pp. 9–10).

The decision goes on to espouse the same definition of entry barriers as that adopted (from Stigler and others) in the contestability literature "... as additional long-run costs that must be incurred by an entrant relative to the long-run costs faced by incumbent firms" (p. 12) and, again citing the contestability literature, it concludes that "if sunk costs are considered an entry barrier, it must be because they create a difference in the risk confronting the incumbent firms who have already committed their resources and potential entrants who have yet to make that decision" (p. 17). The decision finds, along with contestability analysis, that "... we cannot agree that economies of scale and declining markets necessarily create barriers to entry" (p. 18), and that "the absence of past entry, however, does not prove the existence of entry barriers because it is equally consistent with alternative explanations, such as a declining industry or competitive prices" (p. 19).

[2] It should be made clear that the authors of this paper cannot claim to be disinterested reporters, having testified in several of the cases described here.

B. *Deregulatory experience: airlines, buses and trucks*

The intellectual foundations of the deregulation of airlines, trucking, and buses included recognition of the power of potential entry. While it was clear that in many transportation markets efficiency is not inconsistent with the operation of enough active carriers to make the replacement of regulation by competition appropriate, it was also clear that on many routes efficient operation is incompatible with the presence of several carriers. Nevertheless, deregulation proceeded in the expectation that potential competition could adequately protect consumers of transportation services in such arenas.

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 position on this score. We now believe that transportation by trucks, barges and even buses may be more highly contestable than passenger air tranportation. Barges and trucks have business firms rather than individual consumers as their primary customers, and that facilitates the provision of service via contracts on which potential entrants can effectively bid against incumbents. Where the contracts apply to long run relationships, as they often do in transportation, even capital costs that are physically irreversible are not economically sunk in the pertinent time period. Moreover, 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 such as that which prevented People Express from acquiring even a single gate of its own at Denver's Stapleton International Airport, so that it was forced instead to lease gates from other carriers, catch as catch can, during a year of flying to that airport.

In fact, post-deregulation experience in the airlines industry has revealed several elements of the structure of supply that conflict significantly with the conditions necessary for the pure theory of contestability to apply without modification. While these structural elements may be transitory, they nevertheless appear to have influenced the performance of the industry in important ways since the advent of deregulation.

First, as the previously noted difficulties of People Express exemplify, there have been constraining shortages of facilities and services of air traffic control at several pivotal airports. These constrain flights in and out of the affected airports and, in addition, they also restrict the prospects of entry and expansion on routes that would otherwise interconnect efficiently with such flights. Second, technological advances, changes in the 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 the aircraft demanded, with delivery lags frequently stretching

to three years. Third, newly certificated 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. Recently, the older carriers and their unionized work forces have been adapting themselves to this new competitive reality with the aid of more flexible wage contracts, dual wage structures, and less costly contract settlements. Nevertheless, substantial differences in labor costs between older and newer carriers persist.

These conditions make it easy to see why the airline industry does not conform perfectly to the contestability model, even if aircraft are "capital on wings." Thus, it remains to analyze both quantitatively and qualitatively the degree to which the performance of the industry does, or can be expected to, reflect the predictions of contestability theory. This should come as no surprise since most industries can be expected to depart in some important respects from the model of perfect contestability, and it will therefore generally be necessary in applying the theory to assess the economic significance of the deviations.

Several econometric studies have confirmed the imperfection of the contestability of the airline markets (see, e.g., Call and Keeler (1984)). They have shown, for example, that 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, as perfect contestability would require. Moreover, when new entry does occur, established carriers do reduce their fares in response, something one would expect in a conventional oligopolistic market other than one that is perfectly contestable. Even the study of Morrison and Whinston, which concludes that potential entry does constrain price significantly, finds that the coefficient describing the influence of potential entrants does not become significant until the number of such prospective entrants exceeds three.

The econometric study by Graham, Kaplan, and Sibley (1983) permits comparison among the effects on 1981 route-by-route prices of various influences that lie outside the predictions of contestability theory. They find that prices deviate from costs an additional 10 percent on the average and other things being equal, for routes that utilize one of the major slot-constrained airports in New York, Chicago, or Washington, D.C. In contrast, prices are reduced some 22 percent relative to costs on routes where a newly certificated carrier operates. And, prices are some 18 percent lower relative to costs if four carriers fly the route rather than only one.

These results indicate the significance of the physical constraints on entry and expansion in two ways. First, the direct influence of the slot constraints on prices shows up in the data. Second, the substantial effect of the presence of new carriers, with their low labor costs, demonstrates how much lower fares would generally be if those carriers were able to expand their operations to cover more routes. Finally, the significant effect of the number

of carriers on a route shows that active competition plays a role in holding price down toward cost. However, since the size of that effect is so much smaller than would be predicted by most theories of oligopoly that focus on the role of active competitors, it can be concluded that the forces of potential competition still play an important role, despite the structural conditions that impede their workings.

