Summary of the Papers

The Brookings Panel on Economic Activity has been meeting regularly since 1970 to present and critique papers on macroeconomic topics. It now meets twice a year and publishes its results in semiannual issues of *Brookings Papers on Economic Activity*. This publication has become the premier forum for applied macroconomic work and is one of the most widely read and cited journals in economics.

The new Brookings Center for Economic Progress and Employment is using the same approach for its special microeconomic issue of *BPEA*. Papers are prepared by economists in university economics departments and in business schools. The topics and participants, chosen by the editors in consultation with others in the center, differ from those of much university research. The center solicits research papers on the problems and policy issues that form the core of its concern—productivity, efficiency, and job structure. It interprets these topics broadly, to include, for example, applied research relevant to antitrust policy. The purpose of the special issue of *BPEA* is to encourage scholars working in relevant areas of research to focus on issues of importance to a more general readership and to policymakers, and to bring their work to a wider audience. The Brookings editors and the discussants at the meeting provide detailed comments to authors, stressing the importance of drawing out the implications of their work. The new microeconomic *BPEA* is being edited by Martin Neil Baily and Clifford Winston.

Three of the papers in this issue look at factors that affect the productivity of manufacturing plants. The first paper examines how changes in a plant's ownership affect its productivity. The second and third papers analyze data from the auto industry. One considers the productivity of differing arrangements of the work force, the other the speed of new product development.

The fourth paper reports the results of a survey that asked R&D

directors how they protect the intellectual property rights resulting from their companies' research. How do they prevent competitors from simply borrowing their findings? The fifth and sixth papers examine the organization of markets. One is an empirical investigation of the effect the size of markets has on the number of firms that are able to compete. The other is a theoretical analysis of the way in which sunk costs of production, especially the costs of technology development, influence competitive behavior. The paper includes a critique of the influential contestability doctrine.

One striking aspect of these papers is the depth of empirical work they contain. Five of the six analyze large bodies of data that have either been collected by the authors or prepared by them. In-depth empirical analysis and field research do not guarantee that the results obtained are correct, but the chances of capturing something solid and important are increased by such approaches. This kind of work is relatively unusual in the writing of economists.

A detailed summary of each of the papers follows.

Productivity and Changes in Plant Ownership

Since the early 1970s the number of plants and companies involved in changes of ownership has increased tremendously—between 1980 and 1986 their value increased by a factor of six. There is a debate about the effects of this trend, which some consider counterproductive. Company managers, it is argued, are becoming like real estate or stock exchange speculators gambling on capital gains; managers should instead concentrate on improving earnings through more efficient operations. Others consider takeovers a way to improve efficiency by putting plants in the hands of those who can best use them.

Economists have looked at the empirical evidence on the market values of companies and have found that stockholders gain, at least temporarily, from takeovers. Shortly after the announcement of a takeover, the market value of the combined company is greater than the sum of the market values of the individual companies before the announcement, although the gains tend to accrue to shareholders of the acquired firms. The key question, however, is whether these gains reflect real improvements in the efficiency of the companies' operations. Is there evidence of a turnaround or improvement in productivity as a result of the change?

Frank Lichtenberg and Donald Siegel approach this issue using data collected from 1972 to 1981 by the Census Bureau on more than 20,000 manufacturing plants owned by nearly 6,000 firms. One-fifth of the plants changed owners at least once during this period, so the data allow comparison of their productivity with the productivity of those that did not. The authors also compute each plant's productivity over time to see how it changes before and after acquisition by a new owner.

The plants in the sample form a comprehensive data base, accounting for two-thirds of total U.S. manufacturing shipments and over half of manufacturing employment. In certain respects, however, this data base does not represent all plants. Those that opened or closed during the sample period are excluded—an important limitation. And the plants in the sample are larger than the average for manufacturing; about 82 percent have 250 workers or more, compared with 4 percent for all manufacturing plants.

The authors argue that over time the economic performance of an important fraction of operating plants deteriorates, perhaps partly because economic conditions—the pace of technology, composition of demand, factor prices—change, and some managers are less able than others to adapt. This relative deterioration presents an opportunity to improve a plant's performance, and exploiting the opportunity often involves a change of ownership. In other words, some owners may have a comparative advantage and some a comparative disadvantage with respect to certain plants. The sources of advantage could be superior knowledge about a particular technology or better market access or lower-cost product distribution.

The authors argue that the quality of the match between an owner and a plant is an important determinant of productivity. A company with many plants and products may be able to operate some of them effectively but not others, which will cause it to look for a buyer for its weaker operations. And, of course, potential buyers may seek out plants whose operations they think can be improved.

In addition to improving the match, an ownership change can affect productivity in other ways. A plant that has been operating for some time has built up a network of relationships or implicit contracts with its workers, its suppliers, and the community in which it is located. If economic conditions necessitate changes, and the changes are hard for the existing owners to make, new owners may be more willing to break the implicit contracts and restructure the operations of the plant. Such restructuring could increase productivity and profitability, although the gains could cause workers to be laid off and suppliers to be discontinued, an obvious sacrifice for them and a potential long-term loss for the plant.

The measure of productivity the authors use is total factor productivity, which takes account of the contributions of labor, capital, and materials. TFP thus captures the efficiency with which these factors are used in producing output. Lichtenberg and Siegel find that plants with a low initial level of productivity show a significantly higher probability of a change of ownership. The results, therefore, are consistent with the idea that a plant's poor economic performance encourages current owners or potential buyers to seek a change. The authors also compare the average initial productivity of plants that subsequently changed owners with the productivity of plants that did not, after controlling for plant size. They find that the changers had 3.2 percent lower initial productivity than the nonchangers.

