Part V National systems of innovation

Preface to Part V

Richard R. Nelson
Columbia University, New York

The chapters in this section follow on naturally from those in Part IV. In Part IV the focus was on the evolution of particular technologies, and on how technical change was both shaped by and shaped firm behavior and industry structure. Here the focus is wider, being concerned with national systems of technical change spanning the full spectrum of industries. And the institutional concerns are broader.

The thrust of both Nelson's chapter on the United States and Freeman's on Japan is that modern national innovation systems are complex institutionally. While they involve the institutional actors and activities considered in the chapters of Part IV, they include as well institutions like universities dedicated to public technological knowledge, and government funds and programs. While these two chapters deal only with the United States and Japan, a description and analysis of the innovation systems of any of the other major industrialized nations would show a similar complex structure. Private-for-profit firms are the heart of all of these systems. They compete with each other, but they also cooperate. In all nations universities play an important role in the innovation system. And in all modern nations, public funds account for a significant fraction of total R & D spending, in some cases being nearly as great as total private funds.

There are certain essential similarities among the innovation systems in advanced industrial nations. There also are certain interesting and important differences. The contrast between the United States and Japan brings this out strikingly. A comparison of the two national systems is particularly interesting because throughout much of the post-war period the United States has been the clear technological leader in most fields, but over the last decade Japan has caught up in many, and now is ahead in a few, technologies.

As noted, in both nations there is vigorous technological competition among firms in the same industry. However, in Japan, under the auspices of MITI, there also has developed a tradition of inter-firm cooperation in certain kinds of research. In the United States, university–industry interaction has been close for a long time in many industries, and in some technologies universities have been an important source of invention. This seems to be much less the case in Japan, where the distance between industry and universities seems to be greater. In both nations, the government has played an important role, but the nature of the roles has been strikingly different. In the United States, defense and space R & D programs accounted for a large share of total industrial R & D during the 1960s, and spillover to civilian technology was considerable during that era in the fields of aircraft, semiconductors and computers. In recent years the evidence suggests that spillover has been much less substantial. In Japan, government industrial R & D support has been very small compared with the level in the United States, and in particular there is very little defense R & D spending. On the other hand, principally through MITI, the Japanese government has played a significant role in trying to direct and orchestrate the Japanese industrial R & D effort in certain key industries. Much of the current public policy debate in the United States, in Europe and in Japan is concerned with the efficacy of, on the one hand, significant government R & D support for the advancement of particular technologies, in the style of the US Department of Defense, and, on the other, MITI-like coordination of industrial research.

The chapter by Lundvall takes a different cut at analyzing national systems of innovation. His focus is on user–producer interactions, which he argues is an important if often overlooked feature of the innovation process. He argues plausibly that geographical and cultural closeness facilitate effective interaction, and goes on to propose that national borders tend to enclose networks of technological interaction which define national innovation systems. He puts forth several reasons why there are such national systems—common government, as well as common heritage and education (at least in the relatively homogeneous Nordic countries), and obstructions to cross-national flow of labor being prominent on his list. The Nelson and Freeman chapters simply assume that there are national systems, and that borders matter. Lundvall presents a theory as to why this might be the case. Actually, rather little is known about just how borders affect the flow of technological information and capabilities, which is what many governments are concerned about, or the patterns of interaction between upstream and downstream firms, which is Lundvall's focus, or university industry connections, which is another interesting topic. It seems clear that borders matter, but not clear how much, or in what ways.

The chapter by Pelikan is concerned with a theoretical exploration of whether a capitalist innovation system can be out-performed by a socialist one, where by the latter he means one in which officials appointed by a central authority control the use and creation of technology. He argues that it cannot, because of the likelihood that incompetent people will control the process in socialist regimes, whereas under capitalism pluralism and competition tend to assure good management, or at least that there are ports of entry with good management. While Pelikan does not consider in any detail the inefficiencies of the pluralistic capitalist system that have been stressed in several places in this book, or the large potential gains from some centralization in some cases, his stress on the dangers of centralization echo some observations in Nelson's chapter. Of course, one could look at the real question as being not about the relative merits of a fully decentralized versus a fully managed system, but about how much centralization or decentralization is appropriate for what kind of work. This is implicitly the way the matter is being treated nowadays in so-called capitalist economies. And many socialist economies are experimenting with a certain amount of decentralization and even competition. The results of these experiments will be interesting to watch.

15 Institutions supporting technical change in the United States*

Richard R. Nelson
Columbia University, New York

Economists, from Marx to Schumpeter, have touted capitalism as an engine of technical progress. But what kind of an engine is it? What are its key components? How does it work?

This chapter is a preliminary report on my theoretical and empirical research on innovation systems in capitalist economies. That research has been powerfully shaped by the understandings about the nature of technical progress, and perceptions of the role and organization of corporate R & D, discussed earlier. The emphasis here is on mapping and trying to comprehend the wider institutional structures within which corporate R & D and technical change proceed in capitalist economies. If technical change is far more complicated and variegated than it is depicted in standard economic theory, so too are the institutional structures supporting it.

The focus here is on the contemporary US scene. The following chapter is concerned with modern Japan. There are many similarities between the two national innovation systems, but some important differences as well. They will be initially signalled here, and developed at greater length in the following chapter.

Also, one of the striking features of national innovation systems is that they change over time, usually gradually, but sometimes sharply. The occasional large changes in innovation systems are part and parcel of the occasional sea changes in technical–economic paradigms discussed earlier by Freeman and Perez. While I shall sketch where the current US system came from, and how it is different from what it used to be and why, space precludes giving that subject adequate treatment here. And while I will speculate a bit about where it is going, how the new technologies discussed in several places in this book will drive it, I cannot here consider that matter at any length.

The complex capitalist innovation system

Particularly in comparison with Soviet-style systems, there are three rather obvious characteristics of national innovation systems in capitalist

*The research on which this chapter is based has been funded partly by PRA-NSF, and partly by the American Enterprise Institute.

economies. One is the privatization of much of new technology, which harnesses profit incentives and market forces to its creation. A second striking feature is the existence of multiple, independent, generally rivalrous sources of new technology. A third, and related, characteristic is heavy reliance on *ex post* market forces to select on the innovations offered by different firms, and on the firms themselves.

Put tersely, in capitalist countries, technical change is set up as an evolutionary process. The research on technical change, discussed earlier in this book, increasingly is explicitly analyzing it as such.[1] The fact that the process is that way in good part reflects the particular institutional structure of capitalist countries. Technical change proceeds in quite different ways in Soviet-style economies.

Evolutionary processes in general, and technical change in particular, are inherently wasteful, at least with the vision of hindsight. Due to the unplanned, uncoordinated nature of industrial R & D in capitalist economies, R & D allocation is doomed to be inefficient, compared with any kind of an ideal. Looking backward one can see a litter of failed or duplicative endeavors that probably never would have been undertaken had there been effective overall planning and coordination. Economies of scale and scope that might be achieved through R & D coordination are missed. Certain kinds of R & D that would have high social value simply are not done. Also, because technology is to a considerable extent proprietary, one can see many enterprises operating inefficiently, even failing, sometimes at considerable social cost, for want of access to the best technology.

It is something of a puzzle, therefore, why the capitalist innovation system has performed so well. There certainly is nothing like the twin theorems around to support an argument that capitalism 'can't be beat'. But, of course, the key question is: what are the alternatives? Compared with what? Various socialist scholars have observed the wastefulness of capitalism and proposed that a centrally planned and coordinated system, which treated technology as a public good, ought to be able to do better at generating and using new technology. The troubles socialist economies have been having with their innovation systems suggests that this is easier said than done.

What is it about technical change that makes effective central planning so difficult, or perhaps impossible? The basic matter, I would argue, is the uncertainty that almost always surrounds the question, where should R & D resources be allocated, in a field where technology is fluid? There generally are a wide variety of ways in which existing technology could be improved, and several alternative paths toward achieving any of these—not simply uncertainty about where the bets ought to be laid, but also disagreement among exprts. Under such circumstances, attempts to get *ex ante* consensus are likely to be futile or counter-productive. What the capitalist innovation system provides is multiple sources of initiative, and a competition among those who place their bets on different ideas. And it does so in a context where, as I shall elaborate later, there is widespread access to the basic generic knowledge one needs to understand the technical possibilities, and strong incentives to heed market signals. It is left to the market to decide, *ex post*, what were the good ideas. This is an inefficient and wasteful way to do things, and painful for the losers, but, given the nature of technological uncertainties and the way humans and organizations seem to think and behave, it may be hard to do much better than to set up technical change as an evolutionary process.

Also, there is another feature of capitalist innovation systems that partially mitigates the problems discussed above. In capitalist economies technology, or aspects of it, is partly a public good. These public and private faces of technology both complement each other and are at odds. The public aspect of technology helps to control the inefficiencies associated with the private rivalrous aspect, but at the same time builds other problems into the system.

Schumpeter's perceptive analysis about how new technology gets generated and spread in capitalist countries clearly recognizes both the private and public sides.[2] He saw the lure and reward for innovation in the quasi-rents on the private temporary monopoly associated with the introduction of a new product or process. However, in Schumpeter's analysis the monopoly normally is limited. Sooner or later competitors will be able to imitate, or invent around, the initial innovation. The fact that technology ultimately goes public has three benefits.[3] First, this assures that a healthy share of the benefits of innovation go to users, and that 'triangle' costs are kept down. Second, knowledge of the new innovation provides a base and a spur for further innovation by others. Third, by facilitating subsequent competition, the dangers that a company can build a wide and durable industry monopoly out of a particular innovation are kept under control. However, all of these public benefits will come to naught if fear of rapid imitation damps incentive to innovate in the first place.

