EMPIRICAL STUDIES OF RIVALROUS BEHAVIOR

Richard L. Schmalensee

WP #1990-88 February 1988
Empirical Studies of Rivalrous Behavior

Richard Schmalensee

Massachusetts Institute of Technology

Text of a lecture to be delivered May 23, 1988 at the Fondazione Einaudi in Turino, Italy, in a series sponsored by the Associazione Borsisti Luciano Jona. To be published by Oxford University Press in a volume edited by Giacomo Bonanno.

February 1988
The landscape of industrial organization, particularly in the U.S., was for many years dominated by a bitter struggle between Harvard and Chicago. Chicago's partisans assumed that rivalry would generally be intense in the absence of government regulation or effective cartel arrangements. They accordingly relied on the price theory of Alfred Marshall and, for the cartel case, Joan Robinson. Harvard's loyalists, on the other hand, assumed that competition was often seriously imperfect even in the absence of regulation or cartels. They tended to employ less formal analytical frameworks derived from the works of Edward Chamberlin and William Fellner.

The differences in the dialects of economics spoken by the two Schools were such that debates about scientific issues often seemed in need of translators. Empirical work by each school tended to support its basic assumption about market conduct. Policy prescriptions also reflected those assumptions. Harvard called for an activist government policy to deal with serious imperfections in competition, while Chicago, believing that such imperfections were rare, argued for laissez faire.

In the last two decades the study of industrial organization has been transformed by rigorous theoretical analyses of imperfect competition. Graduate courses in industrial organization now cover a host of sophisticated, game-theoretic models of imperfect competition, some which have been reviewed in the other chapters in this volume. There are no deep divisions or distinct Schools in this domain; broadly similar theoretical papers are written and studied at Harvard, Chicago, and other leading institutions in the U.S. and abroad. The dialect of game theory is spoken everywhere.
One might expect that this convergence in theoretical method and the development of a large and widely studied body of theoretical literature would have served to narrow differences in the empirical and policy domains. But this does not seem to have happened. Methods and assumptions employed in empirical work seem if anything more diverse than two decades ago, and the controversies about basic factual questions that raged then continue largely unabated now. And, even though distinct Harvard and Chicago Schools are no longer present in the theoretical literature, defenders of their traditional positions still dominate many policy debates.

In the remainder of this essay I consider causes and potential cures for this fragmentation on matters of substance, concentrating on the actual and potential contribution of empirical studies of rivalrous behavior in real markets.

In the next section I argue that theoretical work in industrial organization has, somewhat paradoxically, made it clear that empirical research is absolutely critical to progress in this field. The current fragmentation of the field reflects in large measure the loss of faith in the two approaches to empirical research—comprehensive industry case studies and cross-section industry-level profitability studies—that dominated empirical work in industrial organization until roughly the start of this decade. These approaches are discussed in Sections 2 and 3, respectively.

The 1980's have witnessed what Tim Bresnahan and I (1987) have been incautious enough to call an empirical renaissance in industrial economics. Compared with the earlier literature, recent research is notable for its methodological diversity. Sections 4 and 5 discuss some particularly promising approaches to industry-specific and inter-industry studies,
respectively, that figure prominently in recent work. Though the organization of Sections 2-5 highlights methods used rather than questions addressed, important empirical findings and problems are discussed as well. Section 6 briefly summarizes some of the main themes that emerge from this overview and offers some modest prescriptions.

1. Theory and Empirics

Industrial organization is primarily concerned with the behavior of business firms in their roles as sellers and with the implications of that behavior for the operation of markets and the design of public policy. While game-theoretic tools are well-suited in principle to the analysis of key aspects of business behavior, their application has neither produced a general theory of market operation nor given one much reason to expect such a theory to emerge in the foreseeable future. Under these conditions, empirical research becomes critical to scientific progress.

The Nature of Recent Theorizing. Game theory was developed in large part to model rational behavior in small numbers situations, and many industrial markets have only a few important sellers. Indeed, in retrospect the only surprise is that game theory took forty years to conquer oligopoly theory. But there is more to the appeal of game theory than this.

Many important aspects of business conduct are inherently dynamic. Actions and reactions take time; productive assets are often long-lived; entry and exit decisions turn on before-and-after comparisons. In addition, information in real markets is rarely complete, perfect, or symmetric. Potential entrants may know less about a market than established firms, for instance, and individual established firms may know only their own costs,
not those of their rivals. Recent advances in extensive form game theory have greatly facilitated the analysis of situations that involve strategic behavior over time in settings in which information is incomplete and asymmetric. There is no comparably attractive approach to the analysis of such situations anywhere in sight.

Unfortunately, however, the general impression that has emerged from a decade's extensive use of the game-theoretic approach is "Anything can happen!" The diversity and growth of the theoretical literature provides a good deal of support for the conjecture that almost any remotely plausible pattern of conduct -- anything that has ever been alleged with a straight face in an antitrust case, say -- can appear in an equilibrium in an apparently plausible game-theoretic model. Policy implications are in some sense even more varied, since efficient policy in many models depends on details of parameter values and functional forms.

This situation flows from two apparently general features of game-theoretic models of market behavior. The first is that many apparently simple multi-period games of incomplete information have multiple equilibria -- often an uncountable infinity of equilibria. Though considerable work has been done to refine the definition of equilibrium in order to mitigate this problem, it still remains endemic. Current practice often involves selecting a single equilibrium on the basis of model-specific plausibility arguments. Sometimes these arguments are compelling; in other cases "anything can happen" is the only way to summarize the multitude of a single model's equilibria.