The authors then analyze how productivity changed in the years before and after an ownership change, relative to productivity in the nonchangers. The timing of ownership changes within the sample period was obviously different for different plants, but by pooling results Lichtenberg and Siegel were able to estimate the average pattern of productivity in the seven years before a change and the seven years after it. The results were striking. On average, relative productivity deteriorates in the years before an ownership change. In fact, this deterioration continues through the year following the change itself. In subsequent years, productivity climbs back to the point at which it is almost identical to the average for those plants with no ownership change.

These results produce a consistent picture of the effect of ownership changes, but the authors double-check the data by comparing the rates of TFP growth over time for changers and nonchangers. Recall that plants that are destined to experience an ownership change start with a lower level of productivity and then climb back toward the average. This should translate into slightly faster growth over the period for the plants that change owners, which is indeed the case. The changers had an annual productivity growth rate 0.6 percent higher than that of nonchangers.

Two important questions, however, have to be considered in assessing the validity of these results. The first concerns a phenomenon called regression toward the mean. An example of this occurs with heredity. The height of children depends on the height of their parents, but there is some randomness in this relationship. Very tall parents have tall children, but on average the children will be a little shorter than the parents. Their height tends to "regress" or move toward the mean or average height of all children. Similarly, although Lichtenberg and Siegel argue that ownership change improves low productivity, regression toward the mean predicts that plants with low productivity may move back toward the average even with no ownership change.

The second, and related, issue was the subject of much criticism at the conference. Plants that have low productivity and then get worse are likely to close altogether. But because the data Lichtenberg and Siegel use include only the survivors, we do not know the patterns of ownership change or of productivity change experienced by plants before they fail.

The authors point out, however, that neither the exclusion of plants that fail nor regression toward the mean can explain their results that plants destined to change owners have falling relative productivity in the seven years before the change and rising relative productivity after the change. Low initial productivity does not simply regress toward the mean; it begins to rise only after the change of ownership has taken place.

The authors have, then, found important links among low initial productivity, ownership change, and productivity improvement for plants that remain in operation. Ownership change can be an important way of improving the match between plants and managers, and hence improving overall productivity. But because data on plants that fail were not available for their study, part of the picture remains unknown. A change of ownership could lead either to a plant's being closed or to its being made more efficient. Additional research will have to determine how important the former effect may be and the extent to which plant closings improve overall productivity by weeding out the inefficient plants, painful though that may be.

Finally, the authors comment on the relationship between their results and those of David Ravenscraft and F. M. Scherer, who argue that mergers have an adverse effect on efficiency.[1] Ravenscraft and Scherer

1. David J. Ravenscraft and F. M. Scherer, *Mergers, Sell-Offs, and Economic Efficiency* (Brookings, 1987).

found that the wave of conglomerate mergers that took place in the 1960s and early 1970s involved companies that were highly profitable before a merger and became less profitable afterward. Lichtenberg and Siegel suggest that one way to reconcile their findings with those of Ravenscraft and Scherer is to consider the difference in the time periods covered in the two analyses. The wave of takeovers in the 1960s may have been unproductive, motivated by the belief on the part of some managers that they could create dynamic companies and boost share values by takeovers. In practice these mergers created mismatches and consequent poor productivity in the plants taken over. This led to sell-offs of assets in the 1970s to companies better able to operate them. The resulting improvements in performance were then picked up in the Lichtenberg-Siegel data for the late 1970s.

If this interpretation is correct, it suggests that not all changes in ownership are alike. Some may reflect an unproductive manipulation of assets. Some may improve efficiency. Any consideration of policies that will influence takeovers or sales of assets should evaluate which kind of ownership changes are being discouraged and which encouraged.

Labor Relations and Productivity in the Auto Industry

The automobile industry has traditionally been a symbol of American industrial strength, but in the past fifteen years U.S. auto companies have found it difficult to compete on the basis of cost and product quality against foreign rivals, particularly the Japanese. It has been asserted that American workers are less productive than foreign workers, but the experience of Japanese-run auto plants in the United States has been similar to that of plants in Japan. This suggests that ineffective management may be a more important source of the industry's problems than the allegedly poor quality of the work force. Can U.S. managements implement work practices that achieve high productivity?

The managements of the American auto companies have apparently accepted the proposition that shop floor industrial relations practices have contributed significantly to their competitive problems. The companies have been negotiating with the unions for changes in work rules, increases in the pace of work, and improved cooperation between management and workers. In return for the UAW's consent to such changes and to only moderate wage increases, General Motors and Ford have agreed to provide greater job security.

Harry Katz, Thomas Kochan, and Jeffrey Keefe use data on performance within individual auto plants to detemine whether work practices really do make a difference to productivity and product quality. They investigate the effects of such practices as the number of job classifications or the initiation of team systems.

The authors have a unique data set collected from fifty-three plants of one of the major U.S. auto companies. The plants are in both the United States and Canada, and all are covered by collective bargaining agreements between the company and either the U.S. or Canadian auto workers' unions. Most are assembly plants, but they also include stamping, body fabrication, engine assembly, and auto parts plants.

The authors surveyed the chief industrial relations manager in each plant, asking questions on such matters as the amount of relief and idle time, the procedures used to allocate work hours, the number of job classifications, and the layoff and recall procedures. They also asked about the pace of work in the plant and the extent of managerial discretion in setting work practices, and they identified those plants using a teamwork approach. A second group of survey questions assessed the extent of worker-management cooperation in the plant through such measures as grievance and absentee rates. The authors also assessed the state of the labor market in the areas where a plant was situated to see how anxious workers might be to keep their jobs. A third group of questions considered worker and union participation in decisions made in the plant and the extent to which information about costs and quality was provided by management.

