

# INDUSTRIAL ORGANIZATION: THE ROLE AND CONTRIBUTION OF ECONOMETRICS\*

By HENRY GRABOWSKI, *Yale University*  
and  
DENNIS MUELLER, *Cornell University*

Industrial organization traditionally has been one of the more empirically-oriented branches of economics. Many of the hypotheses peculiar to it have been inductively derived (e.g., the Berle and Means study [1]). Its literature contains a heavy dosage of case study and descriptive statistical material. Questions of the proper measurement of various phenomena, the quality of data, and index number problems also occupy a prominent place in the literature.

While the field has been somewhat slow to take up the use of econometric techniques in its empirical work, recent years have witnessed the appearance of a large number of studies on concentration, innovation, advertising, mergers, and other industrial organization subjects that have employed multiple regression analysis and similar econometric tools. This development is encouraging and suggests that the IO economist may some day be able to derive more general hypotheses and theories than have been possible up to this point. Unfortunately trends in mergers and firm diversification [2] [5] during the post-World War II period have created (or exposed) deficiencies both in the stock of theories that can be tested and in the data available for testing them. Indeed, these trends have been so broad and strong that they threaten to vitiate future efforts to apply econometric techniques to firm and industry data and force the IO empiricist back to a case-to-case approach—with the unit of investigation being the large, diversified firm. Hence, we stand in the danger of seeing the period of infancy in the application of econometrics to industrial organization coincide with its zenith, unless we are able to develop better theories and/or come up with better data than are presently available. The remainder of this paper is devoted to a discussion of these two problems.

## I. *Problems Associated with a Lack of Theory*

Industrial organization is intimately concerned with matters of economic policy, ultimately attempting to provide insights into questions such as the optimal forms of market structure on

growth and efficiency grounds and the like. While the field therefore has an empirical focus, research on these questions also requires a strong underlying theoretical structure if results applicable for policy are to be meaningfully evaluated. In its limiting extremes, traditional microeconomic theory offers a basis for formulating hypotheses on such issues. Certainly, one of the most powerful input-output generators of normative results in all economic inquiry is the microeconomist's model of perfect competition. With this model at the core and a good deal of partial equilibrium analysis available for other market structures, the empiricist can draw on a considerable body of results in formulating his speculative hypotheses.

Unfortunately, as one begins to pull at the various assumptions underlying the perfectly competitive model, one finds that it is possible to rationalize almost any kind of empirical relation. Consider the range of solutions that are normally possible under an oligopolistic market structure. Suppose, for example, one is interested in investigating how oligopoly affects a firm's allocations for activities such as advertising and R and D. It is possible on a priori theoretical grounds to have a whole range of contrasting solutions including the aggressive behavior of an extreme rivalry situation, the passive behavior of the Cournot solution, and the collusive behavior of a cooperative arrangement. Moreover, the empirical conditions under which one might expect to find any of these solutions are very nebulously formulated and tend to resist quantification. They consist of factors such as the amount of the uncertainty present, the goals of the firms, their attitude toward risk, and the historical pattern of the market's development. It is, therefore, difficult to predict in advance what might happen here solely on the basis of economic theory.

These problems have been compounded by the rise in importance of the large diversified firm. Microeconomic theory is basically a theory of the one-product firm. It has little to offer us by way of testable hypotheses of how multi-industry firms behave. Nor does it provide us with a well-accepted framework for analyzing the behavior of divisions of diversified companies. Yet the only sources of data we have come in these forms. Either we employ the census data for industries,

\* The authors are indebted to the National Science Foundation for financial support under grants GS-2677 and GS-2678 which greatly facilitated completion of this paper.

which are aggregates of the figures for single product firms, the major divisions of diversified firms based in the industry, and minor divisions of firms based outside of the industry, or we resort to data on firms and accept the aggregation over the many diverse products the typical corporation now produces. Indeed, there is a basic dilemma inherent in tailoring our current theories to accommodate these existing data sources; that is, any simplifying assumption we might choose to make to allow us to use one body of data, are likely to destroy the usefulness of the other.