Similar conclusions seem to us to follow from qualitative evaluation of the behavior of the airline markets under deregulation. As we have just seen, the responsive price cuts of incumbents when faced with incursions of low priced entrants is certainly not compatible with the predicted qualitative behavior of contestable markets in the long run. However, there are other qualitative properties that are highly pertinent. In a recent study, Elizabeth Bailey (1986) sets out to examine these. She undertakes to analyze

> "... the consequences that have emerged under deregulation in terms of the following behavioral properties predicted by contestability theory: a variety of products will emerge, each of which will yield zero economic profit; the revenues from any subset of the products must exceed the incremental costs of those products, so that no cross-subsidy can exist; prices for each product will equal or exceed marginal costs; and an equilibrium market structure will minimize costs of the industry. Thus, if the theory has some degree of validity, more of these properties should be displayed after deregulation than before. Cross-subsidy, which was pervasive in the regulatory era, should be significantly eroded. Prices should move nearer to costs. Products capable of producing zero economic profits, but which were formerly excluded, should now appear." (pp. 2–3, footnote omitted)

Here evidence leads her to conclude that, while the performance in each respect has not been absolutely clear-cut, on balance the patterns predicted by contestability theory have indeed emerged. She writes, in summary of the consequences of deregulation of the airlines, that

> "Prices in the cheaper-to-serve long-haul and dense markets were substantially lowered (by about forty percent), whereas prices in the more expensive short-haul and thin markets went up somewhat. A diversity of price-service options arose. Individuals could select between low-service/low-price discount carriers and full-service national carriers. Even among the full-service carriers, prices were lower for customers willing to improve load factors by traveling in off-peak periods or by taking one-stop rather than non-stop flights.
>
> The encouragement to entry under air transport deregulation brought with it a variety of contributions to efficiency. One involved delivery systems. It quickly became clear that hub-and-spoke operations rather than the mostly linear systems imposed by regulation offered savings to the airlines. Hub-and-spoke systems also substantially improved service for consumers. A second efficiency contribution involved input productivity. The post-deregulation period has been characterized by pressures to reduce pay scales toward those in unregulated economic sectors, to increase productivity through changes in work rules, and to choose a more efficient fleet configuration" (p. 22).

In light of these observations and similar evidence drawn from the opening of brokerage commissions to market determination, the deregulation of trucking, and the opening of telecommunications markets to entry, Bailey concludes that performance in these areas has shown no significant inconsistencies with the predictions of contestability analysis.

In short, in terms of the airlines case we can infer that market forces through the pressures of competition, both actual and potential, have done a commendable if imperfect job in protecting consumer interests. This is suggested by a number of developments following deregulation: the decrease in real average prices; the reduction in average time spent in traveling from point to point; the falling real costs of airline operation; the erosion or disappearance of cross subsidies, with the elimination of financing of sparsely traveled routes by those that are heavily traveled, and of peak travelers by off-peak flyers, and the considerable improvements in efficiency through computerized routing and hub-and-spoke operations—approaches that had evaded both executives and the Commission during the era of regulation.

C. *Contestability theory as guide for regulation*

We come, finally, to the arena in which the viewpoint of contestability may make its main contribution—as a guide for regulation, rather than as an argument for its elimination. How does contestability theory help in this domain? After all, perfect competition has long served usefully as the ideal for government intervention to follow; that is, as the model of performance to which it should seek to make the regulated firm adhere. What does contestability analysis have to add to this? The answer is that in some circumstances, notably in the presence of substantial economies of scale and scope, the standard of perfect competition is totally inappropriate. For example, where economies of scale and scope are present society no longer is sure to benefit if firms are required to be small, as they would be under perfect competition. Similarly, a rule that price must be set equal to marginal cost is a prescription for financial disaster. Of course, Ramsey pricing theory is of considerable help here, but in some circumstances its usefulness in practice is limited, particularly where there are no reliable data on elasticities and cross elasticities of demands for the considerable number of products at issue. It is then that contestability theory can come to the rescue (and several times already has). For it can propose, for example, to offer consumers in markets with unavoidable entry barriers just the same sort of protection from excessive pricing that they would have derived from perfect freedom of entry, if such freedom had been possible. This is precisely how the stand-alone cost criterion for price ceilings, which will be described below, emerged from contestability theory and could not have been deduced from the model of perfect competition.