With responses to so many different questions given in so many different forms, the authors could not use statistical analysis on the raw data. Instead they used principal components analysis to consolidate the responses into three indexes. These measured the extent to which a plant used teamwork, the degree of management discretion over operating procedures, and the extent to which workers were involved in implementing new technology.

To determine the effects of the differences in labor relations structures, indicators of the economic performance of each plant were required, and three were available. The first was a corporate quality index used by the plants themselves and derived from a count of the number of faults found by inspections of each plant's products. The second was a productivity measure maintained by the assembly plants, which captured the number of production-worker hours per vehicle. The company being studied also generated an adjusted labor productivity index for each plant that allowed for differences in the complexity of the cars being assembled. The third measure of plant efficiency was the number of supervisors per production worker. While not a direct measure of productivity, it did provide a useful supplement free from any biases introduced by accounting practices.

The authors find that the plants in the sample vary considerably in their labor relations practices and their productivity. Since all plants are part of the same company, this finding raises an obvious question: Why has the company not decided on a best practice and applied it to all plants? The authors argue that because of the histories of collective bargaining and labor relations in each plant, the company is not free to impose a uniform approach. Some plants have much better relationships with their workers than others, and it is difficult to rectify the problems in the bad cases.

Some of the divergences among plants also result from experimentation. Because the company has not settled on a best practice, it is trying alternatives. Some managers but not others have experimented with team systems.

Looking for statistical relationships among performance measures and labor relations characteristics, the authors come up with a surprising result. All else constant, the use of team systems of assembly actually worsens performance, increasing assembly time by seven and a half hours and requiring more supervisors per production worker. The authors were concerned that this result might have come about because some factor had not been held constant. For example, did plants with poor productivity switch to teams in an attempt to improve productivity? Apparently not. Data were available both for 1979 and 1986, and the plants with high productivity in 1979 were actually more likely to have tried team approaches in 1986.

The authors do find that greater managerial discretion over work rules and more worker and union participation in decisionmaking are both associated with better performance. Neither factor turns out to have effects that are measurable with great precision, but the data suggest that the effects may be substantial. The state of the labor market outside the plant gates, however, has little effect on performance. A high unemployment rate in the local labor market does not seem to encourage harder work among autoworkers. Finally, the authors consider whether plants with high rates of absenteeism did poorly, and find that they did. And they look at high grievance rates to see if these signal problems—apparently not. When workers stay home, it signifies that a plant is in trouble; if they try to resolve grievances, the plant is probably not.

The authors focus on the use of teams because many Japanese assembly plants use them, and previous studies have found that Japanese plants have higher productivity on average than U.S. plants. When the authors examined this conclusion using data collected by John Krafcik, the result was confirmed. In Krafcik's data Japanese plants used 20 percent fewer hours to assemble a vehicle than U.S. plants used. European plants were much less productive.

The Japanese have used team systems to achieve higher productivity in the United States as well as in Japan. The NUMMI joint venture between GM and Toyota comes close to matching the high productivity of its sister plant in Takaoka, Japan. This suggests that the apparent failure of teams in U.S. plants does not stem from an inability of unionized U.S. workers to produce under this system.

The results of the study challenge conventional views and are a puzzle that U.S. managers must solve. Clearly the quality of labor relations makes a difference to performance, and teams can work well. But this study shows that simply superimposing the team approach on the existing patterns of labor relations and management in the United States does not pay off.

George Eads, the chief economist at General Motors, discussed this paper at the conference and took issue with the authors' statistical finding that teams actually harm productivity. He pointed out that the authors' sample is basically restricted to assembly plants, and he questioned how well they were able to identify those plants that used a team approach. He said that an internal GM study has confirmed Krafcik's finding that the use of teams has been successful in the NUMMI plant, and he thought this might be a better indicator of the potential gains from using team systems. Despite these criticisms, however, Eads and other participants in the conference agreed with the essential conclusion of the paper: the introduction of team systems by themselves is likely to accomplish very little.

Product Development Efficiency in the Auto Industry

Concerns about the efficiency and productivity of the U.S. auto industry go beyond problems on the shop floor. In 1986, R&D investment in the U.S. industry was $6.25 billion (about 8 percent of sales), most of it devoted to product development. Because product development costs are a significant part of total costs, the efficiency and productivity of R&D are important to the competitive position of the industry.

Auto companies have a tremendous stake in the success of their development efforts. A new model that increases a company's market share can generate a major increase in profits. A product that is efficient to manufacture can reduce production costs and increase profits even more.

But when a new product takes years to develop, its failure in the market can spell disaster for a company. If the company can replace the unsuccessful product quickly, the damage might be limited. This ability to develop new products rapidly to respond to changing market conditions varies significantly across companies.

To examine the product development process and the variations among manufacturers, Kim Clark, Bruce Chew, and Takahiro Fujimoto collected data on twenty-nine passenger vehicle development projects in twenty auto companies in Japan, Europe, and the United States in the 1980s. Six of the projects were in the United States, twelve in Japan, and eleven in Europe. All the projects involved models in which more than half the parts were newly designed, and some projects entailed creating entirely new models. The vehicles included sedans, micro-mini cars, and small vans, ranging in price from $4,300 to $40,000. In extensive field research the authors used in-depth interviews and questionnaires with key members of each project, including project managers, heads of R&D groups, and product and process engineers. The mass of quantitative information was supplemented by more general interview information, particularly when anomalies or inconsistencies turned up.

The new models often required innovation in components and systems, but this did not involve research. Thus the technical feasibility of the process was well established, but the exact nature of the product and its technical performance was uncertain. The authors consider the acquisition of knowledge about the product—about what kinds of cars are likely to be well received in the market and what buyers are looking for—as an essential part of the development process. Firms must also learn about the opportunities for technological advance in materials and design. Such knowledge is obtained through a series of problem-solving cycles, each one carried through by a team of engineers. A typical development project has four main cycles: concept generation, product planning, product engineering, and production engineering. The efficiency with which these cycles are completed goes a long way toward explaining a company's success.