Even game-theoretic models that have unique equilibria possess a second feature that is in some respects even more troubling: the
predictions of game-theoretic models seem delicate and are often difficult
to test. Important qualitative features of equilibria often depend
critically on whether prices or quantities are choice variables, on whether
discrete or continuous time is assumed, on whether moves are sequential or
simultaneous, and, perhaps most disturbing of all, on how players with
incomplete information are assumed to alter their beliefs in response to
events that do not occur in equilibrium. When information is incomplete,
strategies depend on unobservable beliefs, and the often empirically
questionable assumption that key parameters and probability distributions
are common knowledge is frequently central to the analysis. The level of
rationality required of actors in many game-theoretic models seems to exceed
the capabilities of all but the best economic theorists.

Thus game-theoretic modeling has taught us a great deal about what
might happen in a variety of situations, but relatively little about what
must happen conditional on observables. Game theory has proven better at
generating internally consistent scenarios than at providing plausible and
testable restrictions on real behavior. It seems almost certain that the
theoretical literature contains a sizeable number of what Clapham (1922)
called "empty boxes" - internally consistent models that describe no real
markets - and that more such boxes are being constructed daily. But there
is no consensus on which models belong to this class and which others should
be taken particularly seriously because their predictions are often correct.
Thus disputes about issues of fact and policy can and do flourish despite an
impressive array of theoretical results and insights.

Moreover, the very diversity of predictions in the theoretical
literature seems to cast doubt on the value of theory in this field. The
sensitivity of equilibrium outcomes to modeling details seems to suggest that one must know more than anyone is likely ever to know about any real market in order to use theory to make definite predictions about conduct and performance.

The Roles of Empirical Research Absent a theoretical breakthrough, it seems clear that only empirical research can alter this state of affairs, since only empirical testing can definitively establish the domain of applicability of any model. This is the most-cited role of empirical research in this or any field: testing theories in a variety of settings to see under what conditions, if any, their predictions are valid. Because most of the theoretical literature in industrial organization remains untested in this sense, theorists are free to generalize, extend, innovate, and explore with essentially no external constraints.

Game-theoretic models are particularly hard to test, since their predictions are sensitive to market details it is often difficult or (in the case of models in which beliefs and expectations play a critical role) impossible to observe. But, as Friedman (1953) stressed, predictions are ultimately what matters. That is, in situations in which a model's assumptions are at least plausible (Friedman might argue against this condition), one can concentrate on confronting its predictions with the facts. The domain of any model's applicability can then be determined by analysis of the similarities and differences between and among situations in which it does and does not perform well.

While most general discussions of empirical research stress theory testing, such research also has two other critical roles to play in industrial organization. The first is simply to provide facts with which
theories must be consistent. Tests that reject widely accepted theories or indicate that they have a narrow range of validity yield such facts automatically, of course. But empirical studies not primarily aimed at theory testing may also play this role. Early astronomers aimed simply to describe the heavens; their results on the movements of the planets provided key facts in the development of theories of gravitation.

Unfortunately, few theories are ever conclusively rejected in economics. And the low value placed on descriptive research by much of the economics profession depresses the volume of such research. Industrial organization theorists thus confront a relatively small fact base and are free to build models that relate to no well-described real market, since few markets are in fact well-described.

The third role of empirical research in industrial organization is to inform the analysis of particular industries and the design of general policy rules by providing information on the frequency with which particular market structures and patterns of conduct occur in the economy. If, for instance, predatory pricing never occurs, the ideal policy is to ignore charges of predation, since any other policy can only waste resources. This is in many respects the hardest of the three roles. We have concentration ratios for most manufacturing markets in many economies, for instance, but little comprehensive information is available on more subtle aspects of market structure, and essentially no systematic data aside from accounting profit rates is available on conduct or performance. This leaves a factual vacuum in policy debates that is quickly filled by beliefs and assumptions.
2. Classical Industry Case Studies

Industry-level case studies were central to the research program of the Harvard School when it emerged under the leadership of Edward Mason and others in the 1930's. Lacking a satisfactory theory of business behavior in concentrated, dynamic markets with incomplete information, these scholars initially hoped to develop useful generalizations inductively by carefully examining structure, conduct, and performance in many markets. These examinations tended to be reported in comprehensive, book-length studies; see Wallace (1937) for an early and influential study of the aluminum industry and Peck (1961) for an interesting follow-up.

These studies tended to be qualitative and historical in nature and to emphasize the evolution of market structure and patterns of behavior over time. Since such work required both broad and deep information, case studies written in the U.S. often relied on data and documents collected and made public in the course of antitrust cases. Thus concentrated industries that seemed, at least to antitrust authorities, to behave non-competitively tended to be over-represented in the case study literature.

The best of these early industry studies summarized mountains of documents and testimony into a consistent and generally persuasive picture. They regularly described patterns of conduct that were not easily explained by simple competitive or monopoly models. Studies of the interwar U.S. cigarette industry by Tennant (1950) and Nicholls (1951), for instance, found an early period of differing and frequently changed prices, followed by a long period of price leadership, during which list prices were identical and only rarely changed.\(^5\) Rivalry apparently shifted to advertising and away from price early in this second period.\(^6\)
This single data point was highly influential. It showed clearly the limits of both competitive and monopoly models, and it drew attention to the phenomena of price leadership, price rigidity, and shifts from price to non-price rivalry as industries matured. In addition to specific observations of this sort, classical case studies helped to shape the world views of several generations of scholars by providing a wealth of detailed qualitative information about business decision-making and its effects.

The production of high-quality comprehensive industry studies seems to have peaked around 1960. It may have become increasingly clear that the Harvard School's inductive research program, like recent theoretical research, was better at generating interesting examples and observations than useful general rules. Moreover, comprehensive case studies were time-consuming, often involved a great deal of subjective judgement, and tended to cover only a small, non-representative sample of industries for which usually private data had been made public. Systematic comparative evaluations of case studies were difficult because the individual studies were not easily summarized. And an alternative approach had become increasingly attractive.