To see just how contestability theory can be used to guide regulation we

can do no better than to follow the outlines of a remarkable decision of the Interstate Commerce Commission (1985), a decision that encompasses the foundations of its current policies toward those elements of railroad activities in which competitive pressures are judged to be inadequate (that is, in regulatory terminology, in areas in which a railroad possesses "market dominance," or its shippers are "captive").

Early in the discussion of its economic framework the decision provides a section headed "Contestable Markets." However while, as we will see presently, much is made of the logic of contestability, the decision asserts flatly and quite appropriately that the pertinent arena is *not* contestable: "the railroad industry is recognized to have barriers to entry and exit and thus is not considered contestable for captive traffic" (p. 10).

The question, then, is what is best done to control the pricing terms on which such traffic is served. The Commission adopted a set of rules which it termed "constrained market pricing." Before getting to the details of those rules it is desirable to take note of two fundamental attributes underlying the approach: rate of return-rate base regulation of overall earnings, and acceptance of differential (non uniform) pricing. On the first of these, the Commission undertakes to avoid decisions that preclude a railroad from earning in the long run what it refers to as "adequate revenues." Following long regulatory tradition, these are defined in terms of a permitted rate of return on the railroad's rate base (its total invested capital). What is new here is that the Commission adopts for this purpose most of the criteria called for by economic analysis, for example determining that "'adequate' returns are those that provide a rate of return on net investment equal to the current cost of capital (i.e., the level of return available on alternative investments)" (p. 18).

Second, the Commission recognizes that solvency of the railroads is likely to require differential pricing:

> Most importantly, railroads exhibit significant economies of scope and density. Economy of *scope* refers to the fact that the rail plant is indivisible and can produce numerous services at less cost than those services could be produced by separate rail plants for each service. Economy of *density* refers to the fact that greater use of the fixed plant results in a declining average cost. Thus, the marginal cost of rail service is less than the average cost, because the fixed plant is used in a progressively more efficient manner. The differential between marginal costs and average costs cannot be assigned directly to specific movements by any conventional accounting methodology. Hence, we refer to it as the "unattributable costs." These are the costs which must be covered through differential pricing. (pp. 7–8, footnotes omitted)

The decision goes on to point out that where unattributable costs cannot be covered by marginal cost pricing, then demand considerations as well as cost data must enter into decision making, both in order to permit adequacy of revenues and in order to achieve efficiency.

Any means of allocating these costs among shippers other than actual market demand is arbitrary and may not permit a carrier to cover all of its costs. This is because non-demand-based cost apportionment methods do not necessarily reflect the carrier's ability (or inability) to impose the assigned allocations and cover its costs. Thus, they frequently "over-assign" or "under-assign" the carrier's unattributable costs to particular services. If a carrier sought to apply the formula price to all of its traffic, it would lose that traffic for which the demand could not support the price assigned. In that event, the remaining shippers might be required to pay a larger portion of the carrier's unattributable costs because they would lose the benefit of sharing these costs with the lost traffic.

"Ramsey pricing" is a widely recognized method of differential pricing, that is, pricing in accordance with demand. Under Ramsey pricing, each price or rate contains a mark-up above the long-run marginal cost of the product or service to cover a portion of the unattributable costs. The unattributable costs are allocated among the purchasers or users in inverse relation to their demand elasticity. Thus, in a market where shippers are very sensitive to price changes (a highly elastic market), the mark-up would be smaller than in a market where shippers are less price sensitive. The sum of the mark-ups equals the unattributable costs of an efficient producer. (p. 8, footnote omitted)

Nevertheless, the Commission comments,

Ramsey pricing is based on a mathematical formula which requires both the marginal cost and the elasticity of demand to be quantified for every movement in the carrier's system. Thus, the amount of data and degree of analysis required seemed overwhelming. We concluded that while formula Ramsey pricing is useful as a theoretical guideline, it is too difficult and burdensome for universal application. In setting flat rates, that is, rates which do not vary with the volume shipped, Ramsey pricing, in principle, yields the least inefficient price structure. However, even under pure Ramsey pricing, output levels are less than they would be if rates were set at marginal costs. This results in an economic inefficiency because the value of the lost output to the shipper is greater than the value of the resources saved by reducing output. In such a situation, it may be feasible for the parties to negotiate a contract which will leave both parties better off than at the [flat] Ramsey price. Rail freight differs from financing government services (for which shippers are relatively large and few in number). Thus, the feasibility of contracting is more evident for rail freight than for these other services.

As an alternative to pure Ramsey pricing, we proposed Constrained Market Pricing. (p. 9, footnotes omitted)

The key issue to be faced by constrained market pricing is the formulation of a criterion to be used in setting a ceiling over the price to be charged for traffic over which a railroad possesses market dominance. This is an issue with which the Commission had been struggling only since 1978. Before that, curiously, the Commission had been preoccupied primarily with the setting of floors beneath rail prices, to prevent railroads from undercutting barge and truck competitors (the Commission had once described its role as that of a "a giant handicapper").