The number of hours required to complete the tasks that make up each of the project's cycles can be totaled to measure the engineering labor input, the first of the author's key performance measures. The lead time needed to complete a project is the second principal measure. This measure does not bear a simple relation to total labor input because of the way the tasks are arranged and coordinated. Some can be done simultaneously, some can partially overlap, and some must be done sequentially. And for any development project, certain tasks form what is called the critical path. Because these tasks limit the company's ability to reduce development time, any delay in accomplishing them will delay the overall project. Other tasks have slack time, so that delays in completing them will not delay the entire project—up to a point, of course.

To explain what they find in specific development projects, the authors look particularly at two elements that determine total engineering hours and the number of months of lead time. First, complex projects requiring more innovation or resulting in larger and more sophisticated new cars take more hours of labor and more time. Second, the way a project is organized—the degree to which tasks are subdivided, the skill with which they are coordinated, the effectiveness of problem solving—can increase or decrease this time.

The data on total engineering hours and lead time revealed substantial differences between the Japanese projects on one hand and the U.S. and European projects on the other. A Japanese project averaged 1.2 million engineering hours and a lead time of 43 months, compared with 3.5 million hours and a lead time of 62 months for a U.S. project. The European projects averaged 3.6 million hours and 63 months. But the cars developed in these projects differed in complexity, in the fraction of parts developed by suppliers, and in the extent to which parts from previous models were carried over. Much of the remainder of the study's analysis is directed at taking account of these differences so that the basic sources of the Japanese advantage can be isolated.

To adjust for the extent that carryover parts and those developed by suppliers influence differences in engineering hours required, the authors estimate the fraction of the total parts developed in-house. For example, if only half the parts in a particular project were developed in-house, then reported total engineering hours are doubled. On average, the authors find that the Japanese depend on supplier-designed parts more heavily than others. Adjusting for this dependence reduces the gap in total engineering hours between Japan and the United States and Europe but does not come close to eliminating it. The authors then adjust for differences in the complexity of the projects. A car that is larger, has more body types, and has more power or better performance requires a larger development project. Controlling for car size, number of body types, and price (an important indicator of performance and quality), the authors use a regression to determine the effect of project complexity on the number of engineering hours required.

Once these adjustments have been made, the Japanese advantage in engineering hours is reduced, but a base-line project still takes twice as many engineering hours in Europe and the United States as it does in Japan.

The authors do a somewhat similar analysis of the differences in project lead time. They estimate the effects on lead time of the complexity of the project and of using parts designed by suppliers. Before adjustment, the average lead time in Japan is 20 months less than that in the other regions. After adjustment the Japanese advantage drops but is still 12.5 months, about 20 percent of the total needed in the United States and Europe.

What then is it that gives the Japanese their advantage? To answer this question Clark, Chew, and Fujimoto classify the projects into three organizational types. One they describe as a "functional structure," under which development is organized into functional departments (for example, body engineering) and activities are coordinated by rules and traditional patterns, with no overall project manager. A second type, characterized by a "lightweight manager," has departments as before, but a project manager coordinates activities, although he has little influence over the content of each department's projects and has little status in the organization. A third type uses a "heavyweight manager," one who not only coordinates the project but also acts as a champion for the concept being developed. Heavyweight managers have direct responsibility for all aspects of their projects, work directly with the engineers to create project teams, and have high status within the organization.

The authors find that many projects in Europe and one in the United States were of the functional type, which indeed has been historically the norm in Europe. The lightweight structure characterized more than two-thirds of the projects. Only four companies, however, used the heavyweight project manager; none was in the United States. When the performance of the projects was compared, the heavyweight structure tended to require fewer hours and less lead time than the lightweight structure. In turn the lightweight structure usually performed better than the functional structure. There is, however, an important reservation about these results. The effect of the different types of organization could not be estimated precisely. Some lightweight projects worked well. Some heavyweight projects did not. Choosing a particular form of organization does not guarantee success—a result that echoes the findings of Katz, Kochan, and Keefe.

Clark, Chew, and Fujimoto also consider the extent to which the tasks were specialized. Some specialization certainly is needed—an engineer or group must focus on a particular part of the project. But too much specialization can make coordination difficult and encourage duplication of effort. One measure of specialization is the number of people involved in a particular project, and by this measure there is less specialization in Japan and less in projects with heavyweight structures. The authors find a correlation between project performance and specialization. Projects that use a lot of total hours and take a long time have more people involved. The authors interpret this as a manifestation of Parkinson's law: if lots of people are put on a project, they find things to do. Some participants in the conference worried, however, that a correlation like this is almost tautological.

Finally, the authors study the way information was conveyed from one project group to another. They distinguished between "batch" communication, in which a group working on an early stage of the project waits until its activity is completed and then passes on the results in a batch to other groups, and "dialogue" exchanges, in which the upstream and downstream groups communicate continually and use feedback in both directions to guide their work. The method of communication is especially important when activities overlap, that is, when downstream activities start before upstream tasks are completed. Overlap can shorten lead time but can prove costly with only batch transfer of information.

The authors find that Japanese projects regularly overlap tasks within a project, U.S. projects sometimes overlap, and European projects do so rarely. There are also differences in the ways groups communicate. The Japanese use informal methods of communication and frequent interactions among groups. U.S. companies rely on more formal interactions, typically with two stages of information release—a semifinal design and a final design. Europeans are in the middle, using informal communication within a task group but formal communications among groups. The authors find problems with the U.S. structure because, although it uses overlapping, there is often not enough communication to make it efficient. Europeans do not overlap and hence need less communication.