3. Classical Cross-Section Studies

Joe Bain's (1951, 1956) seminal inter-industry cross-section profitability studies were based on the deceptively simple observation that the effective exercise of monopoly power should on average yield monopoly profits. The statistical analysis of profitability differences among large samples of markets thus seemed to promise rapid and objective development of general relationships regarding the incidence of monopolistic behavior.
Bain's own work was marked by extremely careful development of quantitative data based on detailed qualitative knowledge of the markets in his samples. Perhaps because few could confidently imitate this style or because multiple regression analysis was not routine in the 1950's, few scholars followed Bain's lead initially. But, as computation costs fell and government-supplied industry-level data became more widely available, the journals began to fill with cross-section profitability studies in the 1960's, and the production of comprehensive industry studies waned.

Much cross-section research focused on Bain's (1951) original hypothesis: effective collusion, and thus supra-competitive profitability, are more likely when concentration is high. Over time, more effort was devoted to following Bain's (1956) later work and considering the effects of entry conditions. Most early studies tended to support the existence of a positive relation between concentration and profitability, though that relation was often statistically fragile and economically weak. And a number of variables that arguably proxied for the difficulty of entry were also positively related to profitability; some of these relations (notably that involving the advertising-sales ratio) were quite robust.

But during the 1970's critics of this general approach became vocal and persuasive, and a number of empirical anomalies were uncovered. As a result, relatively few scholars at the start of this decade believed that the industry-level cross-section literature had shed much light on the structural determinants of non-competitive behavior.

Measurement Problems The design of Bain's (1951) original study and many that followed involved comparing seller profitability and concentration in a sample of manufacturing markets. The first critics of
this approach argued that neither concentration nor profitability could be observed accurately in practice. As the specifications of cross-section models became more complex and as the theoretical literature developed, additional measurement problems became apparent.

Bain had first to decide what collections of products and regions for which data were available constituted economic markets. This is often not a simple task even when detailed data are available, as any antitrust veteran will testify. Then Bain had to decide how to measure concentration. He chose the total market share of the eight largest sellers, which later work has shown to be highly, though not perfectly correlated with alternative plausible measures.

Finally, Bain had to decide how to measure seller profitability. He chose to use the average ratio of accounting profits to the balance sheet value of net worth (or owners' equity) for the firms in each market for which he was able to collect data. Unfortunately, it is by now well-known that accounting measures of the rate of return on assets or net worth are at best noisy measures of firms' true, economic rates of return. Conventional accounting systems treat inflation and depreciation improperly (Fisher and McGowan (1983)), and accounting practices vary among firms and over time.

Some authors have sought to avoid capital-related accounting biases by using the so-called price-cost margin: \( \frac{\text{revenue} - \text{labor and materials cost}}{\text{revenue}} \). But this measure has little theoretical or empirical support (Liebowitz (1982)); radically different values of the price-cost margin can yield identical rates of return on owners' investment when capital intensities differ. Recently a number of authors have used Tobin's \( q \), the ratio of a firm's market value to the replacement cost of its assets, as a
measure of profitability (Salinger (1984)). But the measurement of replacement cost inevitably relies critically on accounting data.

Even if accounting data were not inherently noisy, most large modern firms sell in multiple markets and have assets and expenses that are not easily allocated among those markets. Even though his data were from the 1930's, when diversification was less of a problem than today, Bain was thus required to define some of his markets broadly (e.g., aluminum products) so as to include most of his firms' revenues. Today, firm-level data are rarely used for market-oriented cross-section studies unless product mix information can be used to construct weighted averages of the features of the markets in which each firm sells.

Many cross-section profitability studies attempt to avoid the diversification problem by using data for individual plants or business units. Since plants' outputs are typically more homogeneous than firms', it is more frequently plausible to assign plants to particular markets than to assign entire firms. But data derived in this fashion, like those in the U.S. Census of Manufactures, omit costs that are not incurred at the plant level and tend to force the use of measures, like the price-cost margin, that ignore capital costs. The obvious alternative is to use data in which firms themselves have allocated costs to each of the markets from which they receive revenue; the U.S. Federal Trade Commission's Line of Business data (Ravenscraft (1983)) and the PIMS data set of the Strategic Planning Institute provide frequently-employed examples. The obvious danger here is that cost allocations are inevitably somewhat arbitrary.

Early defenders of the cross-section approach (e.g., Weiss (1971)) had a ready reply to attacks based on measurement error. They noted that as
long as measurement errors are random, they tend to mask true relations, not exaggerate them. Thus rather than indicting the whole cross-section approach, measurement errors provide an excuse for the weak results it frequently produces.

But the argument does not end there. Random measurement errors should cause the variance of accounting rates of return to exceed that of real, economic rates of return. And a number of authors (e.g., Alberts (1984) and Salinger (1984)) have observed that differences among accounting measures of firm profitability in the U.S. are generally too small to be easily reconciled with the existence of much monopoly power in the economy, even if measurement error is assumed away. That is, even if accounting profit measures were exact and perfectly correlated with concentration, the estimated effect of concentration on market performance would be small. The real effect, if any, must be even smaller because of measurement error.

Data on after-tax returns on equity from Bain's (1951) original study illustrate this point nicely. Suppose the competitive value of this measure of profitability is $r_c$. Then if $r$ is the after-tax rate of return on equity for some firm exercising monopoly power, we must have

\[
(1) \quad r - r_c = \frac{(1-r)(R-C)}{E} = \left[\frac{(1-r)R}{E}\right]\left[\frac{E-C}{R}\right],
\]

where $r$ is the corporate tax rate, $R$ is revenue, $C$ is total cost (including normal profit), and $E$ is owners' equity. A plausible estimate of $r_c$ is the average after-tax rate of return on equity in Bain's 20 unconcentrated industries: 6.9%. It appears that $[(1-r)R/E]$ averages about 1.12 for the firms in Bain's sample. Thus an observed $r$ of 16% corresponds to a markup over total cost $[(R-C)/R]$ of about 8.1% $[(16.0 - 6.9)/1.12]$, which would be
chosen by a monopoly facing a demand elasticity of about 12. Such a high
demand elasticity implies little monopoly power, yet only 3 of Bain's 22
concentrated industries had r's above 16%.