In the early 1980's, in its search for a tenable price ceiling rule, it turned

initially to the tools it had designed long before in its floor setting endeavors. Here, at least for the bulk of the postwar period, it had been using an accounting concept which it called "fully allocated cost." In essence, this assigned to each part of a railroad's traffic a figure intended to approximate the incremental cost of that traffic, plus a share of the remaining (unattributable) cost proportionate to some criterion of "relative use" such as number of carloads, weight or monetary value of the commodity in question carried by the railroad. The apportionment of the unattributable costs was admittedly arbitrary, bore no necessary relation to the cost data (such as marginal costs) that an economist would consider pertinent, and allowed for no adjustments for variations in demand conditions. Even earlier than the landmark Ingot Molds Case (1967) in which the Supreme Court threw up its hands on the matter, economists had begun to argue before the ICC that the use of a fully distributed cost floor was pernicious and a source of substantial inefficiency, and that marginal or incremental costs were the only defensible cost data for use in the calculation of price floors, a view that gradually acquired acceptance among regulatory agencies, including the ICC.

It seems odd, then, that in 1978, when the Commission began its search for a rate ceiling formula (Interstate Commerce Commission (ICC) (1978)) it turned to fully allocated cost, now discredited as a cost floor, as its candidate for a viable cost ceiling. Initially, it proposed for this purpose to allocate a railroad's unattributable costs in proportion to the variable costs (as a proxy for the incremental costs) of the various portions of its traffic. Because this formula soon proved fatally restrictive, particularly in ignoring demand considerations, the Commission permitted the railroads leeway to charge prices as much as seven percent above fully allocated costs where demand conditions permitted. The courts agreed to the legitimacy of this move toward differential pricing but rejected the use of the arbitrary seven percent range of tolerance (see, e.g., ICC (1979)). The Commission made one more attempt to salvage the fully allocated cost approach to price ceilings, but this time adopted weight and distance traversed (calling it the "ton/ton-mile" method) as its basis of allocation (ICC (1980)). However, in response to careful arguments by the railroads and others, it withdrew from this approach altogether in 1981 (ICC (1981)), and early in 1983 offered constrained market pricing as its alternative.

Here, the ceiling proposed by the Commission was *stand-alone cost*, a concept it acknowledged to have derived from the contestability literature.

> " . . . stand-alone cost (SAC) test . . . is used to compute the rate a competitor in the market-place would need to charge in serving a captive shipper or a group of shippers who benefit from sharing joint and common costs. A rate level calculated by the SAC methodology represents the theoretical maximum rate that a railroad could levy on shippers without substantial diversion of traffic to a hypothetical competing service. It is, in other words, a simulated competitive price. (The competing service could be a shipper providing service for itself or a third party

competing with the incumbent railroad for traffic. In either case, the SAC represents the minimum cost of an alternative to the service provided by the incumbent railroad.)

The theory behind SAC is best explained by the concept of 'contestable markets.' This recently developed economic theory augments the classical economic model of 'pure competition' with a model which focuses on the entry and exit from an industry as a measure of economic efficiency. The theory of contestable markets is more general than that of 'pure competition' because it does not require a large number of firms. In fact, even a monopoly can be contestable. The underlying premise is that a monopolist or oligopolist will behave efficiently and competitively where there is a threat of losing some or all of its markets to a new entrant. In other words, contestable markets have competitive characteristics which preclude monopoly pricing (p. 10).

Here it is worth reviewing the logic of the stand-alone cost ceiling a bit more closely. The first-best lesson of the perfect competition model, calling for prices to be set equal to marginal costs, has no doubt contributed to the common regulatory ethos which *equates* price to *some* measure of cost. This doctrine has been used frequently where it is completely inappropriate and without logical foundation, that is, in cases where prices should be based on demand as well as cost considerations, because of the presence of economies of scale and scope. Such arbitrary measures as fully distributed costs cannot substitute for marginal cost measures as decision rules for proper pricing, and the search for a substitute is a remnant of inappropriate reliance on the model of perfect competition for guidance in regulation.

In contrast, contestability theory suggests cost measures that are appropriate guideposts for regulated pricing—incremental and stand-alone costs. The incremental cost of a given service is, of course, the increment in the total costs of the supplying firm when that service is added to its product line. In perfectly contestable markets, the price of a product will lie somewhere between its incremental and its stand-alone cost, just where it falls in that range depending on the state of demand. One cannot legitimately infer that monopoly power is exercised from data showing that prices do not exceed stand-alone costs, and stand-alone costs constitute the proper cost-based ceilings upon prices, preventing both cross-subsidization and the exercise of monopoly power (see Faulhaber (1975) for tests of cross-subsidy and their equivalence). A simple example will show why this is so.