From one point of view, the job of institutional design is to get an appropriate balance of the private and public aspects of technology, enough private incentive to spur innovation, and enough publicness to facilitate wide use. Access to an innovation by competitors should not be so rapid or complete as to dull incentives to innovate. On the other hand, the stronger the restrictions on access and the longer their lasting, the higher the social costs in terms of less than optimal use.[4]

However, while this simple view of trade-offs is illuminating, it is too simple. It represses that different kinds, or aspects, of technology differ significantly in terms of latent publicness. Also, the institutional structure of real capitalist systems is much richer than that depicted in simple models.

It is important to distinguish between two different aspects of a technology. On the one hand, a technology consists of a body of generic knowledge, in the form of generalizations about how things work, key variables influencing performance, the nature of currently binding constraints and approaches to pushing these back, widely applicable problem-solving

heuristics, etc. Dosi (1982) has called these packages of generic knowledge 'technological paradigms'. On the other hand, a technology also comprises a collection of specific ways of doing things, or artefacts, which are known to be effective in achieving their ends if performed or used with reasonable skill in the appropriate context. Much, if not all, of the generic knowledge tends to have properties of a latent public good. Such knowledge tends to be widely applicable, and germane to a variety of users. Access to generic knowledge may be essential if one hopes to advance further the technology with any force. Also, in a system where there is considerable inter-firm mobility of scientists and engineers, generic knowledge is very difficult to keep proprietary. On the other hand, while portions of the set of extant techniques possess latent public good properties, in the sense that certain techniques are widely applicable, a good part of it is not appropriately so characterized. As Keith Pavitt (1984) has stressed, a good portion of technique is of rather narrow application, being tailored to the attributes of the products and processes of particular firms. Thus restriction of access entails little cost.

Also, there is much more to the capitalist innovation system than for-profit firms in rivalrous competition. There are, as well, a variety of mechanisms through which firms share technological knowledge, and cooperate on certain kinds of R & D. There are universities in it, and professional societies. There are public monies, as well as private funds.

Once institutional richness is recognized, along with different aspects and kinds of technological knowledge, the simple 'trade-off' view of the matter needs to be supplemented by another one. One can see the task of institutional design as somehow to get the best of both worlds. Establish and preserve property rights, at least to some degree, where profit incentives are effective in stimulating action, and where the costs of keeping knowledge private are not high. Share knowledge where it is of high cost not to do so, and the cost in terms of diminished incentives is small. Do the work cooperatively, or fund it publicly, and make public those aspects of technology where the advantages of open access are greatest, or where proprietary claims are most difficult to police. Put another way, the design problem involves institution creation and task assignment at least as much as simple trade-offs taking institutional structure as given.[5]

This chapter is organised as follows. In the next section I describe the proprietary parts of the capitalist system, with particular emphasis on the means through which firms appropriate returns to their investments in innovation, and consider some consequences of the uneven effectiveness of these means across industries, and across different kinds of innovation. The subsequent section is concerned with technology sharing, and R & D cooperation among firms. Universtities, a very important part of the institutional structure, are considered in the following section. Next is a treatment of government programs in support of R & D. In the concluding section I pull these strands together.

Proprietary technology: mechanisms, domains and consequences

R & D carried out by business firms, in rivalry with other firms in the same industry and looking to innovation to get ahead or stay up, is the heart of the modern capitalist engine. Christopher Freeman (1982), Nathan Rosenberg (1985), David Mowery (1984) and others have told the story of the growth of employment of scientists by industry, and the rise of industrial research laboratories. Such laboratories proved profitable for the firms because they served to link the increasingly powerful generic knowledge and methods of modern science to the problems and opportunities of industrial technology, in an environment where both kinds of knowledge could be brought to bear on project selection and execution. However, for a modern industrial research laboratory to be profitable for a firm investing in one, it was not sufficient that organized, focused scientific research be able to push forward industrial technology in directions the market would pay well for. The firm had to be able to appropriate a non-trivial share of those benefits. In particular, a firm undertaking expensive R & D had to have assurance that its competitors would not reap on the cheap what they did not sow.

Historians and economists studying technical change in capitalist countries have recognized a wide variety of means through which firms appropriate returns to their investments in innovation. However, until my colleagues and I designed a questionnaire with the purpose of exploring exactly where different means of appropriation were effective, there was no systematic map of the terrain. Since the details of the questionnaire, and the broad results of our probes about appropriability, have been reported in several other places, here I will simply summarize some of our findings that are most relevant to the topic of this chapter.[6]

To oversimplify somewhat, we distinguished three broad classes of means through which firms can appropriate returns to their innovations—through the patent system, through secrecy, and through various advantages associated with exploiting a head start—and asked our respondents in different lines of business to score on a scale from one to seven the effectiveness of these means for product innovation and for process innovation.

There were significant cross-industry differences regarding the means rated most effective for appropriating returns to product innovation. Contrary to widespread lay beliefs, patents were rated the most effective instrument in only a small number of industries.[7] Those tended to be of two types: those producing chemical products, and those producing relatively simple mechanical or electrical devices. In industries like semiconductors and computers, a head start and related advantages were reported as the most effective means through which firms appropriated returns. But while there were significant inter-industry differences in the means reported most effective, most of our industries reported at least one of the means as being very effective in enabling a firm to appropriate

returns from product innovation. However, our respondents from the food-processing and metal-working industries tended to report that there was no effective means for them to protect their product innovations. It is interesting to note that these industries are characterized by very low levels of product R & D relative to sales.

Most industries reported that, with the exception of secrecy, the standard means for capturing returns from innovation were less effective for process innovation than for product innovation. A significant fraction of industries reported that no means was particularly effective. It is interesting to learn that in most industries firms spend very little on process R & D.

My colleagues and I have studied the effect of the ability to appropriate returns on the R & D intensity of an industry through regression analysis. We measured ability to appropriate through the means discussed above by the score assigned to the most effective means. Since a firm that accounts for a large share of a market can appropriate returns, even if the means considered in our questionnaire are not effective, we also included a measure of industry concentration in the regression equations. The bulk of industrial R & D is aimed at new or improved products, and the ability to appropriate variable had a positive and significant effect in the cross-industry product R & D intensity regression. The ability to appropriate returns to process innovation variable had a positive sign in the process R & D intensity equation, but it was not significant. In contrast, industry concentration had a much larger positive coefficient in the process R & D regression than it did in the product R & D regression.[8]

These findings are consistent with an important structural feature of capitalist economies. Except in highly concentrated industries, process innovations come largely in the form of new machines and materials (products) made by upstream suppliers. If one reflects on it, one can see that this 'institutional assignments' solution has some efficiency advantages. Except for highly concentrated industries, if the process R & D is done upstream, innovators will be able to get their innovations applied to a larger fraction of industry production than if the process innovations are done in-house and use is restricted (say by secrecy).

In our questionnaire we asked our respondents to assess the contribution made to technical change in their line of business from various outside sources, particularly upstream suppliers. Our regression analysis suggests that upstream contributions are greater, the less concentrated the user industry. Also of interest is the fact that the contribution of upstream firms was strongly positively correlated with the reported contribution of professional and technical societies. These latter would appear to be important vehicles for sharing information about new process technology, and about technological needs. I turn now to consider this sharing, cooperative aspect of capitalist innovation systems.

Technology sharing and R & D cooperation

Private R & D yields proprietary knowledge, initially. But that knowledge does not stay private; it leaks away and becomes public. The fact that proprietary technology ultimately goes public enhances the ability of the economy to use new technology, both in production and as a base for further research and development. However, to the extent that a significant share of the benefits of a company's R & D goes to competitors, or consumers, its incentives to do such R & D are diminished.

At first thought, one might presume that firms that create new technology ought to exert strong efforts to hold back that technology from going public, and in the normal run of things they clearly do. However, in some cases firms take positive action to make their proprietary knowledge available to others.

Patent licensing, of course, is a widespread practice. Here, while there are exceptions, society gains by enlarging the range of firms that can use a new technique, and the licensor collects a portion of those gains. This is easy enough to explain.[9]

Not quite so easy is the practice of implicit patent pooling, which exists in a number of industries, under which rivalrous firms apparently have an agreement not to sue each other for infringements. Such arrangements reflect an apparent agreement among a group of firms that they are all better off if they make a common, big pool of at least some of their technological knowledge, than if they all try to keep their individual pools strictly private. It seems that, within limits at least, rivalrous companies can and do recognize the 'public good' properties of technology. There is a possible problem here, of course, of free riders. My conversations with people in industries where these arrangements exist indicate that patent pools tend to be limited to firms that are active in R & D and hence are contributing to the pool, and that patent suits are likely to arise when non-contributors to the pool are known to be drawing significantly from it.[10]

Technical and professional societies provide formal structures for the sharing of technological information. As noted earlier, these societies seem to serve as vehicles for communication between suppliers and users. They also facilitate the spread of certain kinds of information among rivals. Industrial scientists and engineers, like academics, take pride in their professional reputations, which to a considerable extent are enhanced, and recognized, through publication of articles in journals associated with professional societies, through giving talks at their meetings, etc. For the members of these societies, new technology is the news they gather to hear about. Their operation almost certainly speeds the 'going public' of new generic technological knowledge.[11] R & D cooperation among firms also long has been a part of the capitalist system. Cooperation between user and supplier has been discussed in an earlier chapter. Firms in the same line of business, but operating in separated markets, sometimes have combined cross-licensing with exchange of information about technological problems and opportunities thus engaging in a form of *de facto* R & D cooperation. This pattern has long been common among firms operating in different national markets. In recent years there has been a surge of explicit joint R & D ventures among companies in different countries, particularly in aircraft and electronics.[12]

Even rivalrous firms may forge agreements to have certain kinds of research carried out cooperatively where the results are difficult to keep proprietary, or where it would be disadvantageous to the group as a whole to do so. Typical examples are industry-wide problems, like learning how better to grade and test raw materials, or to establish appropriate standards for inputs. There is a tradition in some industries of trying to fund this type of work collectively, through some kind of an agreed-upon voluntary tax formula for contribution to a trade association, which can support research at universities, or at independent laboratories. There are obvious free-rider limits on this mechanism, however.