Bain's data also illustrate the general weakness of estimated
concentration/profitability relations. Using this same approach, the
average r in his preferred sample of concentrated industries corresponds to
a demand elasticity of about 22. For other samples (see his Table 3),
implicated elasticities range from 31 to 111, and the corresponding
profitability differences are generally insignificant.

Studies that went beyond concentration to consider conditions of entry
encountered yet another layer of measurement problems. Bain (1956)
performed small-scale case studies for each industry in his sample and
assessed barriers to entry judgmentally. Since entry will eliminate excess
profits in the absence of barriers, Bain tested - and found some support
for - an interactive hypothesis: profits are high only when both
concentration and barriers to entry are high.

Later authors generally eschewed both Bain's labor-intensive and
inherently subjective measurement approach and his theoretically plausible
interactive specification. Following the influential work of Comanor and
Wilson (1967), most studies employed additive regression models in which
both concentration and proxies for entry barriers appear as independent
variables. Not only are additive specifications suspect on a priori
grounds, but it is unclear that commonly-employed proxies, such as the
market share of a medium-sized plant and the advertising/sales ratio,
measure conditions of entry at all well. The theoretical literature
suggests that other factors that are more resistant to measurement, such as
information structures and the extent to which costs are sunk, are at least as important as scale economies. And the theoretical links between the advertising intensities of established firms, which clearly depend heavily on difficult-to-quantify features of the product involved, and entry conditions is tenuous at best.

It is perhaps not surprising in light of the discussion of Bain (1951) above that when variables intended to proxy for conditions of entry and other elements of market structure are added to cross-section profitability regressions, the coefficient of concentration is often negative or insignificant (e.g., Comanor and Wilson (1967), Porter (1976)).

Identification Problems A later and ultimately more potent stream of criticism began with Demsetz's (1973) argument that profitability and concentration could be positively correlated in cross-section even if concentration had no effect on the intensity of rivalry. His argument relied on inter-firm differences and pointed toward the endogeneity of concentration, two key themes in much recent research.

Since the standard presumption is that cross-section studies aim to reveal differences among long-run equilibria, Demsetz's argument is most naturally illustrated in that context. It is then plausible to let quantity (which one can think of as capacity) be the strategic variable and assume constant returns to scale. Thus consider a homogeneous-product industry in which firm i's constant long-run unit cost is \( c_i \). Then if \( P(Q) \) is the industry inverse demand function, \( q_i \) is firm i's output, and \( \bar{q}_i = Q - q_i \) is the output of firm i's rivals, firm i's economic profit is given by

\[
\pi_i = [P(q_i + \bar{q}_i) - c_i]q_i.
\]
The first-order condition for maximizing \( \pi_i \) can be written as follows:

\[
(P - c_i) = -q_iP'(1+\lambda_i) - \eta S_i(1+\lambda_i)P,
\]

where \( S_i = q_i/Q \) is firm i's market share, \( \eta = -P'Q/P \) is the reciprocal of the (absolute value of) the industry elasticity of demand, and \( \lambda_i = dq_i/dq_i \) is firm i's conjectural derivative.

Game theorists tend to become apoplectic at the sight of quantities like \( \lambda_i \), since they do not appear in game-theoretic equilibria. I use conjectural derivatives here, as they are used in much recent industry-specific work (following Iwata (1974)), to summarize conduct that may in fact be an imperfectly collusive equilibrium of a complex game played by real oligopolists. Generally, higher values of conjectural derivatives describe less intense rivalry: all else equal, the higher is \( \lambda_i \), the lower is firm i's output and the larger is the gap between price and its marginal cost.

Continuing the development above, substitution of (3) into (2) yields

\[
\pi_i = \eta(1+\lambda_i)(Pq_i)S_i.
\]

Firm i's accounting profit (neglecting accounting errors) will equal \( \pi_i \) plus \( \rho k_i \), where \( \rho \) is the relevant competitive rate of return and \( k_i \) is the value of firm i's equilibrium capital stock. Adding \( \rho k_i \) to both sides of (4) and dividing by \( k_i \), we obtain an expression for firm i's accounting rate of return on assets:

\[
r_i = \rho + [\eta(1+\lambda_i)]v_i S_i,
\]
where \( v_i = Pq_i/k_i \) is the reciprocal of firm \( i \)'s observed capital/output ratio. In this model, firms with lower costs tend as a consequence to have higher market shares and higher rates of return.

Finally, if \( \lambda_i = \lambda \) and \( v_i = v \) for all firms in the industry, the industry's average accounting rate of return is given by an \( S_i \)-weighted average of the \( r_i \):

\[
(6) \quad \bar{r} = \rho + [\eta(1+\lambda)v]H,
\]

where \( H = \sum(S_i)^2 \) is the Herfindahl-Hirschman measure of seller concentration. The greater are the differences among the \( c_i \), the larger will be \( H \) in this framework; concentration is endogenous in long-run equilibria.

Equation (6) directly rationalizes Demsetz's (1973) assertions: even if the intensity of rivalry, measured here by \( \lambda \), does not vary among industries, equation (6) predicts a positive correlation between concentration and profitability in cross-section. Random inter-industry differences in \( \eta \) and \( v \) could easily account for the general weakness of that correlation in practice. This view of the world cannot be distinguished from that of Bain in industry-level cross-sections. Many recent authors have accordingly turned to the analysis of intra-industry differences between firms, I discuss some of this work in Section 5.