First, suppose that a firm supplies two services, A and B, which *share no costs* and that each costs 10 units a year to supply. The availability of effective potential competition would force revenues from each service to equal 10 units a year. For higher earnings would attract (profitable) entry, and lower revenues would drive the supplier out of business. In this case, in which common costs are absent, incremental and stand-alone costs are equal to each other and to revenues, and the competitive and contestability benchmarks yield the same results.

Next, suppose instead that of the 20 unit total costs 4 are fixed and common to A and B, while 16 are variable, 8 of the 16 being attributable to A and 8 to B. If, because of demand conditions, at most only a bit more than 8 can be garnered from consumers of A, then a firm operating and surviving in contestable markets will earn a bit less than 12 from B. These prices lie between incremental costs (8) and stand-alone costs (12), are mutually advantageous to consumers of both services, and will attract no entrants, even in the absence of any entry barriers. In contrast, should the firm attempt to raise the revenues obtained from B above the 12 unit stand-alone cost, it would lose its business to competitors willing to charge less. Similarly, the same fate would befall it in contestable markets if it priced B in a way that earned more than 8 plus the common cost of 4, less the contribution toward that common cost from service A.

Thus, the forces of idealized potential competition in perfectly contestable markets enforce cost constraints on prices, but prices remain sensitive to demands as well. Actual and potential competition are *effective* if they constrain rates in this way, and in such circumstances regulatory intervention is completely unwarranted. But if, in fact, market forces are not sufficiently strong, then there is likely to be a proper role for regulation, and the theoretical guidelines derived from the workings of contestable markets are the appropriate ones to apply. That is, prices must be constrained to lie between incremental and stand-alone costs.[3,4] This is the approach adopted by the ICC to determine maximum rates for railroad services, and the method has already withstood appeals to the Federal courts.

[3] Note that, properly applied, the SAC criterion should hold for all subsets of a firm's products and not only for its individual products. Thus, suppose a firm produces three items, A, B and C, and that a widget of limitless capacity must be used in order to produce any A, any B or any combination of the two. The cost of the widget will then not enter either A's or B's incremental cost, but it will clearly consititute part of the incremental cost of supplying them both. Then, the combined total revenue of the two items must suffice to cover the cost of the widget as well as any other incremental cost incurred individually or in common. Note also that efficiency requires that a supplier never set price below either marginal *or* per unit incremental cost, because if it is set so low it may lure away customers from a more efficient supplier, i.e., one who can supply the product at lower marginal or per unit incremental cost, and hence supply either one unit of the product or its entire amount at lower resources cost.

[4] It is easy to prove that if a firm's total earnings are exactly equal to its cost of capital, i.e., if it earns exactly zero economic profits, then if all of its revenues equal or exceed the corresponding incremental costs they are *automatically guaranteed* not to exceed their stand-alone costs and *vice versa*. That is, the passing of either test (the SAC ceiling or the incremental cost floor) automatically demonstrates that the other must also be passed by the prices at issue. To show this simply we deal with the case of two products, though the proof is perfectly general. Let

X_1, X_2 = the quantities of the outputs supplied by a firm

P_1, P_2 = the prices of the products

$C(X_1, X_2)$ = the total cost function

$C(X_1, O)$ = the stand-alone cost of product 1 and

$C(X_1, X_2) - C(X_1, O)$ = the incremental cost of product 2.

Concluding remarks

Whatever one's attitude toward contestability theory or the policy recommendations that have derived from it, it must surely be agreed that it has evoked a flood of imaginative and valuable research and writing in opposition, in extension, and in application. To see how rapidly it and associated concepts have spread one need only recall that as recently as 1970 the concept of Ramsey pricing was unknown to most economists, though the analysis, of course, had appeared in 1927. Indeed, the term "Ramsey pricing" was coined by the present authors less than a decade ago. Terms such as "economies of scope," "stand-alone cost" and "contestable markets" were coined in the 1970's as well. Yet today they are used routinely not only in professional journals, but in hearings before U.S. courts and regulatory agencies. Clearly, propagation of the substance rather than the terminology is what really counts, but this too, is being achieved.

Thus, whatever contribution the future will judge contestability theory to have made, it will surely conclude that the analysis has succeeded in stimulating thought in both the realm of academic research and policy formulation.

Princeton and New York Universities.