In industries closely linked with science, the results of generic research have strong, latent public-good properties and are difficult to keep private for long. As a result only a few, generally very large, firms engage in much of this work. Most of it is undertaken at universities, as I shall elaborate in the following section.

However, in the past ten years or so in several different industries, groups of firms have come together to finance and have generic research carried out through formal cooperative research arrangements. These have been particularly prominent in the United States in the semiconductor and computer industries. Some, like the Microelectronics and Computer Technology Corporation, involve a group of member firms who jointly fund an agreed-upon body of research in special laboratories. Others, like the North Carolina Microelectronics Center, have been organized by state government, and involve public, as well as private, funds. NCMC and the Center for Integrated Systems at Stanford University are associated with universities.[13]

These new cooperative R & D organizations clearly were motivated by the belief that cooperative generic research was an important reason for Japan's technological and economic success in high-technology industries. This part of the Japanese innovation system is discussed at some length in the following chapter. As will be stressed, in Japan universities play little role in such endeavors; indeed, with certain exceptions, the connections between university and industry research are relatively weak. In contrast, in the United States university research and industrial research often have been intertwined closely. I turn now to that part of the United States' innovation system.

The role of universities

Since the last part of the nineteenth century, universities have become an increasingly important part of the capitalist engine. They are a recognized repository of public scientific and technological knowledge. They draw on it in their teaching; they add to it through their research.

Within the United States, university science and engineering and our science-based industries grew up together.[14] Chemistry took hold as an academic field at about the same time that chemists began to play an imporant role in industry. The rise of university research and teaching in the field of electricity occurred as the electrical equipment industry began to grow up in the United States. In both cases the universities provided the industry with its technical people, and academic research provided many of the ideas about product and process innovation.

It is important to recognize that these are two quite distinct kinds of contribution. Academics may be able to teach what new industrial scientists need to know, without having their research be particularly relevant to industry. It may be necessary for young scientists to learn the basic principles and research techniques of a field before being able to work effectively in an industrial laboratory, even if the research being done by academics stands at some distance from what is going on in industry. In some technologies, academic research may be illuminating the opportunities and providing the key insights for industrial R & D, but in others the cutting edge of industrial R & D may be far away from academic research.

The situation is dynamic, not static. There is evidence that academic research in chemistry and electrical engineering has over the years diminished as a source of important new knowledge for industry. Academic researchers were very important to technological developments in the early days of the semiconductor industry, but as time went by research and development in industry increasingly separated itself from what the academics were doing. As I shall document in a moment, at the current time certain areas of academic biology and computer science are very important sources of new ideas and techniques for industry. The latter is a new field; the former is experiencing a renaissance.

To probe at some of these issues, in our survey my colleagues and I asked our respondents to score, on a scale from one to seven, the relevance of various fields of basic and applied science to technical advance in their line of business. We also asked them to score, on the same scale, the relevance of university research in that field.

The fact that an industry rated a field of science highly relevant by no means implies that it rated university research in that field so. Thus while seventy-five industries out of 130 rated the relevance of chemistry as a field at five or greater, only nineteen industries rated university research in chemistry that highly. Forty-five industries rated the relevance of physics at five or greater, but only four gave that high a score to university research in physics. This does not mean that academic research in physics is unimportant over the long run to technical advance in industry. However, the impact will likely be stretched out and indirect, operating through influences on the applied sciences and the engineer disciplines, with the ultimate impact on industrial R & D occurring through these.

What fields of university research have widespread reported relevance to industry, in the sense that a number of industries accredited university research in that field with a relevance score of five or more? Computer science and materials science head the list, followed by metallurgy and chemistry. The industries for which these sciences are important tend to look to universities for new knowledge and techniques, as well as training.

Biology, and the applied biological sciences (medical and agricultural science) appear somewhat special today. While these scientific fields were deemed relevant by only a narrow range of industries, those industries that scored these fields at five or higher almost always rated university research in these fields at five or higher too.

Industries where technological advance is being fed significantly by academic science naturally look for close links with university scientists and the laboratories where that work is going on. In recent years there have been a large number of new arrangements established whereby a single firm or a group of firms funds research at a university laboratory, and receives some sort of advantaged access to that research or its findings. Not surprisingly, the industries most engaged in these activities are ones where firms are large and academic research is rated as highly important to technological change. The major such industries are pharmaceuticals, computers and semiconductors. And the fields of university science being tapped tend to be those where academic research was judged highly relevant to technological advance in those industries—the biological sciences and computer sciences.

My conjecture is that in the United States these kinds of new arrangements for support of industry-oriented generic research will prove more durable than the self-standing industry cooperatives. The same free-rider problems are there, and this limits the magnitude of industry funding. However, the universities themselves are interested in sustaining these programs, and federal and state governments are likely to provide funding for them.

In contrast, because of the presence of MITI, in Japan there is a mechanism for pulling together and subsidizing groups of firms to do generic cooperative research, without involving a university connection. And, as noted, Japan's efforts in this area have preceeded largely without university involvement. Which route will turn out to be more effective in the long run is hard to say.

Government programs in support of technical change in industry

Particularly since the Second World War, government R & D support has been an important part of the capitalist innovation system. However, there are major differences across countries in the roles played by government. In Japan, in certain key industries MITI has orchestrated the R & D attack, getting companies to think through together which directions ought to be followed and what goals pursued if commercial primacy were to be achieved. The United States has no organization like MITI, and the government role generally has not been to try to plan and coordinate the development of broad technologies, except in fields highly relevant to the military, and where it has tried to do so the experience usually has been bad. Rather, the governmental involvement, while large and important in some areas, has tended to be quite selective. In good part this is because where firms are rivalrous, government R & D is viewed with latent hostility as potentially helping one's competitors. To be accepted, government funding must be viewed as benefiting the industry as a whole, or as being justified by overriding natural interests.

I have found it analytically useful to distinguish three different kinds of government R & D support programs. One kind is concerned with basic research. The second is tied to government procurement needs. A third class, of relatively modest importance in the United States, is expressly aimed to advance the commerical competitiveness of a particular industry, or group of firms.[15]

While acceptance of a governmental responsibility for broad funding of basic research at the universities was reached in the United States only after the Second World War, the tradition of governmental support of selected kinds of research at universities developed much earlier. Thus the Hatch Act of 1887 provided for federal funding of agricultural research. During the 1930s there began to be a trickle of federal funds into health-related research. Prior to the Second World War, the armed services funded a considerable amount of research at universities to foster the development of computers.

However, government funding of research at universtities was piecemeal until after the Second World War. During the 1920s and 1930s various influential scientists in industry were trumpeting the importance of university research, as well as teaching, to the health of their industries, but the policy message they and their academic colleagues gave was that industry should band together to support universities. For reasons that are obvious to economists, this call for industrial voluntary contributions to finance a public good failed to achieve its purpose. By the end of the Second World War, industry scientists and university scientists were ready to appeal for public funding for university basic research, and they got it.

Despite a popular impression that the National Science Foundation is the principal governmental source of funds for academic research in the United States, in fact a significantly larger amount of government research support to universities comes from government agencies with particular applied missions, which are seeking to build the scientific understanding to advance those missions. Thus the National Institutes of Health are the dominant source of university funding in the biological sciences. The Department of Defense and the Department of Energy are the major sources of university research support in the fields of concern to them. It is interesting that the fields of academic research deemed most relevant to

technical advance in industry by our questionnaire respondents have tended to be those in which a mission-oriented government agency, as well as the National Science Foundation, has been providing significant support.

Of course, governmental funding of academic research presently is puny compared with governmental support of R & D on products and systems which it wants to procure for its own purposes. The massive defense-procurement-related R & D programs of the last quarter century are so familiar to contemporary observers that it seldom is recognized that this phenomenon, like government support of university research, dates from the Second World War. Prior to then, much less R & D went into the design of military equipment, and a large share of what did was financed by companies themselves as an investment in possible future government sales.

There are various issues about the efficacy of the post-Second World War military R & D programs that can be raised. The central one is the effectiveness of these programs in assuring US and world security. This issue involves both the appropriateness of the goals of the programs, and the effectiveness with which these were pursued. This is a massive complex topic and I cannot address it here in any detail. However, a strong case can be, and has been, made that many of these programs have been frighteningly misguided and disgracefully mismanaged. One major alleged problem is lack of a sophisticated mechanism for appraisal of objectives, of the sort provided by customer markets. Another is lack of serious competition among producers, once an R & D contract has been let.

Another issue is the effect of these programs on the character of US civilian technology. While prior to the Second World War there were many instances where defense demands served to pull into existence technologies of subsequent value to the civilian economy, only a few of these were of major importance. Usually advances in military technology exploited advances originally achieved for other purposes. In the two decades following the Second World War, however, military R & D pulled into place a number of technologies of enormous civilian significance, including modern semiconductors, the electronic computer, and jet aircraft. Various observers have remarked on this and have gone on to argue that DOD R & D has been the key to US technological supremacy during the 1960s and 1970s.

However, during the post-war period military R & D has absorbed a large share of total industrial R & D in the United States, and likely has squeezed out a certain amount of civilian R & D by bidding away scarce scientific and technical resources. Also, with few exceptions, the civilian pay-off per dollar of military R & D certainly has been very small. The technologies pulled into place in the two decades after the war may be exceptional; since 1972 it is arguably the case that military R & D has cost the US considerably in terms of foregone civilian alternatives.

While NASA funding of new technologies has been far smaller than

DOD funding, project Apollo complemented DOD R & D funding and procurement in pulling into existence important advances in semiconductor and computer technologies, and a variety of new materials of subsequent widespread use. The NASA programs have differed from the DOD programs in one very important respect. From its beginnings NASA has advertised itself as a vehicle for advancing American technology across a broad front through technology-stretching procurement.