It is important to note that endogeneity problems of this sort are ubiquitous in cross-section studies in industrial organization. Such studies can at best detect differences among long-run equilibria. But essentially all observable quantities in any market are determined by what Scherer (1980, ch. 1) has called the market's basic conditions (particularly
the nature of the product and the available technologies for production and marketing) and by business strategies, government policies, and historical accidents. Not only are basic conditions difficult to observe and quantify, but they too are endogenous in the long run as firms invest in product and process innovation.

This argument casts considerable doubt on the ability of cross-section studies in general to reveal structural relations. Consider, for instance, the strong positive cross-section relation between industry advertising-sales ratios and profitability first detected by Comanor and Wilson (1967) and subsequently found by numerous other authors. Even if this relation is not an accounting artifact, reflecting merely the failure to treat advertising with long-lived effects as an investment, it surely cannot be structural. If it were, it would imply that colluding firms could always increase their profits by increasing their advertising budgets, and this is most implausible. In this case, and in others, cross-section studies are best understood as revealing descriptive relations among endogenous variables. Such relations can be informative, but they must be interpreted with considerable care in light of the generally unobservable differences in exogenous variables that they reflect.

Finally, it is worth noting that there is some tension between the weak and inconclusive results of industry-level cross-section analyses and the fairly stark picture painted by many case studies. The latter show many apparent instances of non-competitive behavior in concentrated industries; the former find relatively little evidence that concentrated industries earn monopoly profit. This suggests that profits are often
eroded by forces other than rivalry among established firms. I discuss this suggestion further below.

4. Econometric Industry Studies

The discussion so far indicates that substantive progress in industrial organization depends critically on the productive use of non-classical approaches to empirical research. And a number of promising methods of this sort have been developed - some recently and some during the heyday of classical cross-section profitability studies. I begin here with promising approaches to the analysis of individual industries and then consider inter-industry analysis in Section 5.

In recent years a number of authors have begun to heed Leonard Weiss's (1971, p. 398) call to go "back to the industry study, but this time with regression in hand." By focusing on a single industry, they can control for the unobservable differences basic conditions that often plague the interpretation of cross-section studies. This focus of course also means that no single study can yield more than one observation on the industries that make up the economy. But many scholars have learned from the history of cross-section profitability regressions that it is better to understand one industry well than to collect difficult-to-interpret data on many.

By exploiting econometric techniques that were in large part unavailable to the authors of the classic industry studies discussed in Section 2, scholars today can exploit more fully the information in available data. In some studies, which I discuss first, this information comes from differences among separated markets in the same industry; in others the useful variations occur over time in a single market.
Inter-Market Variation  An early exemplar of this line of research is Benham's (1972) study of the advertising of eyeglasses. Benham observed that some states in the U.S. made it illegal for eyeglass vendors to advertise, while others barred only advertising that mentioned prices, and still others had no restrictions at all. Benham gathered data on prices charged for eyeglasses in various states and found that the more severe were advertising restrictions, the higher were prices on average. This does not prove that advertising tends to increase rivalry in all settings, of course, though it does seem to indicate that it can do so under some conditions -- most plausibly by reducing consumers' search costs.

A number of authors have reacted to the difficulty of measuring profitability and studied the relation between price and concentration in geographically separated markets, often with proxy variables for cost differences inserted as controls. Cotterill's (1986) study of supermarket pricing in Vermont towns provides a recent example. Most studies of this sort find seller concentration to be positively related to price. This work thus seems to provide relatively strong support for a link between concentration and collusion, since Demsetz's arguments would associate concentration with efficiency and thus with low prices. But the sources of spatial variations in concentration have not been systematically explored. And both trivial and substantial concentration effects have been detected, suggesting that the concentration-collusion relation, if any, has a substantial product-specific dimension.

In an interesting recent variation on this general theme, Bresnahan and Reiss (1987) study the number of retail establishments of various types (including veterinarians, beauty parlors, and movie theaters) operating in a
set of small, isolated U.S. towns. For each type of establishment, they basically estimate two parameters: \( P_1 \), the minimum population at which a single firm enters, and \( P_2 \), the minimum population at which a second firm appears. If \( (P_2/P_1) \) is approximately 2, entrants into one-firm markets must expect monopoly pricing to continue after their entry; values of \( (P_2/P_1) \) much above 2 indicate less favorable post-entry expectations and are thus consistent with entry deterrence. Bresnahan and Reiss find a good deal of variation in this ratio; they wisely resist the temptation to explore possible sources of inter-industry differences in their small sample.

Finally, Bresnahan (1987) uses quality differences to identify year-specific inter-market variation within the U.S. automobile industry. He employs a model of vertical product differentiation, which implies that small changes in the price of any one model of automobile affect only the two models with "adjacent" quality levels. Because rivalry is thus localized along the quality spectrum, the overall auto market can be decomposed into a set of linked submarkets. Estimated relations between pricing behavior and the identities of the participants in each submarket can then be used to test hypotheses about firm behavior. Bresnahan cannot reject collusive behavior in 1954 and 1956; he cannot reject competitive behavior for the boom year of 1955.

**Intertemporal Variation** A number of studies have analyzed changes in market behavior following some arguably exogenous event, often related to government policy. Rose (1987), for instance, compares the U.S. trucking industry before and after deregulation and finds strong evidence that labor unions captured a large fraction of the rents created by regulatory restrictions on competition. Such rent-sharing seems generally plausible a
priori, and it may account at least in part for the generally small interindustry variations in profitability discussed in Section 3. But it has proven hard to detect rent-sharing in many inter-industry studies, in part because concentration and unionization are highly correlated in U.S. cross-sections.13

Mergers naturally lend themselves to before-and-after analyses of this sort, and a large literature on the effects of mergers has emerged in recent years. Barton and Sherman (1984), for instance, studied a merger that substantially increased concentration and found that prices rose sharply after it was consummated. Many authors have found that acquired firms' shareholders generally benefit from mergers. Eckbo (1985) found that horizontal mergers also tend to benefit shareholders of rival firms. This finding might suggest that horizontal mergers tend to increase the likelihood of collusive behavior, but Eckbo's finding that rival firms' stock price increases were unrelated to the level of or change in seller concentration indicates that something more than a simple concentration-collusion relation is at work.