REFERENCES

APPLEBAUM, E. and LIM, C., "Contestable Markets Under Uncertainty, *Rand Journal of Economics*, 16, 1, Spring 1985, pp. 28–40.

BAILEY, ELIZABETH, E., "Price and Productivity Change Following Deregulation: The U.S. Experience," *Economic Journal*, March 1986 (forthcoming).

BAILEY, ELIZABETH E. and BAUMOL, WILLIAM J., "Deregulation and the Theory of Contestable Markets," *Yale Journal of Regulation*, 1984, 1, pp. 111–137.

BAILEY, E. E. and FRIEDLAENDER, A. F., "Market Structure and Multiproduct Industries," *Journal of Economic Literature*, Vol. 20, September 1982, pp. 1024–1048.

BAUMOL, W. J., PANZAR, J. C. and WILLIG, R. D., *Contestable Markets and the Theory of Industry Structure*, San Diego: 1982.

BAUMOL, W. J., PANZAR, J. C., and WILLIG, R. D., "Contestable Markets: An Uprising in the Theory of Industry Structure: Reply," *American Economic Review*, 73, June, 1983, pp. 491–496.

BITTLINGMAYER, G., "The Economics of A Simple Airline Network," mimeo, 1985.

Then, zero economic profit requires

(1) $P_1X_1 + P_2X_2 = C(X_1, X_2)$

and the incremental cost test for product 2 requires

(2) $P_2X_2 \geq C(X_1, X_2) - C(X_1, O)$.

Subtraction of (2) from (1) immediately yields the stand-alone cost criterion for product 1,

(3) $P_1X_1 \leq C(X_1, O)$.

Similarly, subtraction of (3) from (1) immediately yields (2). Q.E.D.

BROCK, W. A., "Contestable Markets and the Theory of Industry Structure: A Review Article," *Journal of Political Economy*, 91, 6, December 1983, pp. 1055–1066.

BROCK, W. A. and SCHEINKMAN, J. A., "Free Entry and the Sustainability of Natural Monopoly: Bertrand Revisited by Cournot," in Evans, D. S. (ed.), *Breaking up Bell: Essays on Industrial Organization and Regulation*, 1983.

CALL, GREGORY D. and KEELER, THEODORE E., "Airline Deregulation, Fares, and Market Behavior: Some Empirical Evidence," in *Analytical Studies in Transport Economics*, 1984.

CASSIDY, PETER A., *Australian Overseas Cargo Liner Shipping*, University of Queensland, 1982.

CAVES, D. W., CHRISTENSEN, L. R., and HERRIGES, J. A., "Modelling Alternative Residential Peak-Load Electricity Rate Structures," *Journal of Econometrics*, Vol. 24(3), 1984, pp. 249–268.

CAVES, D. W., CHRISTENSEN, L. R., and SWANSON, J. A., "Productivity Growth, Scale Economies, and Capacity Utilization in U.S. Railroads, 1955–74," *American Economic Review*, December 1981, pp. 994–1002.

CAVES, D. W., CHRISTENSEN, L. R., and TRETHEWAY, M. W., "Flexible Cost Functions for Multiproduct Firms," *Review of Economics and Statistics*, Vol. 62(3), August 1980, pp. 477–481.

CAVES, D. W., CHRISTENSEN, L. R., and TRETHEWAY, M. W., "Productivity Performance of U.S. Trunk and Local Services Airlines in the Era of Deregulation," *Economic Inquiry*, Vol. XXI, July 1983, pp. 312–324.

CHARNES, A., COOPER, W. W., and SUEYOSHI, T., "A Goal Programing/Constrained Regression Review of the Bell System Breakup," Center for Cybernetic Studies Report 513, University of Texas at Austin, May 1985.

COURSEY, D., ISAAC, R. M., and SMITH, V. L., "Natural Monopoly and Contested Markets: Some Experimental Results," *Journal of Law and Economics*, Vol. 27, April, 1984, pp. 91–113.

COURSEY, D., ISAAC, R. M., LUKE, M., and SMITH, V. L., "Market Contestability in the Presence of Sunk (Entry) Costs," *Rand Journal of Economics*, Vol. 15, Spring 1984, pp. 69–84.

COWING, T. G. and HOLTMANN, A. G., "Multiproduct Short-Run Hospital Cost Functions: Evidence and Policy Implications from Cross-Section Data," *Southern Journal of Economics*, January 1983, pp. 637–653.

DASGUPTA, P. and STIGLITZ, J. E., "Sunk Costs, Competition, and Welfare," mineo, 1985.

DAVIES, J. E., "The Theory of Contestable Markets and its Application to the Liner Shipping Industry," Canadian Transport Commission, Ottawa-Hull, 1986.