Perhaps because political support for NASA has not proved durable, in recent years a number of observers have come to see DOD R & D support programs as a vehicle to be used explicitly to enhance civilian technology. However, national security needs to have a strong imperative of its own, and the DOD is unlikely to be willing to see its programs diverted much to generate more 'spillover'. And, as noted, in recent years at least the civilian harvest from military R & D has been thin. Nonetheless, it is likely that a new governmental role, that of advancing civilian technology across a broad front by financing ambitious technology-stretching projects, is gaining political currency. In Japan, MITI has played that role. The United States has no MITI. Just how the United States will organize the provision of government funding for stretching civilian technology, in particular whether it will find a different instrument than the DOD, is hard to say.

Compared with the magnitude of DOD R & D, public finance of R & D to develop civilian technologies has been small change. However, two such programs warrant attention.

One is the long-standing support of nuclear power reactor R & D. This program has been closely controlled from Washington, has tended to concentrate resources on a small number of designs liked by government officials, and arguably has led the nuclear power industry of the United States into a dead end. This, and similar programs in Europe, provide strong support for my earlier warnings about problems of trying to plan and control technical advance centrally.

The other is the more-than-century-old program of support of R & D to lift agricultural productivity. This program has been highly decentralized and quite responsive to farmers' demands. Various studies show that the rate of return has been very high.

While many observers have looked to the agricultural R & D support program as a model that might be widely extended, in fact most industries have resisted government R & D support, not welcomed it, for the reasons suggested earlier.

What is it about farming that has led the 'firms' in this industry actively to demand, not resist, government R & D support of applied as well as generic, work? A principal factor is that individual farmers, unlike individual firms in industries such as pharmaceuticals, are not in rivalrous competition with each other. It does not hurt one farmer if his neighbor becomes more productive. Put another way, there is very little proprietary knowledge among farmers. At the same time, regional groups of farmers perceive that their sales can be enhanced and their profits increased if their

productivity and the quality of their crops are improved. Thus a regional group of farmers has a strong interest in getting effective applied research and development undertaken aimed to enable them, as a group, to become more competitive. The circumstances here clearly are rather special. It is unlikely that the agricultural model can be used in many industries.

Reprise

This essay has been concerned with laying out the institutional structures supporting technical advance in modern capitalist countries, with particular focus on the current US scene. I hope I have been persuasive that capitalist innovation systems are far more complex than commonly recognized, and far more complex than their depiction in extant economic models.

I began this essay by observing that, while the capitalist engine obviously has been a powerful one, inefficiency is built into its basic design. While privatization of new technology harnesses the profit motive to its creation and leads innovators to be sensitive to market opportunities, privatization causes waste both in the use and generation of technology. While I do not believe that the power that has been generated could have been with an engine of radically different design, that is far from saying that the engine is optimal in any way.

However, the engine undoubtedly is less inefficient than simple models of it, which miss the complexity, would suggest. The system has managed to avoid a good share of the costs of privatization, while preserving the profit motive for industrial innovation, by treating as public large parts and aspects of technology, through involving institutions like universities and providing liberal amounts of public money. In my view the capitalist innovation system has solved the institutional assignment problem not optimally, but tolerably well.

Of course, the institutional assignment problem never is solved once and for all. As science and technology change, and as the nature of research and development in different areas of science and technology change, so too do the institutional structures appropriate to the endeavor. In the United States there is no single agency responsible for looking at the national innovation system as a whole and recommending or mandating needed changes. Rather, new institutions and institutional assignments are created pluralistically, and the structure itself changes through an evolutionary process.

The modern corporate R & D laboratory itself, the heart of the system, grew up during the early twentieth century as a result of bets made by a few scientists and businessmen. Events proved them right, and the lesson was learned by other firms. The nature and the source of funding of university research in the United States have also followed an evolutionary pattern, with different kinds of universities doing different things and being open to different kinds of initiatives on the part of business and government. As noted earlier, the system of government funding of university research arose only after various other proposals for funding had been tried and failed.

At the present time there are a lot of organizational experiments going on. Industry cooperation in the finance of generic research is now the rage. In the United States there has been a surge of new arrangements linking industry to university research, in some cases initiated by industry, and in some cases initiated through government programs. It is too early to judge which new departures will be fruitful and survive and which will not. But they are going on. Such institutional experimentation may be the most durable strength of the system.

As noted, at the present time there are many voices in the United States suggesting that the country experiment with a structure like MITI. Others have suggested that MITI, while appropriate to catching up technologically, is not a useful institutional vehicle for a country at the frontier. Still others argue that it is, but that is hard luck for the United States because our historical traditions and political structure make it impossible for us to put in place such a structure, or not to abuse it were it in place. Whether the DOD should be used in a MITI-like way in technological fields of national security concern is, as observed, a matter of current dispute.

I shall make a strong bet that over the coming decade, through one route or another, that the US innovation system will adopt a number of features that are noteworthy on the Japanese scene. The following chapter describes some of these features.

Notes

1. Earlier chapters have noted that this view of technical change is flatly at odds with the treatment of technical change contained in virtually all neo-classical models that address the topic. This conflict is discussed at length in Nelson and Winter (1982), where we also attempt to lay out a general structure for formal evolutionary models of technological change.
2. The basic Schumpeterian model was mapped out in his *Theory of Economic Development*, first published in 1911. The characterization in *Capitalism, Socialism, and Democracy* (1942), published nearly thirty years later, differs mainly in that large corporations have replaced individual entrepreneurs as the locus of innovation.
3. This and subsequent characterizations of the benefits and costs of making technology public are based on Nelson and Winter (1982), particularly Chapter 14.
4. Perhaps the best analysis that views this problem in terms of balance of trade-off is by Nordhaus (1969).
5. Thus this is a study in the spirit of 'the new institutional economics'. Oliver Williamson's work in this area (1975, 1985) has been particularly influential, but many other economists now are concerned with the details of institutional structures and how the present ones came into being. For a survey of much of this work, see Langlois (1986).
6. A description of the questionnaire and a preliminary report of some of the findings is contained in Levin *et al.* (1984).
7. Taylor and Silberston (1973), Scherer *et al.* (1959), and Mansfield, Schwartz, and Wagner (1981) also have provided evidence of the limited number of industries where patents are important.
8. These results are consistent with those of Pavitt (1984) in his more systematic study bearing on upstream–downstream relationships.
9. Unfortunately, economists have only begun to map out the domain of patent licensing. For a sketch, see Caves *et al.* (1983).
10. For case studies of technology sharing, see von Hippel (1986) and Allen (1983).
11. Unfortunately, there is very little published research on just what technical societies do, and the nature of the information passed around. My observations above are based on limited evidence.
12. For the story of aircraft, see Mowery (1986); for electronics, see Haklisch (1986).
13. These and other partnerships between universities and industry are described in *New Alliances and Partnerships in American Science and Engineering*, put out by the Government-University-Industry Research Roundtable, National Academy Press, 1986. See also the survey by Peters and Fusfeld (1982).
14. See Thackray (1982), Noble (1977), and Rosenberg (1985).
15. This section draws heavily on Nelson (1982, 1984).

References

Allen, R.C. (1983), 'Collective invention', *Journal of Economic Behavior and Organization*.

Caves, R., Crookell, H. and Killing, J.P. (1983), 'The uncertain market for technology licenses', *Oxford Bulletin of Economics and Statistics*, August.

Cohen, W. and Leventhal, D. (1985), 'The endogeneity of appropriability and R and D investment', Carnegie Mellon University, mimeo.

Dosi, G. (1982), 'Technological paradigms and technological trajectories', *Research Policy*, vol. 11.

Freeman, C. (1982), *The Economics of Industrial Innovation*, London, Frances Pinter.

Haklisch, C.S. (1986), *International Alliances in the Semiconductor Industry*, Center for Science and Technology Policy, New York.

Haklisch, C.S., Fusfeld, H.D. and Levenson, A.D. (1984), *Trends in Collective Industrial Research*, Center for Science and Technology Policy, New York.

Hippel, E. von (1982), 'Appropriability of innovation benefit as a predictor of the source of innovation', *Research Policy*.

—— (1986), 'Competition between competing firms: informal know-how trading', Sloan School MIT Working Paper No. 1759–86.

Langlois, R.W. (ed.) (1986), *Economics as a Process: Essays in the New Institutional Economics*, New York, Cambridge University Press.