Many authors have recently used variations in behavior over time in periods without clearly significant exogenous events to measure the extent to which monopoly power is being exercised.14 Their studies are the most direct descendants of the classical industry studies discussed in Section 2. This research involves a particularly heavy investment in data set construction and in developing modeling strategies tailored to available industry-specific data. Accordingly, a large number of techniques for econometric industry analysis have been developed, but most have been employed only once or twice.
Lieberman's (1987) analysis of capacity expansion decisions in chemical process industries is an interesting, though somewhat atypical example of this general approach. Because of scale economies, new capacity in these industries is generally added in sizeable lumps, typically after a period of high utilization of existing capacity. Lieberman looks for significant differences between the decision rules used by established firms to add capacity and those used by new entrants, and he finds none. This finding casts doubt on the empirical validity of the many models in which overinvestment in capacity is used to deter entry. Slade's (1987) use of daily gasoline station price data (along with daily data on the wholesale price of gasoline) to estimate reaction functions provides another recent example of the direct estimation of firm decision rules.

Many recent econometric industry studies rely on variants of the static first-order condition written above as equation (3):

\[ P_i = MC_i + [(1+\lambda_i)P_i^*]q_i, \]

where \( MC_i \) replaces \( c_i \) to emphasize that marginal cost is what matters, and \( P_i \) replaces \( P \) because different firms may charge different prices when products are differentiated. If sufficient time-series data on cost and demand conditions are available, one can use this relation to estimate or test hypotheses about the intensity of rivalry, measured here by \( \lambda_i \).

In an influential early paper, Iwata (1974) studied a homogeneous-product industry and used estimates of the industry demand function and accounting estimates of firm-specific marginal costs to estimate conjectural derivatives. Most subsequent authors have avoided accounting estimates of marginal cost and have instead used data on its determinants. These
usually include input prices and capacity utilization, though Ashenfelter
and Sullivan (1987) employ data on changes in state-specific cigarette taxes
to construct (non-parametric) tests on \( \lambda \).

Baker and Bresnahan (1985) observe in effect that (7) implies that a
firm with market power will price according to

\[
(8) \quad \frac{(P_i - MC_i)}{P_i} = - \frac{(1+\lambda_i) P_i q_i}{P_i} = 1/\epsilon_i^F,
\]

where the last equality defines \( \epsilon_i^F \), firm \( i \)'s net or residual elasticity of
demand. This quantity measures the sensitivity of firm \( i \)'s demand to
changes in its price, taking into account the expected responses of \( i \)'s
rivals. Assuming expectations are on average correct, Baker and Bresnahan
show how to use data on demand and on firm-specific determinants of marginal
cost to obtain estimates of firm-specific residual demand elasticities - and
thus of markups over marginal cost.

As Bresnahan (1988) notes, econometric industry studies generally
reject competitive hypotheses in favor of alternatives involving less
intense rivalry. Like the classical industry studies discussed in Section
2, this work often relies on data made public in antitrust proceedings. It
is thus perhaps no great surprise that non-competitive behavior is
frequently detected. And, and I noted in Section 3, profit rates do not
appear to be extraordinarily high in many industries in which price is found
to be well above marginal cost. Thus the results of econometric industry
studies do not serve to rule out Chamberlinian monopolistic competition or
other models in which entry or inflated fixed costs eliminates potential
monopoly profits.
Recent work by Hall (1987) illustrates both the potential value of time series data in industrial organization and a potential pitfall in their analysis. Suppose that capital is fixed in the short run and that a firm's production function is simply \( Q = vL \), where \( v \) is a constant and \( L \) is labor input. If \( w \) is the wage rate, it follows that short-run marginal cost is equal to \( w/v \). Suppose that the firm sets price, \( P \), equal to \( \theta \) times marginal cost, where \( \theta \) is a constant. Taking first differences in the production function then yields

\[
\frac{\Delta Q}{Q} = \frac{\Delta L}{L} = \left[ \frac{wL}{PQ} - 1 \right] \frac{\Delta L}{L},
\]

where the last expression follows because \( P = \frac{\theta w}{v} \). Since all quantities in (9) except \( \theta \) are observable, simple time-series regressions can be used to estimate \( \theta \). And, since labor productivity is observed to vary procyclically (\( Q/L \) rises when \( L \) rises and falls when it falls) and labor's share of revenue (\( wL/PQ \)) is less than one, Hall's estimates of \( \theta \) are generally well above unity for two-digit U.S. industries.

Hall's analysis rests on a static model of firm conduct. In fact, as Carlton (1986) and others have noted, prices tend to be rigid over time, particularly in concentrated industries. Rotemberg and Summers (1988) argue that price rigidity and the related practice of labor hoarding (workforce rigidity) can produce procyclical variations in labor productivity in competitive industries. It would seem that dynamic models of the firm should be used to analyze changes in business behavior over time; the traditional cross-section assumption of long-run equilibrium may be highly misleading in a time-series context.
5. Inter-Industry Studies

Even though cross-section industry-level profitability studies are still out of fashion, studies comparing multiple industries appear regularly in leading journals. Only this sort of research can directly reveal patterns that hold for the economy as a whole. A hallmark of recent inter-industry research is the development of data sets that contain information not present in the industry-level cross-sections on which the bulk of the earlier literature was based. As in Section 4, the discussion here is organized around the sources of that information.