For example, suppose we decide to assume that the diversified firm acts as a monolithic profit maximizing institution. This would imply that each of its subdivisions would act in a manner consistent with the maximization of profit for the whole enterprise. This might lead it to buy from other divisions of the parent firm, or from outside suppliers as part of a reciprocity strategy, at prices which do not maximize the division's profit. Similarly it might undertake R and D and advertising which are only profitable because of their strong complementarity with the activities of other divisions. Whatever the merits of this assumption with regard to the diversified firm, its use is obviously inconsistent with the assumption that the units of observation maximize profit when data by industries as gathered by the Census of Manufacturing are employed.

The sheer growth in the size of corporations has produced bureaucratic structures within the firm that make traditional models of entrepreneurially led or stockholder controlled firms seem frightfully obsolete. Economics and some of its sister disciplines have responded admirably to this challenge, generating a large number of interesting behavioral hypotheses based on managerial motives, organizational theory, etc. However, this proliferation of competing hypotheses only makes the empirically-oriented economist's job that much more difficult by expanding the range of plausible alternatives from which he must choose.

Most of these "managerial" models have retained the central postulate of economic rationality embodied in the maximization assumption and have merely attempted to recast it to new situations. Others have suggested, however, that a "satisficing" and rule of thumb decision-making approach is more fruitful for explaining managerial behavior in an uncertain environment. The implications of this motivational postulate for the empiricist are to fractionalize his world view completely and make his research efforts wholly inductive in nature. For this reason, most economists have chosen not to adopt this approach.

Yet anyone who has attempted to deal empirically with dynamic behavior under uncertainty will find some common bonds with this approach. For example, the stock reaction model prevalent in investment studies is at base an error learning model with close ties to the satisficer's rule of thumb decision making. Indeed, because traditional theory is basically static in nature, mainly for reasons of analytical tractability, investigation of dynamic behavior and lag structures must proceed very much on an inductive trial and error basis.

For a variety of reasons, therefore, one seldom has much beyond a few speculative presumptions when relating existing theory to the policy-oriented issues of industrial organization. What are the consequences of this lack of a general theoretical structure in our hypothesis formulations? There would seem to be at least two basic concerns. First, the lack of theory makes it very difficult to discriminate and select among the various hypotheses relevant to a particular case. A theory serves the useful function of implying certain events and ruling out others as inconsistent with its basic assumptions. Testing the consistency or inconsistency of a theory over a variety of empirical relations, then, allows one to make informed judgments on its general acceptability. If one observes a particular set of relations without a well-defined theoretical framework, however, it is difficult to discern subtle contradictions either internally or with past results and it thus becomes difficult to reject hypotheses which are not consistent with the data. Undoubtedly, this is one reason why opposing hypotheses tend to linger on with such vigor in the industrial organization literature.

A second problem, related to the one discussed above, concerns the stability over time and across various cross-sections of any relationships observed in this manner. If, for example, one observes a particular relation between, say, a market structure variable and a measure of industrial efficiency which ultimately rests on a more complex set of unknown and unobserved relations, then the observed relation is subject to unexpected changes even if the basic structure remains invariant. Unless one has a comprehensive theoretical framework for analyzing a particular structure, the generalization of results beyond the limiting confines of one's sample is not warranted.

Nevertheless, as Professor Mason reminds us, "no one who is other than eclectic, methodologically speaking, has any business in the field of industrial organization" [4]. Decisions of public policy resting on economic considerations cannot be postponed while our theories are being refined. In many policy areas, research therefore must

necessarily go forth as best it can on current analytical structures. Research of this nature must frequently be inductive in spirit and consider as many alternatives as possible over whatever data samples are available. In reporting results, the researcher in these areas should also set forth his basic research design and in particular indicate how other approaches compare with the main findings. Too often one sees only the "tip of an iceberg" of this type of empirical analysis, preceded by an elegant but artificial attempt to make the study appear deductive in nature. However, in such circumstances, it is much more instructive to know, at least in a general way, what is the explanatory power of other variables and functional specifications. Comparison of results with more "naive" models is also highly desirable and frequently ignored in these circumstances.