Levin, R.C. Klevorick, A.K., Nelson, R.R. and Winter, S.G. (1984), 'Survey

research on R and D appropriability and technological opportunity: Part I–Appropriability, Yale University, mimeo.
Mansfield, E., Schwartz, M. and Wagner, S. (1981), 'Imitation costs and patents: an empirical study', *Economic Journal*, December.
Mowery, D.C. (1983), 'The relationship between the contractual and intra-firm forms of industrial research in American manufacturing, 1900–1940', *Explorations in Economic History*.
—— (1984), 'Firm structure, government policy, and the organization of industrial research: Great Britain and the United States, 1900–1950', *Business History review*.
—— (1986), 'Multinational joint ventures in product development and manufacture: the case of commerical aircraft', Carnegie Mellon University, mimeo.
Nelson, R.R. (ed.) (1982), *Government and Technical Progress: A Cross-Industry Analysis*, New York, Pergamon Press.
—— (1984), *High Technology Policies: A Five Nation Comparison*, American Enterprise Institute, Washington D.C.
Nelson, R.R. and Winter, S.G. (1982), *An Evolutionary Theory of Economic Change*, Cambridge, Mass., Harvard University Press.
Noble, D. (1977), *American by Design*, New York, Knopf.
Nordhaus, W.D. (1969), *Invention, Growth and Welfare: A Theoretical Treatment of Technological Change*, Cambridge, Mass., MIT Press.
Pavitt, K. (1984), 'Sectoral patterns of technical change: towards a taxonomy and a theory', *Research Policy*, vol. 13.
Peck, M.J. (1986), 'Joint R and D: the case of the MCC', *Research Policy*.
Pelikan, Pavel (1986), *Institutions, Self Organization, and Adaptive Efficiency: A Dynamic Assessment of Private Enterprise*, The Industrial Institute for Economic and Social Research, Stockholm, May.
Peters, L. and Fusfeld, H. (1982), *Current U.S. University and Industry Research Connections in University and Industry Research Relationships*, National Science Board, Washington, DC.
Polanyi, M. (1962), *Personal Knowledge: Towards a Post-Critical Philosophy*, New York, Harper Torchbook.
Polsby, N.W. (1984), *Political Innovation in America: The Politics of Policy Initiation*, New Haven, Yale University Press.
Reich, L. (1985), *The Making of American Industrial Research: Science and Business at G.E. and Bell*, New York, Cambridge University Press.
Rosenberg, N. (1985), 'The commercial exploitation of science by American industry', in K.B. Clark, R.H. Hayes and C. Lorenz (eds.), *The Uneasy Alliance: Managing the Productivity and Technology Dilemma*, Boston, Harvard Business School Press.
Scherer, F.M. *et al.* (1959), *Patents and the Corporation*, Boston, privately printed.
Schumpeter, J.A. (1934), *The Theory of Economic Development*, Cambridge, Mass., Harvard University Press (first published in Germany, 1911).
—— (1942), *Capitalism, Socialism, and Democracy*, New York, Harper.
Taylor, C.T. and Silberston, Z.A. (1973), *The Economic Impact of the Patent System: A Study of the British Experience*, Cambridge, Cambridge University Press.
Teece, D.J. (1986), 'Capturing value from technological innovation: integration, strategic partnering and licensing decisions', Berkeley, mimeo.

Thackray, A. (1982), 'University–industry connections and chemical research: an historical perspective', in *University–Industry Research Relationships*, National Science Board, Washington, DC.
Williamson, O.E. (1975), *Markets and Hierarchies*, New York, Free Press.
—— (1985), *The Economic Institutions of Capitalism*, New York, Free Press.

16 Japan: a new national system of innovation?

Christopher Freeman

Science Policy Research Unit, University of Sussex, Brighton and MERIT, State University of Limburg, Maastricht

Technology gaps and institutional innovation

When Britain opened up a major 'technological gap' in the first industrial revolution, this was related not simply to an increase in invention and scientific activities, and to a cluster of innovations in the textile, iron and engineering industries, but to novel ways of organising production, investment and marketing and novel ways of combining invention with entrepreneurship (see Chapters 2 and 3). Similarly, when Germany and the United States overtook Britain in the latter part of the nineteenth century and in the twentieth century, their success was also related to major institutional changes in the national system of innovation, as well as to big increases in the scale of professional research and inventive activities and new clusters of radical innovations. In particular both countries developed new ways of organising the professional education of engineers and scientists and of organising research and development activities as specialised departments within firms, and employing the new graduate engineers and scientists in design, production, marketing and management as well as in research.

Again today, when Japan is drawing ahead in some important new technologies, this is related not simply or even mainly to the scale of R & D, but to other social and institutional changes. Japan drew ahead of the United States in the relative intensity of *civil* industrial R & D already in the 1970s and is now well ahead (Patel and Pavitt, 1987). But more significant are the measures of 'output' of the science and technology system which suggest a Japanese lead in exploiting the *results* of R & D (Freeman, 1987). Japanese trade performance in the 1970s and 1980s is further indirect evidence of this success, based as it is on new product and process design, and high quality.

This chapter therefore considers the nature of this new 'technology gap' and those qualitative aspects of the Japanese 'national system of innovation' which might help to explain it.

It seeks to identify some of the distinguishing features of the Japanese 'national system of innovation', not because they are unique, but on the contrary because they are likely to be emulated increasingly as international technological competition intensifies. This chapter first of all discusses the role of central government, especially MITI; it then considers the role of firms, especially the 'Keiretsu'; and finally other social and educational innovations.

The role of MITI

Recently, a voluminous literature has grown up on the role of MITI in explaining the technological performance of the post-war Japanese economy. Some of this literature tends to attribute Japanese success mainly to government policies inspired by MITI, while another part concentrates on debunking this notion. It is clear that a large part of the Japanese success must be attributed to the management of technical change by numerous Japanese enterprises, but this success was related to social and institutional changes promoted and sometimes initiated by MITI, and to the persistent pursuit of certain long-term strategic goals.

The not-so-invisible guiding hand of MITI shaped the long-term pattern of structural change in the Japanese economy and this influence was largely exerted on the basis of judgements about the future direction of technical change and the relative importance of various technologies. The central point of interest from the standpoint of this analysis is that in the immediate post-war period, after an intense debate, Japan specifically rejected a long-term development strategy based on traditional theory of comparative advantage (Shinohara, 1982). This was apparently at that time being advocated by economists in the Bank of Japan and elsewhere who subscribed to the free-trade doctrines of the classical school. They had advocated a 'natural' path of industrial development, based on Japan's relatively low labour costs and comparative advantage in labour-intensive industries such as textiles. One of the central points at issue was whether Japan could hope to compete in the automobile industry and whether special steps should be taken to encourage its growth, but the debate affected industrial and trade policy in its entirety.

In the early days, according to G.C. Allen (one of the few European economists who consistently attempted to study and learn from Japanese experience), the views of the Bank of Japan had some influence. But on the whole the bureaucrats and their advisers at the Ministry of International Trade and Industry (MITI) prevailed. According to Allen:

> Some of these advisors were engineers who had been drawn by the war into the management of public affairs. They were the last people to allow themselves to be guided by the half-light of economic theory. Their instinct was to find a solution for Japan's post-war difficulties on the supply side, in enhanced technical efficiency and innovations in production. They thought in dynamic terms. Their policies were designed to furnish the drive and to raise the finance for an economy that might be created rather than simply to make the best use of the resources it then possessed. [Allen, 1981]

Thus, MITI saw as one of its key functions the promotion of the most advanced technologies with the widest world market potential in the long term. In this respect MITI differed from almost all other analogous Ministries and Departments in Western Europe or North America, which mostly did not see themselves as responsible for long-term technology policies until much later (i.e. in the 1970s or 1980s) and were guided by very different conceptions of comparative advantage. Even today most do not accept the same responsibility for technology as MITI does.

As early as 1952 the 'Enterprises Rationalisation Promotion Law' provided direct government subsidies for

> the experimental installation and trial operation of new machines and equipment, plus rapid amortisation and exemption from local taxes of all investments in research and development; second, it authorised certain industries (to be designated by the cabinet) to depreciate the costs of installing modern equipment by 50 per cent during the first year; and third, it committed the central and local governments to building ports, highways, railroads, electric power grids, gas mains and industrial parks at public expense and make them available to approved industries. [Johnson, 1982]

From this and many other examples it is clear that MITI (and some other ministries) saw it as one of their main responsibilities to encourage the introduction of new technologies through new investment. Furthermore this law is particularly significant in its clear recognition of the crucial role of externalities and infrastructure for innovative firms (see Chapter 21 by Perez and Soete), and of the role of government in ensuring the availability of the necessary infrastructural investments. This tradition has continued right through to the 1980s in the development of regional policies which lay great stress on the development of science, education, communications and transport infrastructure (Kuwahara, 1985). Regional policies have consistently sought to strengthen technological capability throughout the country, particularly in small and medium-sized firms. There are nearly 200 'prefecture laboratories' (i.e. an average of four in each prefecture) which offer research and technical advisory services, varying with the structure and needs of local industry (Ergas, 1984).

However, although the young generation of aspiring technocrats in MITI espoused enthusiastically their role in charting and promoting the adoption of advanced technology in post-war Japan, they never attempted to do this alone. On the contrary, they established a mode of working which depended upon a continuing dialogue on questions of technological development, both with industrial R & D people and with university scientists and technologists. This meant that they were well informed about the broad trend of new developments and were well placed to form an overall view or 'vision' of what was required. Later, this mode of continuous consultation with the Japanese scientific and technological community was systematised and this has been described by Irvine and Martin (1984) in their book, *Foresight in Science Policy: Picking the Winners*.

The informal mode of continuous consultation remains, however, even more important and has been vividly portrayed by Vogel (1980). In his book, *Japan as No.1*, he lays great stress on these informal contacts as fundamental to MITI's success in restructuring the Japanese economy and orienting the leading firms to a desired course of action.

> The agreements they promote among companies in a given industrial sector require a higher level of trust than can be achieved through formal contacts. When a specific issue arises, interested parties schedule even more informal gatherings than usual. When necessary, bureaucrats from other ministries join these meetings and other knowledgeable experts and men of influence are called in . . . American government officials and business men negotiating economic matters feel at a great disadvantage because Japanese officials are much better informed, not only about Japanese companies but often about American companies . . . [p. 76]

Although the *responsibility* for promoting advanced technology was accepted throughout the post-war period, the *mode* of fulfilling this responsibility changed over time. The policies of MITI have been nothing if not pragmatic and eclectic in their choice of *means*, although quite consistent in the long-term goals. Consequently there have been big changes in the methods and instruments of industrial technology policy, as both external and internal circumstances changed. As Koshiro (1985) has shown, the post-war period can be divided into distinct phases: in the 1950s direct physical controls, allocations and priorities were still important, but in the 1960s indirect fiscal and other incentives were used to stimulate and channel the huge investment boom. In the 1970s and 1980s there has been a clear recognition of the strategic importance of information technology.

Clearly the success of such a system must depend on a reasonably accurate identification of the key areas in which to concentrate technological effort and new investment, both at national level, at 'Keiretsu' level and at company level. Serious mis-specification for new areas of investment could lead to enormous waste of resources and clearly 'accuracy' in this context refers both to future trends in technology and future trends in world markets and society generally. A lot of effort in Japan goes into the preparation of such long-term 'visions' of the future, both at national level and company level, as well as in universities and in numerous think-tanks (e.g. STA, 1972). There is no other society where financial institutions, banks and even the Ministry of Finance devote such attention to the future direction of technical and social change.