The authors point out that in theory there should be a trade-off between lead time and engineering hours. If there were efficient alternative ways of carrying out a development task, companies could choose more rapid development, which requires more labor hours, or slower development, which requires fewer resources. But the data did not show evidence of this trade-off: the projects with fewer hours also had shorter lead times. The authors suggest that projects lacking strong leadership and efficient methods of communication run into problems that then require more hours of work and delay completion.

The conclusion, therefore, is that the Japanese advantage in fewer total hours and shorter lead times results from differences in project organization. Strong managers, good communications, and the avoidance of excessive specialization improve performance.

To test the robustness of this conclusion, the authors consider alternative explanations of differences in performance. One possibility is that the Japanese use more automated design technology, particularly computer-aided methods, to save on engineering labor. This does not, however, seem to be the case: U.S. companies also use computer-aided techniques extensively. A second possibility is that in the United States the final designs consist of autos of higher quality that are easier to manufacture than the Japanese designs. Again, studies of production productivity and product quality do not support this possibility. Participants in the conference argued that longer U.S. lead times had been the result of radical product redesigns forced by adjustment to the energy crisis of the 1970s. The authors acknowledge this possibility, but point out that the Europeans, who also have longer lead times than the Japanese, did not have to downsize their cars and that much of the shift to smaller cars in the United States took place before the 1980s.

The authors conclude by speculating that the industries use different approaches because of their varied histories. The Japanese companies traditionally lacked engineers and were forced to adopt the kind of "skunk works" approach used in wartime projects here.

Participants in the conference raised concerns about the authors' methodology. John Meyer pointed out that using the price of an auto to adjust for quality is difficult when some of the cars are not sold in the U.S. market. Other participants noted that because of the international character of auto companies, comparing engineering hours and lead times by country may make no sense. F. M. Scherer wondered why strong managers were not generally used, given that, in his opinion, the advantages of this structure have been well known for twenty-five years. He speculated that perhaps development projects are not forced to meet a market test. He also suggested that the quality of the engineers involved may have been a problem, noting that military R&D pulled engineers away from civilian work during the 1980s.

Appropriating the Returns from Industrial R&D

Technological change is a primary source of productivity growth for the economy, and strength in the development of new technology is an important element in U.S. competitiveness. To encourage private investment in R&D and to generate technological change, companies must receive an adequate return on their investment in innovation. But appropriating returns may become a problem if other companies can borrow or copy new developments in products or processes.

The inability to appropriate returns from research may result in two kinds of social losses. First, too little R&D will be undertaken. Second, R&D will be misallocated as firms concentrate on the type of innovation for which returns are appropriable, even though these may provide fewer social benefits than other projects.

Mechanisms that increase appropriability provide more incentives for R&D, but at a cost. Once a new product or process has been developed, for example, it would be efficient for it to be used throughout an industry, which suggests that knowledge should be made available to all. But efficient allocation of R&D resources suggests that innovators should be allowed to keep their knowledge to themselves. This fundamental conflict was recognized and a resolution of it attempted in the establishment of the patent system. In principle, patents reward an innovator by granting the exclusive right to the commercial use of the knowledge that has been generated, including the sale of licenses to encourage its widespread use. There is a time limit on this monopoly, however, because patents expire, so that benefits to society may be diffused as rapidly as possible.

In practice, patents do not solve the problems created by the special nature of technological knowledge as a public good. Competitors "invent around" patents, and patent restrictions may be difficult and costly to enforce. A study by Edwin Mansfield, Mark Schwartz, and Samuel Wagner found that the protection provided by patents varies among industries and is often ineffective.[2] With the exception of the pharmaceutical industry, the existence of a patent raises the costs of imitation by only 7 to 10 percent.

Richard Levin, Alvin Klevorick, Richard Nelson, and Sidney Winter have launched a major project to assess the effectiveness of patents as a method of appropriating returns and to compare their effectiveness with that of other methods. Companies may, for instance, appropriate returns to R&D by keeping their technology secret, relying on the advantages of being the first to enter a market, and emphasizing marketing efforts to build brand loyalty.

The authors developed a questionnaire and administered it to high-level R&D managers in companies in more than a hundred industries. The managers, knowledgeable about both the technology and market conditions in their industries, were asked not to describe their own companies but to respond as representatives of the line of business in

2. Edwin Mansfield, Mark Schwartz, and Samuel Wagner, "Imitation Costs and Patents: An Empirical Study," *Economic Journal,* vol. 91 (December 1981), pp. 907–18.

which they operated. The questionnaire was divided into four parts, the first two dealing with issues of appropriability and the second two with the respondents' perceptions of technological opportunities and advance. This study reports the results of the first two parts.

The authors obtained responses from 650 individuals representing 130 lines of business that were generally representative of business units performing R&D, except that companies without publicly traded securities were excluded. The authors found significant variability across industries in the responses, indicating the potential for analysis to explain why some industries are different from others, but they also found wide variance within each industry, indicating that random differences among respondents might make analysis more difficult.

In this paper the authors look first at the effectiveness of different methods of protecting new or improved processes and products. The most striking finding is that respondents rated patents the least effective means of protecting process innovations, ranking them behind secrecy, superior sales and service effort, learning and experience, and lead time. Patents are more important for product innovation, ranking ahead of secrecy, but they still trail the other methods. Industries that did consider product patents the most important method of protecting appropriation included pharmaceuticals, pesticides, and industrial organic chemicals. The authors suggest that patents are important in the chemical industry because new chemical products can be clearly defined. As a consequence, patents can be easily defended, while without them, imitation would be easy.

The authors find that the relative failure of the patent system implied by their results is generally mitigated by the existence of the alternative methods of protection. However, 34 of the 130 industries lacked any clearly effective means of protecting innovations, particularly process innovations. The industries were diverse, but there were several from the fabricated metals and machinery groups.