At the same time, empirical research in IO will most certainly benefit in the future as it has in the past from continued research efforts on a number of analytical frontiers. Given the significant current trends toward large size and diversification in our economy, a high priority certainly exists for theoretical studies investigating the firm in a more disaggregate fashion and ones delving into how the pieces of such organizations fit together to form rational decision-making entities. It would seem that only by gaining a greater knowledge of what is happening at the very "micro" level can we begin to make any real progress in understanding relations at the market structure level. While this kind of disaggregate approach is likely to generate theoretical models involving very complex sets of relations, simultaneous equation techniques are becoming increasingly available to help facilitate the use of these models in econometric applications. On the other hand, obtaining data at the intrafirm level to the degree required by such models is a serious problem as data sources now stand and one which will be intensively considered below.

## II. *Problems Associated with the Lack of Data*

Data problems are both universally acknowledged and ignored by economists. The individual scholar is at the mercy of the private firm or government bureau for most of his data, for he possesses neither the financial nor in some cases legal power to collect them himself. Even if he is astute enough to choose the best data available for his particular model, they are frequently inadequate for the task. Rather than abandon his model to the limbo of untested hypotheses, the economist typically "adjusts" the data and truncates the model to the point where an uncomfortable compromise between theory and data is

met. The results that emerge from all of this are rightly regarded as only tentative and at best indicative of the potential for further, more intensive research on the hypothesis. Unfortunately, such research is seldom undertaken, since data of higher quality than were used in the original study rarely appear.

The industrial organization economist probably has been no more guilty of committing these sins than his colleagues in other branches of economics. However, high levels of merger activity in recent years and the trend toward conglomerate organizational entities have rapidly increased the magnitude of data problems in this area.

Mergers can effectively destroy data in both cross-section and time series studies. In time series work a merger may change the characteristics of the firm sufficiently to preclude entirely the combining of the firm's data for before and after the merger. At a minimum it forces the analyst to give up one or more degrees of freedom by deleting an observation or adding dummy variables to the model. Mergers can create similar problems when cross-section data are used. If lagged variables are employed, the merging firm must be dropped from the sample in the year of the merger. Since the number of firms in the sample and their identities will thus vary over time, the researcher's ability to draw inferences by comparing the results for homogeneous cross-sections is reduced or eliminated. Even when no lagged variables are employed, the potential for comparing cross-sections over time may be impaired by the changing composition of the sample as a result of mergers.

The increasing extent of firm diversification across product and industry lines also creates serious data problems. The implicit assumption in cross-section studies is that all of the firms in the sample follow similar behavioral patterns. This assumption seems most plausible if all of the firms produce a single well-defined product and least appealing if they each manufacture a different array of products only a few of which may be produced by the other firms with which they are grouped.

The empiricist who employs Census of Manufacturing data on industries faces problems different in kind but similar in nature to those cited above. For example, in addition to the usual aggregation problems he now must contend with the fact that his data are aggregations of whole firms, divisions of firms, and parts of divisions of firms. If cross-section regressions are estimated, he must contend with the fact that parts of a single firm may be included in a number of industries and hence that the observations are not strictly

independent. And so it goes. These are problems with which we are all familiar, and yet all ignore.

Many researchers have tried to avoid some of these problems by simply dropping the troublesome observations (firms or industries) from their sample. While this may have been a reasonable procedure as long as these phenomena occurred with some rarity, this no longer is the case. At a minimum this procedure is now very costly in terms of lost degrees of freedom. More importantly, the remaining samples of nonmerging and undiversified firms or industries with only single-product firms can no longer be assumed to be in any sense randomly generated. Indeed, they are rather peculiar groups of enterprises that seem intent upon swimming against current economic trends.