The Japanese system of informal and formal technological forecasting permits the formulation of technological and industrial policies not so much on the basis of particular products or of existing industrial statistics or the weight of established firms and industries, but rather on the basis of those new technologies which are likely to transform the established existing pattern.

The recognition of the importance of information technology occurred very early in Japan. Of course there were individual experts and companies in Europe and America who recognised it at least as early. But the importance of the national Japanese system of forecasting lay in the diffusion and generalisation of these expectations through a large number of companies in a great variety of industries. This helped to create a climate where firms would make investments in new products and processes associated with the new technology on a much larger scale than elsewhere in the OECD area, where this confidence was sometimes lacking. This was especially important in the diffusion of information and communication technology (see Chapter 3 on uncertainty and investment at times of change in techno-economic paradigm).

The promotion of 'generic' technologies, especially information technology, has now become a regular feature of technology policy and industrial policy in almost every member country of the OECD in the course of the 1970s and 1980s. The extent to which such efforts are successful will depend not simply on the scale of resources, which are committed in the public and private sectors of the economy, but also on social conditions and attitudes and appropriate institutions, particularly the 'national system of innovation'. The Japanese system seems particularly well adapted to take advantage of the enormous potential of information technology for several reasons:

(1) the systems approach to process and product design;
(2) the flexibility of the industrial structure;
(3) the capacity to identify crucial areas of future technological advance at national and enterprise level;
(4) the capacity to mobilise very large resources in technology and capital in pursuit of strategic priorities;
(5) the horizontal flow of information within and between firms.

We now turn to consider the role of firms, especially the 'Keiretsu' in relation to these five points.

The role of firms in the Japanese national system of innovation

Ever since 1868 Japanese central government had worked closely with industry in modernising the Japanese economy and importing foreign technology. But Japanese national policies for technology in the 1930s were driven mainly by military imperatives. Dependence on imported oil, rubber and other industrial materials and the early development of a war economy led to a one-sided concentration on the promotion of technologies primarily designed to strengthen autonomous Japanese military capability.

As Yamauchi (1986) pointed out,

> the government and industrialists did not pay as much attention to reducing production costs or increasing productivity as to how they could modify imported technology to produce first class defence equipment and adjust it to Japanese raw material sources. . .

It would of course be a mistake to underestimate the technological advances already achieved in Japan before and during the Second World War. Japanese industry was not only capable of independent design, development and production of complex military equipment, such as air-

craft and tanks, but also of a wide range of machinery. However, Japan lagged behind in the techniques of mass production of cars and consumer durables and in some branches of the chemical industry such as flow processes in petrochemicals and synthetic materials.

In the post-war period it was recognised from the outset that the import of mass and flow production techniques was crucial for economic success. It is not so widely recognised that from the earliest days this was accompanied by a systematic policy designed to improve these technologies.

The method of assimilating and improving upon imported technology was mainly some form of 'reverse engineering' (Tamura, 1986; Pavitt, 1985). This involved trying to manufacture a product similar to one already available on the world market but without direct foreign investment or transfer of blue-prints for product and process design. The widespread use of reverse engineering in the 1950s and 1960s had several major consequences for the Japanese system of innovation.

(a) Japanese management, engineers and workers grew accustomed to thinking of the entire production process as a system and of thinking in an integrated way about product design and process design. This capability to redesign an entire production system has been identified as one of the major sources of Japanese competitive success in industries as diverse as shipbuilding, automobiles and colour television (Jones, 1985; Sciberras, 1981; Peck and Wilson, 1982). Whereas Japanese firms made few original radical product innovations in these industries they did make many incremental innovations, as shown for example, by the Japanese lead in vehicle patenting in the United States. They also redesigned and reorganised many processes so as to improve productivity *and* raise quality. The automobile industry is probably the most spectacular example (Altschuler *et al.*, 1985; Jones, 1985).
(b) Japanese engineers and managers grew accustomed to the idea of 'using the factory as a laboratory' (Baba, 1985). The work of the R & D department was very closely related to the work of production engineers and process control and was often almost indistinguishable. The whole enterprise was involved in a learning and development process and many ideas for improving the system came from the shop-floor.

    Since almost all studies of the management of innovation in Western Europe and the United States point to the lack of integration between R & D, production management and marketing as a major source of failure, the integrative effect of learning by creative reverse engineering conferred a major competitive advantage on many Japanese firms. It also gave production engineering a much higher status than is usually the case in Europe or the United States. These horizontal links are probably the single most important feature of the Japanese national system of innovation at enterprise level.
(c) Reverse engineering in such industries as automobiles and machine

tools also involved an intimate dialogue between the firm responsible for assembling and marketing the final product and numerous suppliers of components, sub-assemblies, castings, materials and so forth. The habits, attitudes and relationships engendered during this prolonged, joint learning process did much to facilitate the high degree of cooperation with subcontractors which finds expression, for example, in the 'just-in-time' system. Another factor fostering such intimate relationships was the conglomerate 'Keiretsu' structure of much Japanese industry (see below).

(d) The emphasis on high quality of products which is characteristic of Japanese technology policy also owed much to the experience of reverse engineering. In the 1950s, the first production models, whether in automobiles, TV sets or machine tools, were often of relatively poor quality (Jones, 1985; Baba 1985). A determined effort to overcome these defects led to a widespread acceptance of such social innovations as 'quality circles' (originally an American innovation) and to the development of greatly improved techniques of quality control not simply at the end of the production run but at every stage, including all components from subcontractors. Some of the most important (and most closely guarded) Japanese innovations have been on-line inspection, test and quality control equipment and instrumentation arising in this process. Where the quality of components from subcontractors was particularly bad, as in the case of castings, this led to intense pressure from MITI for the restructuring of the entire industry and a drastic change in its techniques. In other cases these problems were dealt with by joint technical effort between assemblers and subcontractors.

This Japanese approach to the import of technology may be compared and contrasted with the methods used on the one hand in the Soviet Union and on the other hand in many Third World countries. The Soviet Union was also engaged in the large-scale development and import of technology in the twentieth century (Sutton, 1971) and also used reverse engineering, but in the Soviet system much of the responsibility for diffusion and development rested with central research institutes or Project Design Bureaux. This meant that much of the 'technological learning process' took place there, rather than at enterprise level, and acute problems were experienced in the transfer of technology from the specialised R & D institutes to factory level management. This weakness has been increasingly recognised and the institutional arrangements have been changed considerably in the 1970s and 1980s to strengthen R & D at enterprise level and to regroup research activities in close relationship with enterprises. This process has probably gone furthest among the socialist countries in the DDR, where in the 1980s responsibility for R & D, design, production and world marketing is now under one management system in the large groups called 'Kombinats'.

In many Third World countries, on the other hand, the method of technology transfer was very often either through subsidiaries of multinationals or by the import of 'turn-key' plants designed and constructed by foreign contractors. Neither of these methods is likely to result in an intense process of technology accumulation in the (relatively passive) recipient enterprise. Dissatisfaction with both these methods has led, on the one hand, to pressures on multinationals to set up local R & D activities in addition to training, as in Brazil. On the other hand, it has led to efforts to 'unpackage' imported technology and to devolve part of the design and development to local enterprises. The Japanese policy of rejecting foreign investment and putting the full responsibility for assimilating and improving upon imported technology on the enterprise is more likely to lead to 'systems' thinking and to total systems improvement.

At an earlier stage in economic development, by establishing specialised in-house R & D departments in the chemical and electrical industries, German and American firms were able to achieve a world lead in the newest technologies in the latter part of the nineteenth century. At that time the R & D department provided a point of entry for the findings of fundamental scientific research to be assimilated and used in contemporary industrial technology and for some of the brightest scientists and engineers to be recruited into the new industries. However, as the R & D department became a part of the regular organisation of most large industrial firms, some of the disadvantages of specialisation began to be apparent, especially the cultural barriers between 'academic' R & D, and the production and marketing divisions of companies (see, for example, Burns and Stalker, 1957). One of the reasons for the relatively successful performance of small entrepreneurial firms in areas where development costs were low was that they were able to overcome the internal communication and cultural barriers much more easily through the integration of research, production and marketing by the innovative entrepreneur working with a few colleagues.

The Japanese success, however, seems to have been based far more on an integrative approach within *large* firms. Their approach to product and process design, often originally developed through reverse engineering, created a new style of innovation management which reintegrated R & D with engineering design, procurement, production and marketing even in the largest organisations. As development, production and marketing went ahead, the whole organisation was committed to the new products and processes in a way that was relatively uncommon in other countries. Moreover, once development work began, lead times were often very short, especially in the electronics industry.

Horizontal information flows increasingly characterised Japanese management organisation rather than the vertical flows so characteristic of the hierarchical US corporation (Aoki, 1986). It is interesting that this feature of Japanese management was singled out not only by an academic economist (Aoki) but also by the head of research of one of the most successful world-wide competitors of the Japanese electronics companies—Northern Telecom—in his centenary address to the Canadian engineering institutes.

These horizontal information flows are also an important feature of the inter-company cooperation within the 'Keiretsu', as was emphasised by the MIT study of the world automobile industry (Altschuler *et al.*, 1985):

> The Japanese practice of group coordination, simultaneously attains the scale and coordination advantages of Western-style vertical integration and the flexibility of decentralisation. Its aim is cooperation and mutual information flow between the parts rather than rigid top-down hierarchy . . . Because the members of an industrial group are coordinated through their equity links and management working groups, they are able to share research facilities, support staff (such as accountants and marketing responsibilities) and production capacity . . . Finally, the Japanese have developed relationships among the automotive industrial groups, their financing sources, and industrial groups in other sectors that seem to carry competitive advantages over the typical arrangement in the United States . . . [p. 148]

Clearly, the Japanese model of competition has some distinctive features. It permits and encourages a long-term view with respect to research, training and investment. For this reason it is a vital element in the national system of innovation. Firms which are in more perfect competitive market situations would not be able so easily to amass or to allocate resources for these long-term objectives. Indeed there would be strong pressures from the capital markets for them to improve short-term profitability by sacrificing long-term investments. This model does not appear to correspond very closely to the model of a competitive capital market and evolutionary competitive survival of competence postulated in Chapter 18.