**International Differences** Comparisons between the same industry in countries at a similar stage in development hold basic conditions of products and available technologies at least approximately constant. The best-known example of this sort of study may still be Pryor's (1972) comparison of concentration ratios in manufacturing industries. Pryor found that rank correlations of these ratios among industrialized nations were high, suggesting the importance of basic conditions -- as opposed to business strategies, government policies, and historical accidents -- as determinants of concentration. Moreover, he found that concentration did not tend to decline noticeably with the size of the national market except for very small nations; for one reason or another, larger countries tend to have larger firms. (See Scherer, *et al* (1975) for more on this.)

Several Canadian and U.S. scholars have taken advantage of language and geography and studied correlates of differences between Canadian manufacturing industries and their U.S. counterparts. The comprehensive study by Caves, Porter, and Spence (1980) is an important example of this strand of research.
Survey and Interview Data Perhaps because economists tend to compare themselves with natural scientists, particularly physicists, most economic research in recent years has been based entirely on data about what firms and households actually do in the market, rather than on direct measures of attitudes, information, and beliefs. Most econometric industry studies, for instance, eschew reliance on the qualitative documentary information and testimony that were key data for the authors of the classic industry studies.

To some extent this practice reflects a belief, not present in the other social and behavioral sciences, that one can learn about people only by observing their actions, not by asking them questions. The result is that no information at all is obtained regarding the likely outcomes of experiments that nature does not perform, and variables that are not captured by conventional accounting systems or directly reflected in securities prices are not measured. Moreover, the difficulty of obtaining comparable numerical data over long periods tends to discourage work on the evolution of market structure on conduct, themes that were central to the classic case studies.

Some very interesting inter-industry work has been produced by setting these prejudices aside. There is a long tradition of "engineering" studies of scale economies, which are essentially based on structured interview; Scherer, et al (1975) is a leading example. This research aims to map out the features of best-practice technology directly, rather than trying to infer it from numerical data on plants of various (typically unknown) vintages and scales.
Porter's (1976) work on the influence of retailer behavior on market performance gives an example of the potential value of judgmental classifications of industries — in the tradition of Bain (1956). Porter made use of common sense and information about distribution channels to divide consumer products into "shopping goods," for which retailers are an important source of information, and "convenience goods," which retailers simply make available. Estimation of standard cross-section profitability equations produced very different results in these two samples, suggesting, consistent with recent theory, that information transmission mechanisms are important for market performance.

Most recently, Levin, et al (1987) conducted a large survey study aimed at detecting, among other things, what factors prevent rapid imitation from dissipating the rewards to innovation. Their findings that patents are unimportant in this regard in many industries and that the important factors vary considerably from industry to industry have a number of implications for theoretical and empirical research on technical change.

**Intra-Industry Differences** As Section 3 noted, Demsetz's (1973) critique of the classical interpretation of cross-section profitability regressions focused attention on differences among rival firms. The first studies of these differences seemed strongly to favor Demsetz's position. Porter (1979) found (as had Bain (1951)) that the profitability of leading firms was positively correlated with concentration, but the profitability of firms with small market shares was not. Ravenscraft (1983) included both concentration and market share in equations designed to explain business unit profitability. He and a number of later authors found that the
coefficient of market share was positive and significant in such regressions, while the coefficient of concentration was not.

Later studies that looked more closely at the inter-industry pattern of intra-industry differences produced less clear-cut results. A number of authors found that the intra-industry relation between profitability and market share on which Demsetz's argument seemed to rest was not particularly strong in many cases. Schmalensee (1987a) estimated equation (5) for a matched sample of U.S. manufacturing industries in 1963 and 1972 and replicated this finding. He then studied the inter-industry differences in the estimated coefficients and found little support for simple models of either the Bain or Demsetz variety. He did detect, as had several other authors, a positive cross-section relation between the coefficient of market share in (5) and the industry advertising-sales ratio. A possible explanation for these mixed results is that both Bain and Demsetz are right, but the relative importance of the mechanisms they stress varies considerably across the economy.

A number of studies support the importance of further research at the firm and business unit levels. Schmalensee (1985) found that the intra-industry variation in business unit profit rates considerably exceeds the inter-industry variation and that knowledge of a firm's profitability in one of its lines of business does not in general help predict how well its other businesses will do. Later work by Mueller (1986) and Cubbin and Geroski (1987) suggests that the performance of individual firms, particularly market leaders, over time tends to be at most weakly related to average performance of the markets in which they operate. Mueller finds that market shares of leading firms are surprisingly stable over time in many
industries. All of this says that Demsetz was clearly right in one important respect: industry averages generally hide a great deal of interesting intra-industry variation.

**Intertemporal Differences** If a cross-section of industries is observed over time, changes in the pattern of inter-industry relations may provide valuable information. Some recent studies have employed this sort of panel data, though little use has yet been made of the sophisticated econometric techniques that have been developed for the analysis of such data.

In a study concerned with the Demsetz (1973) critique, Peltzman (1977) examined the relation between changes in concentration and changes in price and productivity. He found that concentration increases tended to be associated with above-average increases in both price-cost margins and productivity and (because the productivity effect was stronger) below-average increases in prices. Later work by Gisser (1984) and others finds that productivity gains are associated with both substantial increases and substantial decreases in concentration. These results suggest a generalized version of Demsetz's world view: innovators gain at the expense of their rivals, so that concentration is endogenous in the long run, but concentration may rise or fall depending on whether initially large or small firms are the innovators.

Domowitz, Hubbard, and Petersen (1986) have constructed an industry-level panel data set for U.S. manufacturing that covers the 1958-81 period. Their research reveals, among other things, that the concentration-profitability correlation fell dramatically in the 1970's and that it moved pro-cyclically around this trend. In related work, Schmalensee (1987b) found that the average intra-industry profitability advantage of large firms
over their smaller rivals also declined over this period but that the large-firm advantage moved counter-cyclically.