### III. *Some Examples*

The extent to which these deficiencies in data combine to frustrate the attempts to test hypotheses is probably best illustrated by relating them to an area in which the authors have participated: the recent proliferation of studies testing the Schumpeterian hypothesis. The central question being investigated here is whether the level of innovative activity tends to increase more or less proportionately than size in a particular industry or line of industrial activity. Microeconomic theory provides conflicting strands of thought concerning the expected relationship here and, therefore, an empirical inquiry into the structural relations across various industry samples and time periods exemplifies the inductive approach suggested for such situations above.

The most desirable measure of a firm's inputs to innovational activity for such a study is data regarding the level and character of R and D expenditures, standardized by some uniform definitional criteria. The National Science Foundation has been collecting such data since 1957 for a very large sample of firms but they are able to provide information on it only aggregated over broad industrial classifications, roughly comparable to the SIC 2-digit level categories. These reported data are therefore not very useful in applications at the microeconomic level indicated above. Many firms publicly report their annual R and D expenditures, but usually say little concerning the nature of such expenses or how their definitions relate to the NSF criteria. The researcher in this area must therefore work through an array of questionable figures often surrounded by a strong sense of firm security and secrecy. Even if he is successful in obtaining R and D data privately from firms, he will have problems getting a complete or representative sample in any industry.

Indeed, for reasons of keeping the data confidential, he usually can identify the characteristics of his sample only by phrases like "ten large chemical firms" or the like.

The aggregate character of data, even at the corporate level, will provide obstacles in such a study. For example, the size variable normally available will be the firm sales or assets across all the industries in which it produces rather than the corresponding measure in its major industrial activity. Hence, unless diversification by firms into more or less research intensive industries activities is negligible, or if not, at least independent of size, this hybrid measure will influence the Schumpeterian relations estimated for each industry grouping. Where diversification is substantial, the resulting estimate may bear little relation to the true structural relation pertaining to a particular kind of industrial activity.

To help illustrate these points let us take a closer look at the firms in the chemical industry. *Moody's* for 1969 lists 183 firms which produce chemicals. Of these 183 firms, 101 can be classified as based outside of the industry, or in the case of the conglomerates, as having no real base industry at all. These range in scope from companies whose bases are in related chemical processing industries such as petroleum refining, rubber products, and drugs to companies whose base industry involves a distinctly nonchemical technology such as electronics, alarm systems, shoe machinery, and uranium mining. Obviously with over half of the firms producing chemicals being based outside of the industry, any assumptions regarding behavioral homogeneity for all chemical producing firms are quite hazardous. Even among the 82 firms that are based in the chemical industry there is a great range in the products and constellations of products they manufacture. Park Chemical, for example, is a fairly specialized manufacturer of metallurgical and automotive chemicals. Union Carbide also produces metallurgical and automotive chemicals. In addition, however, it produces a whole line of chemicals that includes agricultural pesticides, alcoholates, fluorocarbon propellants and refrigerants, fumigants, latexes for paints, lubricants, plasticizers, silicones, oils, resins, urethane foams, etc. Moreover, the 1966 *Fortune Plant and Product Directory* indicates that it produces products in no less than 22 different 3-digit industries outside of chemicals, and it has divisions manufacturing carbon products, consumer products, electronics, foods, gases and equipment, metals and minerals, nuclear products, plastics, and textile materials. Despite the chemical base they share in common there is probably a greater degree of homogeneity in the product mix of

Union Carbide and some of the large diversified petroleum and acknowledged conglomerate companies than there is in the product lines of Union Carbide and Park Chemical.

Nor are things much better if one limits one's sample to the dozen or so largest firms. While each large chemical company produces any single product in common with a number of the other large chemical firms, the array and rankings of firms differs from product to product. Hence, it is difficult to speak of a single leader for the chemical industry although DuPont is by far the largest. Nor is it possible to predict which firms are most likely to introduce innovations in a specific area or what and from whom reactions can be expected.