The formation of large conglomerates and of vertically integrated groups of companies is of course not confined to Japan. But most commentators on the Japanese economy agree with the MIT study that they have a specially important role there particularly in relation to technology, finance for long-term investment and world marketing strategies and networks (the 'Shosha').

Goto (1982), in an original analysis of the 'Keiretsu' explains the peculiar success of the Japanese groups in terms of Williamson's (1975) theory of transaction costs:

> resource allocation in a market economy is implemented within the firm under the direction of the manager, or through the market by the market mechanism. The division of labour between these two modes is determined by their comparative efficiency as an instrument to implement all transactions. Here we can see there is a third possibility. From the standpoint of the firm, by forming or joining a group, it can economise on the transaction costs that it would have incurred if the transaction had been done through the market, and at the same time, it can avoid the scale diseconomies or control loss which would have occurred if it had expanded internally and performed that transaction within the firm.

He notes that Williamson refers to the Zaibatsu as an alternative mode in a footnote but does not develop the point because he regards them as a

'culturally specific' phenomenon. Goto argues that they are not a uniquely Japanese type of institution and that they have a wider economic significance. Chapter 12 by Teece strongly confirms this.

Imai and Itami (1984) have given a review of other ways in which Japanese firms use the 'organisational' mode of resource allocation rather than the market mode. The lower cost of capital for strategically important areas of long-term investment is recognised by many other commentators as one of the most important.

Education, training and related social innovations

None of the developments at enterprise level which have been described above would have been possible without changes in the education and training of the work-force and a set of related social changes which broke down the barriers between 'blue-collar' and 'white-collar' types of employment.

The Japanese education and training system is remarkable for two features: first of all, in the absolute numbers of young people acquiring secondary and higher levels of education, especially in science and engineering; secondly, in the scale and quality of industrial training, which is carried out mainly or entirely at enterprise level. Japan is now, together with the United States and the Soviet Union, among the leading countries in the world in the extent of educational opportunity, both at secondary and tertiary level. However, this provision for secondary and higher education has been complemented increasingly by an intensive system of education and training in the large firms. The high level of training intensity goes back a long way in the leading firms and was directly related to their efforts to assimilate foreign technology. Thus, for example, Fukasaku's (1987) study of training and research in Mitsubishi shipbuilding before the First World War shows that already then the company had its own training establishments providing a high level of technical education. Sakurai's (1986) study of Mitsubishi's diversification from mining and shipbuilding to chemicals and electrical engineering between the two world wars shows again that the training and retraining of their own employees from the mining companies was a crucial element in their success, as well as the setting up of a central R & D laboratory to develop new products and processes, which also acted as a training establishment.

The combination of a high level of general education and scientific culture with thorough practical training and frequent up-dating in industry is the basis for flexibility and adaptability in the work-force and high-quality standards. The Japanese system of industrial training is distinguished further by its close integration with product and process innovation. The aim is to acquaint those affected by technical change with the problems that are likely to arise, and give them some understanding of the relationship between various operations in the firm. This again greatly facilitates the horizontal flow of information. Thus the 'systems' approach is inculcated at all levels of the work-force and not only at top management level. Obviously, the availability of a large number of good-quality professional engineers, not just in R & D but in production engineering and management too, generally has played a vital part in the Japanese success in the import of technology, the redesign of processes and products, and increasingly now in autonomous innovations.

This success was also based on a high level of *general* education. Prais (1987), in his comparison of Japanese and British education, concludes that the standard of attainment is higher in Japan especially for the average and below-average pupils. One of its consequences is to encourage all-round capability at lower levels in the work-force so that problems of breakdown and maintenance are more rapidly dealt with. Another advantage of this approach is a smoother assimilation and readier acceptance of new process technology. Finally, it has facilitated the participation of the work-force in the improvement of quality, which is widely acknowledged as one of the outstanding successes of Japanese innovation management in the post-war period.

The huge increase in the scale of education and training in Japan since the Second World War was accompanied by a number of other major social innovations which reinforced the capacity of Japanese society to implement technical change at a high rate. Some of these were only indirectly related to the national system of innovation but Aoki (1985) and Dore (1985) are probably right in seeing them as complementary to those changes which directly affected the management of technical innovation.

The abolition of the distinctions between blue- and white-collar workers started earlier and has probably gone further in Japan than most other industrial countries. Before the Second World War the distinctions in both status and income were probably at least as rigid as in other industrial countries. But immediately after the Second World War, a rather curious combination of circumstances led to a social levelling process which went further and faster in Japan than in most other countries. During the period of direct American military occupation and government, a number of radical 'New Dealers' were able to put into practice tax reforms, which they were unable to implement during Roosevelt's 'New Deal' in the United States. At the same time the strength of the Communist-led unions in some industries was sufficiently great in the immediate post-war period to induce the newly restructured large Japanese firms to make major concessions to blue-collar workers with respect to their status and forms of remuneration. The change of management in many of the larger firms also facilitated a new approach both to social and technical issues. This unique combination of circumstances led to rather low income differentials between management, white-collar and blue-collar workers—probably lower than almost any other capitalist industrial country—and to the abolition of most of the status differentials, which continued to characterise British and some other European firms (such as different hours of work, different eating arrangements, etc.). It should be remembered, however, that income statistics do not tell the whole story since the upper groups receive many 'perks' and benefits in kind.

These social changes, together with the system of annual bonuses related to company performance and the 'lifetime employment' system in the large company sector, provided a powerful combination of incentives. Like the education and training system and the systems approach to technical innovation, they provided a basis for continuous improvement in productivity through work-force involvement in technical change, although 'lifetime employment' never covered the entire work-force even in the large firms (McCormick, 1985). This certainly did not mean any form of 'workers' control'. The social system remained hierarchical and deferential but with considerable opportunities for climbing ladders.

The new techno-economic paradigm

It has been argued that, with respect to the role of central and local government, the organisation of firms for the management of innovation and the role of education and training, the Japanese national system of innovation has some important competitive advantages. These advantages are likely to be particularly important in relation to the change of techno-economic paradigm which is such an important feature of contemporary structural change in the world economy (see Chapter 3). Indeed the developments in the Japanese national system of innovation have received an important stimulus from information technology as well as reciprocally promoting its diffusion.

Although Japanese firms were not the major contributors to the original radical innovations in computers and telecommunications, the Japanese technological forecasting system did indeed identify the main elements of the emerging 'ICT' paradigm earlier than elsewhere, and this enabled Japanese firms to exploit the potential of the new paradigm in such areas as robotics, CNC machine tools, flexible manufacturing systems, construction and financial services more rapidly than most other countries. Japanese policy-makers recognised the crucial importance of information technology at a very early stage and the goal of a 'knowledge-intensive' economy was proclaimed already at the beginning of the 1970s. The OPEC crisis accelerated a trend in the policy-making of Japanese government and industry, which was already well articulated.

Various measures to support the computer industry, the semiconductor industry and the electronics industry go back to the 1950s, but at that time they were not regarded as part of a new technological paradigm which would pervade the entire economic system, but simply as important new sectors with very high growth rates and very big world market potential. The need to economise in space and materials were also important considerations. Originally, it was the consumer-goods electronics industry which was the main focus of attention, as one part of the overall expansion of the then dominant assembly-line, mass-production technology. The semiconductor industry was seen as important mainly in this context. The capacity of the Japanese semiconductor industry to overtake the US industry provided the basis for an increasingly original contribution to innovations in consumer electronics.

The first computer development programmes were initiated in Tokyo University, the laboratories of the Japan Telegraph and Telephone Corporation (NTT), and MITI institutes in the 1950s, and a succession of laws were passed from 1957 onwards giving MITI broad powers to promote electronics industry developments including computers.

The combined effect of these various measures, persistently pursued by MITI and other agencies for more than two decades, was to establish a semiconductor industry which was drawing ahead of the US industry by the 1980s and a viable Japanese computer industry able to compete with IBM when trade was liberalised (Gregory, 1986). No European companies had comparable success.

But the explicit and successful fostering of these two strategic industries was only one strand of Japanese information and communication technology policy in the 1970s and 1980s, albeit a very important one. Two other strands were equally, if not more important from the standpoint of the economy as a whole: the diffusion of IT *outside* the electronics industry itself and the policy for the telecommunications network.

The distinctive feature of a new techno-economic paradigm (see Chapter 3) is that it has effects in every sector of the economy, providing scope everywhere for renewal of productivity increases through a combination of organisational, social and technical innovations and for a broad range of new and improved products and services. The main problem in periods of change of paradigm is not so much in the leading-edge industries (in this case, computers and VLSI) as in the adaptation of the rest of the economy. It is here that the type of structural and institutional inertia problems identified by Perez (1983, 1985) are acute, and that national policies and new regulatory regimes are especially important.

The Japanese economy has been particularly successful in adapting the new paradigm to the needs of some other industrial sectors, especially mechanical engineering, vehicles and now construction. Kodama (1985, 1986) points out that the expression 'Mechatronics' was first coined in Japan in 1975 and that even before that in 1971 an explicit policy designed to induce 'fusion' was initiated with the 'Law on Temporary Measures for the Development of Specific Machinery and Electronics Industries', which spoke of 'consolidation of machinery and electronics into one', and 'systematisation' of them.