DENNY, M., FUSS, M., EVERSON, C., and WAVERMAN, L., "Estimating the Effects of Diffusion of Technological Innovations in Telecommunications: The Production Structure of Bell Canada," *Canadian Journal of Economics*, February 1981, pp. 24–43.

EVANS, D. S. and HECKMAN, J. J., "A Test for Subadditivity of the Cost Function with an Application to the Bell System," *American Economic Review*, September, 1984, pp. 615–623.

FAULHABER, G., "Cross-Subsidization: Pricing in Public Enterprise," *American Economic Review*, 65, December 1975, pp. 966–977.

FAULHABER, G. and LEVINSON, S., "Subsidy-Free Prices and Anonymous Equity," *American Economic Review*, 71, 1981, pp. 1083–1091.

FEDERAL COMMUNICATIONS COMMISSION, *Notice of Inquiry, in the Matter of Long-Run Regulation at AT & T's Basic Domestic Interstate Sources*, Washington, D.C., 1985.

FEDERAL TRADE COMMISSION, *Final Order: In the Matter of the Echlin Manufacturing Company, and Borg-Warner Corporation*, Docket Number 9157, Washington, D.C., June 28, 1985.

FRIEDLAENDER, A., WINSTON, C., and KUNG WANG, D., "Costs, Technology and Productivity in the U.S. Automobile Industry," *Bell Journal of Economics*, Spring 1983, pp. 1–20.

FROEB, L. and GEWEKE, J., "Perfect Contestability and the Postwar U.S. Aluminium Industry," Tulane University, December 1984.

GILLIGAN, T. and SMIRLOCK, M., "An Empirical Study of Joint Production and Scale Economies in Commerical Banking," *Journal of Banking and Finance*, vol. 8, 1984, pp. 67–77.

GILLIGAN, T., SMIRLOCK, M., and MARSHALL, W., "Scale and Scope Economies in the Multi-Product Banking Firm," *Journal of Monetary Economics*, Vol. 13, 1984, pp. 393–405.

GORMAN, I. E., "Conditions for Economies of Scope in the Presence of Fixed Costs," *Rand Journal of Economics*, 16, 3, Autumn 1985, pp. 431–436.

GRAHAM, D. R., KAPLAN, D. P., and SIBLEY, D. S., "Efficiency and Competition in the Airline Industry," *Bell Journal of Economics*, Spring 1983, pp. 118–138.

GROSSMAN, S. J., "Nash Equilibrium and the Industrial Organization of Markets with Large Fixed Costs," *Econometrica*, 49, 1981, pp. 1149–1172.

HARRISON, G. W., "Experimental Evaluation of the Contestable Markets Hypothesis," in E. E. Bailey (ed.) *Regulation at the Crossroads*, Cambridge, Mass., 1986.

HARRISON, G. W. and MCKEE, M., "Monopoly Behavior, Decentralized Regulation, and Contestable Markets: An Experimental Evaluation," *Rand Journal of Economics*, Spring 1985, 16, pp. 51–69.

HARRISON, G. W., MCKEE, M., and RUTSTRÜM, E. E., "Experimental Evaluation of Institutions of Monopoly Restraint," in Green, L. and Kagel, J. H. (eds.), *Advances in Behavioral Economics* (Norwood: Ablex, 1987, forthcoming).

HELPMAN, E. and KRUGMAN, P. R., *Market Structure and Foreign Trade*, 1985.

INTERSTATE COMMERCE COMMISSION, Ex Part No. 347, "Western Coal Investigation—Guidelines for Railroad Rate Structure", 43, *Federal Register*, 22151, May 22, 1978.

INTERSTATE COMMERCE COMMISSION, *Annual Volume Rates on Coal—Wyoming to Flint Creek, Arkansas*, 361, ICC 539, 1979.

INTERSTATE COMMERCE COMMISSION, Ex Parte No. 347 (Sub-No. 1), "Coal Rate Guidelines—Nationwide," 45, *Federal Register*, 80370, December 4, 1980.

INTERSTATE COMMERCE COMMISSION, Ex Parte No. 347 (Sub-No. 1), "Coal Rate Guidelines—Nationwide," December 21, 1981, unpublished.

INTERSTATE COMMERCE COMMISSION, "Coal Rate Guidelines, Nationwide," Ex Parte No. 347 (Sub-No. 1), Washington, D.C., decided August 3, 1985.

KELLNER, S. and MATHEWSON, G. F., "Entry, Size Distribution, Scale, and Scope Economies in the Life Insurance Industry," *Journal of Business*, Vol. 56, No. 1, 1983, pp. 25–44.