Given these facts, it would appear very hazardous to set forth an estimated structural relation on the Schumpeterian question based on a regression over aggregate corporate figures for the largest dozen or so chemically-based firms. As shown above, such a procedure ignores the significant contributions of nonchemically based firms and includes the nonchemical production of those based in the industry.<sup>1</sup> Indeed, the pointed discussions and debates over the validity of observed results which have turned on questions of regression specifications and proper deflating procedures and the like (for a survey here, see [3]) may very well be academic, given the weaknesses of the data bases common to all these studies.

#### IV. Some Recommendations

By focusing attention on current trends in concentration, the publication of Berle and Means's classic study in 1932 [1] helped launch systematic efforts to gather data on industry and aggregate concentration. In this we are very fortunate. The important questions concerning the absolute levels and trends in industrial concentration are among the best documented issues in the field.

If similar progress is to be made on the many other hypotheses that warrant testing, then data appropriate for testing them must be collected and made available to researchers in the field. Of

<sup>1</sup> Some indication of the extent of intra- and inter-industry coverage to be expected in a sample of large chemically-based firms is provided from Bureau of Census concentration ratio figures for 1963 [6]. For example, the percent of the value of shipments accounted for by the eight largest firms whose primary enterprise category code is 28A (basic chemicals, plastics, synthetics) is only 36 percent in organic chemicals, 25 percent in inorganic chemicals, and 30 percent in plastics [6, p. 289]. At the same time, this source indicates that for the 16 largest firms in 28A enterprise category, approximately 30 percent of value added and 40 percent of total employment is in a different enterprise category [6, p. 310].

particular importance here are data on firm operations on a fairly disaggregated basis. In seeking these data we have a potentially powerful ally in the stockholder. Both the stockholder and the economist have exhibited a heightened interest of late in finding out precisely what is taking place within the sprawling diversified corporation.

For this reason we would like to join in supporting the suggestion that a large data bank be established on a national level. In an age where great technological advances have been made in the data communication and processing field, one could certainly devise a system whereby such data could be made available for analysis, while at the same time respecting their confidential nature at the individual firm level. A computer system hookup from a national data bank to various academic centers which would involve central processing and the communication of results without any sight of certain confidential data inputs is an obvious development which is now feasible and highly desirable if the rational development of public policies is to be encouraged.

Barring the creation of a supply of data of this proportion, econometric work in industrial organization can be expected to move forward rather slowly. Indeed, a return in emphasis to the case study approach might even be in order—with the subject of each case study being a single large diversified firm. These organizations each possess a sufficient amount of economic importance to warrant such special treatment. And with the analytical tools and information currently available to us only a very thorough analysis of each giant firm can be expected to yield convincing conclusions. Certainly, the days when one can toss the total sales, profit, capital investment, etc., figures for G.M., 3M, Sperry Rand, and the other diversified giants into the electronic hopper and view the results with any confidence are numbered.

#### REFERENCES

1. A. A. BERLE AND G. C. MEANS, *The Modern Corporation and Private Property* (New York, 1932).
2. H. F. HOUGHTON, statement in *Economic Concentration*, Hearings before the Subcommittee on Antitrust and Monopoly of the Committee on the Judiciary (Washington, 1964), pp. 156-78.
3. J. W. MARKHAM, "Market Structure, Business Conduct and Innovation," *A.E.R.*, May, 1965, pp. 323-32.
4. E. S. MASON, *Economic Concentration and the Monopoly Problem* (Cambridge, Mass., 1957).
5. W. F. MUELLER, statement in *Economic Concentration*, *op. cit.*, pp. 502-18.
6. U. S. Bureau of the Census, *Concentration Ratios in Manufacturing Industry, 1963*, Part II (Washington, 1966).