To interpret these results, it is useful to re-write equations (5) and (6) in somewhat more general form:

\[(5') \quad r_i = \rho + \delta S_i \quad \text{(firm level),}\]

\[(6') \quad \bar{r} = \rho + \delta H \quad \text{(industry level).}\]

Schmalensee finds in effect that $\delta$ on average declines over time. Equation (6') indicates that this can explain (in a mechanical sense) the secular decline in the correlation between concentration and profitability reported by Domowitz, Hubbard, and Petersen. But Schmalensee also finds that $\delta$ on average moves counter-cyclically in (5'), while Domowitz, Hubbard, and Petersen find the correlation between concentration and profitability to be pro-cyclical. This suggests that $\rho$, which is best interpreted here as the profit rate earned by small firms, is more strongly pro-cyclical in concentrated than in unconcentrated industries. This in turn suggests that the incidence of collusive behavior in concentrated industries, which can be expected to raise the profits of both small and large firms, may be pro-cyclical.

Finally, Dunne, Roberts, and Samuelson (1987) use a panel data set constructed by linking information gathered by the U.S. Census of Manufacturing at five-year intervals at the plant and firm levels. They are thus able to study a relatively complete record of entry and exit in U.S. manufacturing industries over time and to distinguish among entry by new firms, firms that build new plants, and firms that change the product mix of existing plants. Among their more interesting findings are that both entry
and exit are frequent events in most U.S. industries and that entry rates and exit rates are positively correlated in cross-section. It is hard to come away from their study with the impression that entry deterring behavior is important in many U.S. manufacturing industries.

6. Conclusions and Implications

Until game-theoretic analysis either begins to yield robust, unambiguous predictions or is replaced by a mode of theorizing that does so, any major substantive advances in industrial organization are likely to come from empirical research. And the recent increase in the volume of high-quality empirical studies makes it quite possible that industrial organization will once again become a field driven by facts rather than theories.

The examples discussed in Sections 4 and 5 indicate that well-designed empirical research can reveal a good deal about how market structure, conduct, and performance are shaped. But knowledge of many phenomena is clearly still thin in important respects. One of the few substantive conclusions that can be confidently asserted at this stage is that while market concentration may indeed have some impact on conduct and performance, it is much less important than Bain and some of his early followers seem to have believed. It seems unlikely that any very simple model of business behavior will prove empirically robust.

It is important to note that virtually all of the persuasive empirical studies discussed here share one important feature: they employ carefully-constructed data sets. Because few real data sets confess their secrets easily, advances in modeling techniques and econometric methods are important. But the main lesson that seems to emerge from recent
developments in empirical research in industrial organization is that the quality of the results obtained depends critically on the quality of the data employed. The key dimension of quality has more to do with experimental design than with measurement error; what matters most is whether the data contain information that can be used to identify the answers to questions of economic interest.

This is in some respects a discouraging conclusion. Economists, unlike historians or anthropologists, are formally trained only in the analysis of data sets, not in their construction. The economics profession does not much reward the tedious labor necessary to construct sound and interesting data sets. And data set construction is particularly difficult in industrial organization, and not only because accounting data are imperfect. Large modern firms are complex collections of business units operating in different markets, each with a whole array of tangible and intangible assets and employing a large number of workers and managers, linked by top executives and their staffs. Business firms typically produce large quantities of data to guide their own decisions, but they are almost universally reluctant to make much of this information public. The labor economist's task of describing workers and their behavior seem to pale beside the difficulties inherent in constructing data sets on firms and their decisions and operations.

The economics profession seems unlikely dramatically to change its collective attitude toward data collection any time soon. Thus progress in industrial organization may depend critically on the extent to which the construction of informative data sets is supported by government agencies and other sources of research financing.
REFERENCES


FOOTNOTES

1. Perhaps the clearest examples of adherents of the two Schools shouting past each other can be found in Goldschmid, Mann, and Weston (1974).

2. Tirole (1988) provides a superb general overview of the theoretical literature in industrial organization.

3. I thus resist the temptation to discuss the emerging literature on laboratory experiments in industrial organization; see Plott (1988) for a comprehensive survey.

4. This paragraph and the next have been heavily influenced by Fudenberg and Tirole (1987) and Milgrom and Roberts (1987).

5. An earlier study by Cox (1933) is much less informative - largely, I think, because Cox did not have available the antitrust trial record on which Tennant and Nicholls relied heavily.

6. This pattern of conduct persisted after World War II, and product innovation became an important form of non-price rivalry. See Schmalensee (1972, pp. 125-133) for a brief discussion of the early postwar period.

7. A number of the points made in the remainder of this section and in Section 5 are discussed in more detail and with more complete references to the literature in Schmalensee (1988).

8. This is basically why studies of the aggregate welfare costs of monopoly power that follow Harberger (1954) and base their analysis on differences in profit rates tend to find small costs.

9. This number was derived from U.S. Internal Revenue Service, Statistics of Income for 1938, the middle year in Bain's sample.
10. Tests of interactive specifications were performed by a few later authors (e.g., Caves, Porter, and Spence (1980) and Salinger (1984)) and produced generally negative results.

11. This development follows Schmalensee (1987a); see also Cowling and Waterson (1976) and Clarke, Davies, and Waterson (1984).

12. See Schmalensee (1987a) for a discussion of more complex cases.


14. See Bresnahan (1988) for a comprehensive and useful survey of this work.

15. Hall (1987) in fact works with a general neoclassical production function with multiple inputs, but his analysis is quite similar to that which follows.

16. See Caves (1988) for a comprehensive and stimulating survey of research of this general sort.

17. Porter (1979) argued that his results did not imply that Demsetz was right but were rather more consistent with the existence of important strategic groups within industries and with mobility barriers that prevented movement into the more profitable groups.