The authors then investigate why patents do not provide much protection. Sixty percent of the respondents reported that competitors can easily invent around a patent. In fact patents may encourage such activities because filing involves making public the nature of the innovation, although in general this was not found to be an important limit on their effectiveness. The survey did find a number of industries in which stringent requirements for obtaining a valid patent mean that an innovator cannot be assured of obtaining patent protection. Respondents did not consider the existence of compulsory licensing arrangements as an important limitation on the protection provided by a patent.

The responses reported so far are ones that focused on protecting the returns from an innovation. But the survey also asked respondents how companies go about getting information about new technology developed by others. On average, performing independent R&D was rated as the most effective means both for products and processes. Although this may suggest wasteful duplication of R&D, the authors note that such duplication may not be as important as it first appears, since independent research to check what rivals are doing may be a significant part of a company's overall research strategy anyway. Licensing of technology and reverse engineering—that is, disassembly of a product to discover how it was made—were also rated as important sources of information.

The survey also asked about the time and cost involved in duplicating different types of innovations. Respondents' estimates of what would be involved varied widely, which made it hard to identify industry differences. Despite the earlier findings on the comparative weakness of patents, the results did indicate that, on average, patents raise both the cost and time involved in duplication. This result, of course, conceals a wide range of outcomes. Some respondents reported that patents actually reduced the cost of imitation through the disclosure they entail. The authors compare their results on the effects of patents on the costs of imitation with the results of Mansfield, Schwartz, and Wagner and find them to be similar.

The results of the authors' survey have been used in earlier studies that have looked at the relationship between industry concentration and R&D. These studies found that, when measures of appropriability are ignored, R&D intensity is low in industries with many firms and high in industries with several large firms that control significant shares of the market. A similar relationship between concentration and the rate of innovation was also found. But when the measures of appropriability derived from the survey are included in the statistical analysis, the systematic relations between concentration and either R&D or the innovation rate become very much weaker. It looks as if R&D spending and the rate of innovation depend on the ability of companies to appropriate returns.

There is an important policy debate going on at present concerning the protection of intellectual property rights. Changes in the patent system to provide more protection for innovators have been proposed, and some have been enacted. The results of the authors' survey do suggest that patents are not providing much protection but also that other methods can be used and are effective. In weighing the costs and benefits of changing the rules on intellectual property, policymakers should keep in mind the full range of options available to innovators.

In his discussion of the paper Richard Gilbert questioned the validity of survey results in which R&D directors used an arbitrary seven-point scale to rank the importance of different items. And he also wondered if the sample was automatically self-selected. Firms that could not appropriate the returns from their R&D may have failed and so disappeared from the sample. Zvi Griliches, who worked with the data from the authors' survey, expressed concern that R&D directors in the same line of business answered the questions quite differently. Robert Litan pointed out that patents are often very important for small and beginning firms, which were not well represented in the sample.

Entry Thresholds in Local Markets

Economic efficiency is enhanced if there is an appropriate balance between the advantages of having many firms compete against each other and the advantages of exploiting the economies of large-scale production. Understanding how the number of firms in a given industry and locality is determined and when this balance might be achieved is important in fostering efficiency. Timothy Bresnahan and Peter Reiss contribute to the theory and evidence on the determinants of industry structure by analyzing how and at what size small local markets can accommodate a second firm.

They begin by thinking of the size of a market as the size of its population. With their data, they can discover how large a market must be to sustain one firm. This size is called the entry threshold. The data also reveal how large a market must be to sustain two firms, which provides an estimate of the "duopoly" threshold. The authors then focus on the ratio of the two threshold values, a concept they call the entry threshold ratio. If the ratio is two, it will take twice the number of people to sustain two firms as it does to sustain one.

The authors argue that their ratio reflects the importance of so-called strategic behavior, which occurs when one firm sets its price or some other important variable to influence the behavior of other firms. In particular, a lone firm—a drugstore, for instance—in a town may not price its products at the level that would give it the most profit in the short run because it wants to discourage a rival from entering the market and so increase its profit in the long run. If strategic behavior is important, then the entry threshold ratio will be affected. If an established firm uses its pricing strategy to discourage a new firm, then a larger population will be needed to encourage the new firm—hence a larger entry threshold ratio.

To apply their framework to the data, the authors have to deal with some complications. The first is that market size is not simply synonymous with population. The per capita income of a town and the demographic characteristics of its inhabitants also determine market size. Second, the decision of a potential rival to enter will depend both on the size of the market and on whether it is growing or declining. A firm will be willing to enter sooner if it believes a market is getting bigger.

The authors' analysis focuses on retail and professional establishments, including auto dealers, drugstores, opticians, veterinarians, and physicians. An initial sample of 149 towns was selected, all from counties with fewer than 10,000 people but with distinct centers of population. The authors then used maps to determine whether the population center was isolated, that is, with no other town of population greater than 1,000 lying within 25 miles and no large town within 125 miles. By concentrating on small isolated markets, the authors could study the effect of market size without having to give very much consideration to the potential effects of competition from neighboring markets. A similar issue determined the choice of industries. Tobacconists, for example, are not a distinct industry because tobacco is sold in drugstores, gas stations, and so on. The authors thus selected retailers and professions whose products could not easily be sold through other outlets. The validity of the authors' claim to have identified distinct geographical and product markets was severely challenged during the discussion of their paper.

Data on the number of firms in a market were collected from American Business Lists, a company that processes telephone books. The lists include each firm that advertised itself in June 1987 as part of a yellow pages industry. The authors then checked the accuracy of the lists by analyzing telephone company classification schemes, comparing the yellow pages listing of auto dealers with auto companies' lists of franchises, and visiting some of the towns. They found the yellow pages lists both accurate and comprehensive. When there were duplicate listings—for example, a physician listed separately and again as part of a clinic—the duplications were eliminated. Professionals working together in a group practice were considered to be a single firm.