KIM, H. Y. and CLARK, R. M., "Estimating Multiproduct Scale Economies: An Application to Water Supplies," U.S. Environmental Protection Agency, Municipal Environment Research Laboratory, Cincinnati, Ohio, July 1983.

KNIEPS, G. and VOGELSANG, I., "The Sustainability Concept under Alternative Behavioral Assumptions," *Bell Journal of Economics*, 13, 1, Spring 1982, pp. 234–241.

KREPS, D. and SCHEINKMAN, J., "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes," *Bell Journal of Economics*, 14, 2, Autumn 1983, pp. 326–338.

MASKIN, E. and TIROLE, J., "A Theory of Dynamic Oligopoly," mineo, 1982.

MAYO, J. W., "The Technological Determinants of the U.S. Energy Industry Structure," *Review of Economics and Statistics*, January 1984, pp. 51–58.

MESTER, L. J., "A Multiproduct Cost Study of Savings and Loans," essay 1, doctoral dissertation, Princeton University, October 1985.

MIRMAN, L. J., TAUMAN, Y., and ZANG, I., "Monopoly and Sustainable Prices As a Nash Equilibrium in Contestable Markets," Mimeo, 1983.

MIRMAN, L. J., TAUMAN, Y., and ZANG, I., "Supportability, Sustainability and Subsidy-Free Prices," *Rand Journal of Economics*, 16, 1, Spring 1985, pp. 114–126.

MORRISON, S. A. and WINSTON, C., "Empirical Implications and Tests of the Contestability Hypothesis," manuscript, 1985.

MURRAY, J. D. and WHITE, R. W., "Economies of Scale and Economies of Scope in Multiproduct Financial Institutions: A Study of British Columbian Credit Unions," *Journal of Finance*, June 1983, pp. 887–901.

PANZAR, J. C. and POSTLEWAITE, A., "The Sustainability of Ramsey Optimal Non-linear Prices," mimeo, 1985.

PERRY, M., "Sustainable Positive Profit Multiple-Price Strategies in Contestable Markets," mimeo, 1982.

QUIRMBACH, H., "Vertical Integration, Contestable Markets, and the Misfortunes of the Misshaped U," mimeo, 1982.

SCHWARZ, MARIUS, "The Nature and Scope of Contestability Theory," Georgetown University, September 1985, unpublished.

SCHWARTZ, MARIUS and REYNOLDS, ROBERT J., "Contestable Markets: A Uprising in the Theory of Industry Structure: Comment," *American Economic Review*, June 1983, 73, pp. 488–490.

SCOTT, J. T., "Multimarket Contract and Economic Performance," *Review of Economics and Statistics*, August 1982, pp. 368–375.

SHARKEY, W. W., "Existence of Sustainable Prices for Natural Monopoly Outputs," *Bell Journal of Economics*, 12, 1, Spring 1981, pp. 144–154.

SHARKEY, W. W., *The Theory of Natural Monopoly*, 1982.

SHEPHERD, W. G., "'Contestability' vs. Competition," *American Economic Review*, 74, September 1984, pp. 572–587.

SPENCE, A. M., "Contestable Markets and the Theory of Industry Structure: A Review Article," *Journal of Economic Literature*, 21, 1983, pp. 981–990.

SPULBER, D., "Scale Economies and Existence of Sustainable Monopoly Prices," *Journal of Economic Theory*, 34, 1984, pp. 149–163.

TAUCHEN, H., FRAVEL, F. D., and GILBERT, G., "Cost Structure and the Intercity Bus Industry," *Journal of Transport Economics and Policy*, January 1983, pp. 25–47.

TEECE, D. J., "Economies of Scope and the Scope of the Enterprise," *Journal of Economic Behavior and Organization*, Vol. 1, September 1980, pp. 223–247.

TEECE, D. J., "Towards an Economic Theory of the Multiproduct Firm," *Journal of Economic Behavior and Organization*, March 1982, pp. 39–63.

TEN RAA, T., "Supportability and Anonymous Equity," *Journal of Economic Theory*, 31, 1983, *pp.* 176–181.

TEN RAA, T., "Resolution of Conjectures on the Sustainability of Natural Monopoly," *Rand Journal of Economics*, 15, 1, Spring 1984, pp. 135–141.

U.S. DEPARTMENT OF JUSTICE, *Merger Guidelines*, Washington, D.C. 1982.

U.S. SUPREME COURT, American Commerical Lines, Inc., *et al.* vs. Lousiville and Nashville Railroad Co., *et al.* (Ingot Molds case), October term, 1967, pp. 571–597.

WEITZMAN, M. L., "Contestable Markets: An Uprising in the Theory of Industry Structure: Comment," *American Economic Review*, 73, June 1983, pp. 486–487.