The authors classified each market by size and by the number of firms in each industry—none, one, or two or more. Markets and industries showed considerable variation. For example, 51 percent of the towns had two or more dentists and only 16 percent had none, but only 22 percent of towns had two or more movie theaters and 58 percent had none. The variables used to measure market size were taken from Census Bureau data. The authors added to the list of variables described earlier some specific ones for industries serving particular markets—the number of veterinarians and farm equipment dealers was related to such agricultural variables as heads of livestock, cooling contractors were related to the number of cooling degree days, and so on.

The authors first discuss professional services industries, finding considerable variation in the size of the market needed to support one firm: from 664 people for auto dealers to 1,886 for opticians. The entry threshold ratio would be 2.0 if the market needed to be twice as large to sustain two firms. In practice, only auto dealers (at 2.3) came close to this. The ratio was 4.3 for veterinarians; physicians, dentists, and pharmacists all required about 3 times the market size to support two firms. Turning to retail industries, the authors find that as few as 618 people can sustain a tire dealer, but almost 7,000 are required for a cooling repair service. The entry threshold value was actually below 2.0 for heating contractors, but as high as 3.0 for movie theaters and 3.1 for electrical contractors. Bakers, plumbers, and tire dealers all had ratios between 2.0 and 3.0.

The authors carried out various experiments to check the robustness of their findings. Taking account of the amount of livestock in a county, for example, dropped the threshold ratio for veterinarians. They also created a variable to measure the number of people who commuted outside a county to work (and who are thus more likely to shop or use professional services away from their hometowns) to see how well they had constructed a data set of isolated markets. In general they found that leakage of demand into adjacent areas was not a problem.

The authors point out that the professional industries generally had higher entry ratios than the retail industries and suggest that this results from the higher sunk costs faced by physicians and dentists than by beauty parlors and such. For example, professionals depend on their reputations and would find it costly to move to a new town; beauticians could move with lower cost. The authors also find that variations in government licensing arrangements do not seem to represent a plausible explanation of the differences they observe across industries.

The focus of the discussion on this paper was on the authors' ability to define clearly separated markets and identify all the firms in them. Several participants argued that in small towns there is often someone who does things like heating repairs on weekends without setting up a formal establishment with a yellow pages listing. Other participants questioned whether the authors' results allowed them to identify why entry threshold ratios vary by industries. Differences in the elasticities of demand in different industries may be important. And they noted that the entry threshold ratio in a given industry will also depend on the extent to which the second firm entering an industry has costs that differ from those of the first firm.

Sunk Costs and Competition

Joseph Stiglitz starts his essay by noting that economists typically analyze the benefits of a competitive market within a static framework in which technological change is excluded. That is a mistake, according to Stiglitz, who cites Josef Schumpeter's work in which competition is tied directly to the development of new products and processes. In such cases conventional models of price competition may be irrelevant. Instead, competitive behavior must be analyzed within a context in which firms incur sunk costs to develop products. Sunk costs mean that firms cannot easily enter and leave an industry and that the average costs of production for firms in the industry decline as output increases, since development costs can be spread over a larger production run when output is high. An industry in which incumbent firms have declining average costs will not sustain an efficient competitive market structure with many firms. If antitrust policy were to contribute to such a structure, it could compromise efficiency.

In recent years, however, "contestability" theory has been developed. It predicts that firms will set their prices at competitive levels even though there are a limited number of them or even a single firm in an industry. If, in some industry, average costs decline with output at all levels of production, then the cost-efficient way of meeting demand is to have only one firm. The contestability doctrine states that this single firm will operate at the highest level of output at which price just equals the average cost of production. And because the average cost of production includes a normal rate of profit on capital, the argument is that even a single firm will set its price at the level that eliminates any excess or monopoly profit. The level is competitive even with no actual competition because at any higher price some new firm (a potential entrant) would come in and contest the market. If the contestability doctrine is correct, an antitrust policy that forced the breakup of a monopoly firm or discouraged mergers is an industry that has only a few firms would reduce efficiency and would yield no benefit in terms of more competitive pricing.

Stiglitz's analysis addresses the validity of the contestability doctrine by reviewing a variety of different models to test the robustness of conclusions that potential competition is as effective as actual competition in ensuring efficient pricing.

The first model sets out a forceful critique of the doctrine's principal result. Contestability theory assumes that even as few as two firms are enough to engender cutthroat competition that will drive prices down to the competitive level. And if there is only one firm in an industry, it will be so afraid of the potential for such competition that it will set a competitive price to keep out entrants.

Stiglitz points out that if cutthroat competition really does take place with only two firms, then potential entrants to an industry would indeed be discouraged from entering, knowing what will happen if they do. But by the same token, if cutthroat competition is the norm, as contestability theorists assume, then the potential for it would protect a single existing firm against entrants who might contest the market. In particular, if there are even very small sunk costs required to develop or establish a product, no entrant will pay them only to get caught up in such devastating competition. In such a world, sunk costs thus provide an effective barrier against any entry of new firms. As a result, Stiglitz suggests that an established firm can actually charge a high price when it is alone in a market, provided that any firms that contemplate contesting the market must pay some initial fixed cost and that these entrants believe the established firm will immediately cut prices if they decide to enter.

Stiglitz explores the robustness of his argument by presenting a model in which firms have sunk costs of some limited durability, such as a plant that lasts for some given period of time. In principle, allowing for the limited lifetime of sunk costs might open the door to the possibility that new entrants may wish to contest a market. Suppose a single firm is established in a market and has its plant in place. At some point the plant is worn out and the firm has no advantage relative to potential new firms. Stiglitz argues, however, that there will be no actual competition in this case either. The established firm can anticipate the problem of possible new firms and build a replacement plant before the old one is worn out. It can use its investment decision as a strategic device to forestall entry. Thus a single firm can still charge a high price, make an excess profit, and protect itself against entry. Efficiency will suffer because capacity will be built too early. But consumers will not benefit through lower prices. Stiglitz drives home the role of sunk costs in discouraging entry when he asks how the above conclusion is affected if an investment has only a very short life. His results continue to apply.

Up to now Stiglitz has assumed that when there are as few as two firms in an industry, there will still be cutthroat price competition. He has allowed the assumption not because he is sure that actual markets operate this way—rather the contrary—but because contestability theorists have used the assumption in arguing that even a firm with no actual competitors will price competitively. Stiglitz has shown that while cutthroat competition makes the potential entry of new firms a very effective way to police pricing, such behavior also makes it very unattractive to be an entrant. Having developed this point, he turns to the case in which competitive behavior is less fierce. What happens if two firms in an industry work together to keep price high? He finds that an intuitive idea about the effect of potential competition may be false.

The lesson from traditional economic theory has been that the more competitors, the closer prices will be to competitive levels. This might lead one to think that the more potential competitors, the more effective the restraint on the pricing of established firms. Not so, says Stiglitz. Suppose a single established firm has been setting a high price. One potential entrant into this industry might anticipate that if it were to enter, it could divide the available profit with the existing firm. But if this entrant foresaw the possibility of other firms jumping into the market in the future, its own participation would seem less attractive. The strength of potential competition cannot be assessed in any very simple way from the number of potential entrants.

Advocates of the contestability model have responded to these criticisms, and Stiglitz next addresses some of their objections. One argument for the doctrine is that an established firm has advertised and settled its price and is locked into that decision; it cannot respond quickly to a new entrant who undercuts its price. This allows a potential competitor to come in and scoop some of the market, earning enough to offset the sunk costs of entry if they are not large. The disadvantage faced by an established firm if entry were to take place thus forces it to price competitively in the first place.

Stiglitz considers this objection weak. He argues that the costs to an established firm of adjusting its price and the time involved are likely to be small relative to a new entrant's costs to develop a new product or build production facilities. Other participants in the meeting shared this view and noted that it takes time for a new entrant in an industry to establish its price and product quality and then draw customers away from an established producer.

Contestability theorists also argue that a firm considering market entry can sign up potential customers to long-term contracts before incurring any of the costs of developing a new product or building production facilities. Examples of this have occurred when aircraft manufacturers and airlines have signed agreements based on an aircraft yet to be developed. This possibility also, the argument goes, imposes market discipline on established firms and forces competitive pricing even when there are costs of entering an industry.

Stiglitz argues that contracting cannot eliminate the problems created for entrants by the existence of sunk costs. A potential entrant would have to incur substantial development costs and have demonstrated its ability to produce a product well before customers would be willing to sign long-term contracts.

Participants in the conference agreed that contracting does not eliminate the problem of sunk costs, but the issue raised is clearly important. In markets with only a single firm, customers have the option of looking for alternative suppliers willing to enter the market with contractual guarantees. This will limit the market power of a monopolist, although clearly many industries are not suitable for this approach.

Finally, Stiglitz returns to the theme he began with, the idea that markets with changing technology need new analytic approaches. He points particularly to the role of R&D in providing protection against the entry of new firms. This is exactly the kind of sunk cost he has been discussing. An established firm can do a small amount of R&D early on and then maintain a credible R&D presence in the industry. This signals to any potential rivals that it will outcompete them by grabbing a technological lead if they decide to enter the industry.

In discussing the paper Sam Peltzman expressed some reservations about what had been learned from the results. He felt that the sensitivity of the contestability doctrine to the existence of small sunk costs was already known. And the trouble with much theoretical modeling in this area is that for every model proving one thing there is another proving the opposite. Empirical work in industrial organization has not benefited as one might have hoped from the theoretical firepower the discipline has provided. Stiglitz, supported by Timothy Bresnahan, responded that theoretical work is not intended necessarily to generate empirical predictions, but by presenting alternative models, theorists can establish a framework for sorting out what market conditions are more or less likely to lead to competitive or noncompetitive behavior. One can note also that while the weaknesses of the contestability doctrine may have been evident to some early on, the doctrine remains very influential in Washington. It is important for theorists to continue exploring its validity.

Future Issues

This microeconomic issue of the *Brookings Papers* has brought together economic research that is relevant and important to the problems of the American economy. The conference to be held in December 1988 promises to continue along this path. One of the ways economic policy has tried to foster efficiency is by reducing the adverse effects of regulation. Based on current plans for the December meeting, Paul Joskow will examine regulation in the electric power industry, Clifford Winston and Steven Morrison will look at deregulation in the airline industry, and Sam Peltzman, one of the leading analysts of the political economy of this issue, will assess where the United States stands on regulation and deregulation.

Three papers will address issues of technological change. Zvi Griliches and Bronwyn Hall will consider technological opportunities, Ariel Pakes will report on patents, and David Teece will provide a comparative study of industrial organization in the United States and Japan, asking whether technological change is fostered more easily by the Japanese structure. A final paper, by Lawrence Katz and Lawrence Summers, will consider the efficiency of the U.S. labor market, particularly in its response to trade competition.

My coeditor, Clifford Winston, and I have learned from the studies prepared for this issue of the *Brookings Papers,* and we expect the second issue to be equally informative. We hope that this summary provides an introduction to the ideas and results of the 1987 conference for the nonspecialist and encourages the specialist to read the studies. The work is worth studying.

MARTIN NEIL BAILY