The successful Japanese development of robotics, CNC machine tools and flexible manufacturing systems has also owed a great deal to the group structure of Japanese industry. The collaboration of the leading Japanese vehicle firms with 'their own' robotics suppliers and similar collaboration with electronic firms in the same groups is one obvious instance.

In the long run perhaps even more significant is the Japanese infrastructural policy with respect to telecommunications. Both before and since its privatisation the telephone utility (NTT) has pursued its programme for an 'Integrated Network System' (INS). This is envisaged as ultimately providing a connecting web for information services to private and industrial users throughout the country. It will 'put a digital broad-based infrastructure in place in anticipation of its uses, while simultaneously developing those uses through model programmes and pilot projects' (Borres *et al.*, quoted in Arnold and Guy, 1986, p.97).

This appears to be a bolder strategy than that of most other OECD countries, which is largely based on responding to current new business opportunities and is therefore more conservative in the scale and scope of new infrastructural telecommunications investment. The INS concept 'represents a discontinuous shift in the character of telecommunications networks' since it will enable broad-bandwidth services using video signals or other forms of very rapid data transmission to use the same network as telephone services. It will ultimately permit a wider range of new information service provision, linking computer and telecommunication networks as a digital technology. The Japanese strategy therefore tackles the key strategic problems of diffusing information technology both to the service and manufacturing industries and to household consumers. This is not simply 'deregulating' and 'privatising'. It is an important step in establishing a new 'regulation regime' as well as a new infrastructure.

As Melody (1986) has forcefully pointed out, in all OECD countries during the 'era of "deregulation" more regulatory activity has been generated than ever existed in the so-called "regulation" era'. National and international policies will have an enormous influence on the provision of new services, on technical standards, on terminal and systems interconnection, on value-added network services (VANS), on the radio frequency allocation and on prices. Japanese policy has clearly recognised the fundamental strategic importance of this regulation underpinning the whole change of techno-economic paradigm in the same way as an electric power distribution system, a highway network or a railway network in previous periods.

In the case of the Fordist mass-production paradigm, as Yamauchi (1986) insists, Japan was two decades behind in the 1950s but had the great advantage of following a production system already well established in the United States and elsewhere, and improving upon it. In the case of information and communication technology, Japan is among the leaders. This was already evident in consumer electronics in the 1970s and is now increasingly apparent in capital goods, manufacturing systems, information systems and infrastructure.

Conclusions

In this context the organising and energising role of the Japanese forecasting system is important. The 'Visions' of the future produced by STA, MITI, NIRA and other government and private sources do not pretend to be accurate predictions, nor do they commit companies to inflexible plans. They chart the broad direction of advance for the economy and for technology and give companies sufficient confidence in this vision to make their own long-term investments in research, development, software, equipment and training. In this respect technological forecasting plays a role similar to that of project evaluation in sophisticated research-intensive companies. Nobody believes that it is possible to eliminate uncertainty, but a thorough discussion within the firm of the range of probabilities and alternative strategies serves to mobilise resources, to expose difficulties and bottlenecks, and above all to energise the participants, secure consensus and heighten awareness.

It may be that this consensus tends to diminish the pluralism and range of evolutionary alternatives which Nelson distinguishes as so important in the US system. Nevertheless, the contemporary success of the Japanese national system of innovation in catching up and moving ahead in information technology confronts other countries with a formidable challenge and has had severe disequilibrating effects in the world economy (Chapter 3). But catching up with Japan is certainly not an impossible task. Nor does this mean simply imitating Japan. Potentially there are a variety of different social and institutional alternatives: Japanese solutions are certainly not perfect. In some respects Sweden has found equally good or better solutions.

The Swedish example is of special interest because this is a case of a small European country with fairly limited resources which is nevertheless among the leading countries in the production and the diffusion of robotics, in the design and manufacture of telecommunication equipment and generally in computer applications. Sweden's successful diffusion of ICT has been achieved whilst maintaining excellent social services, a rather high degree of consultation with trade unions and safeguards for civil liberty. Swedish industry has made particular efforts to keep in touch with Japanese developments and in general to take the best from world technology. Sweden was also committed fairly early on towards giving ICT a high priority and probably has the most advanced training and retraining system in Europe. This clearly demonstrates the feasibility of catching up with Japan and perhaps of doing better.

In the case of the US national system of innovation analysed by Nelson in Chapter 15, there are clearly some important similarities and some major differences. The strength of the Japanese challenge in world markets has already led to determined effort in the United States to recover the US technological lead in many industries and services. The US automobile industry, for example, is striving in a variety of ways to assimilate Japanese organisational and technical innovations and emulate its success. At a more general level there is increasing evidence of the emergence of a national technology policy with respect to the semiconductor and computer industries. As Nelson observes, a whole number of initiatives for joint government-sponsored R & D programmes linking industry and universities have been taken in the 1980s. However, a distinguishing feature of the newly emerging technology and industrial policies in the United States is that they are mainly conducted under the auspices of defence agencies such as DARPA (Arnold and Guy, 1986). This limits their effectiveness and their strategic objectives in terms of the performance of the civil economy. As we have seen, the priority given to strategic defence objectives in Japan in the 1930s delayed the assimilation of the most advanced civil mass-production technology (Yamauchi, 1986).

Another significant difference between the Japanese and US national systems of innovation lies in the area of fundamental research. The United States has clearly been the world leader in basic scientific research since the Second World War and, as Nelson's chapter has shown, the interaction between this (largely university-based) research and the private industry-financed applied research and development has been an extremely fruitful source of radical innovation in the post-war period, although the intensity of this interaction varies very much by industry and field of sciences.

It is sometimes suggested that the relative weakness of Japanese fundamental research by comparison with the United States may now prevent the Japanese economy from further strengthening its technological competition. This is indeed an important topic of debate and policy-making in Japan and some of the most powerful leading companies have taken major decisions in the 1980s to build up their basic research. However, there is widespread agreement with Nelson's thesis that successful fundamental research probably flourishes most in an environment which stimulates controversy and pluralism, and when is conducted mainly in universities and publishes in open scientific literature. Lundvall (Chapter 17) has also shown how important are the interactions between 'users' and 'producers' in the national system of innovation. These links are particularly strong in the Japanese system.

But while there would be much agreement in Japan with Nelson's thesis on the advantages of the United States (and Europe) in fundamental university research, there are two qualifications which should be taken into account in considering the performance of the two systems.

The first is that the contribution of Japanese universities both to world scientific research and to Japanese industrial innovation is often underestimated. Whilst it is certainly true that Japan's contribution to world science is relatively much less significant than her contribution to world technology, it is by no means negligible and, more importantly, it is increasing rapidly. Even in the 1940s, 1950s and 1960s, Japanese universities were making major contributions which were an essential basis for the development of new technologies in Japanese industry. For example, Tokyo University played a leading part in the work which ultimately led to the success of the Japanese robotics and computer development programmes. There is much more interaction between Japanese universities and industry than is commonly realised in the West even though the networks operate under different financial arrangements than those customary in the United States. It would have been impossible for Japanese industry to succeed in mastering new technologies without a fairly strong interaction with university basic research in some ways very similar to that described by Nelson.

The second qualification is that it is not actually essential to be the world leader in basic science in order to be a world leader in technology. It is necessary to have a strong capability in basic research in order to assimilate and advance most important new technologies today, but since world scientific literature is an open literature and since world-wide exchange of ideas is still a characteristic of the world's scientific community, it is not essential to *lead* in science in order to develop and exploit new technologies in advance of competitors, as the United States itself demonstrated in the latter part of the nineteenth century.

References

Allen, G.C. (1981), 'Industrial policy and innovation in Japan', in C. Carter (ed.), *Industrial Policy and Innovation*, London, Heinemann, pp. 68–87.

Altschuler, A., Anderson, M., Jones, D. and Womack, J. (1985), *The Future of the Automobile: The Report of MIT's International Automobile Program*, Cambridge, Mass., MIT Press.

Aoki, M. (1985), 'The Japanese firm in transition', Japan Political Economy Research Committee Conference, Department of Economics, Stanford Center for Economic Policy Research, Stanford University, Publication No. 39, mimeo.

—— (1986), 'Horizontal vs vertical information structure of the firm', *American Economic Review*, vol. 76, no. 5, pp. 971–83.

Arnold, E. and Guy, K. (1986), *Parallel Convergence: National Strategies in Information Technology*, London, Frances Pinter.

Baba, Y. (1985), 'Japanese colour TV firms: decision-making from the 1950s to the 1980s: oligopolistic corporate strategy in the age of micro-electronics', D.Phil. dissertation, University of Sussex.

Burns, T. and Stalker, G.M. (1961), *The Management of Innovation*, London, Tavistock.

Dore, R. (1985), 'The social sources of the will to innovate', Papers in Science, Technology and Public Policy, no. 4, Imperial College Lecture, London.

Ergas, H. (1984), 'Why do some countries innovate more than others?', Centre for European Policy Studies, Brussels.

Freeman, C. (ed.), (1986), *Design, Innovation and Long Cycles in Economic Development*, London, Frances Pinter.

—— (1987), *Technology Policy and Economic Performance: Lessons from Japan*, London, Frances Pinter.

Freeman, C. and Perez, C. (1986), 'The diffusion of technical innovation and changes of techno-economic paradigms', paper prepared for the Conference on Innovation Diffusion, Venice, DAEST, March 1986.

Fukasaku, Y. (1987), 'Technology imports and R & D at Mitsubishi Nagasaki Shipyard in the pre-war period', *Bonner Zeitschrift für Japanologie*, vol. 8, pp. 77–90.

Goto, A. (1982), 'Business groups in a market economy', *European Economic Review*, pp. 53–70.

Gregory, G. (1985), *Japanese Electronics Technology: Enterprise and Innovation*, New York, John Wiley.

Imai, K. and Itami, H. (1984), 'Inter-penetration of organisation and market: Japan's firm and market in comparison with US', *International Journal of Industrial Organisation*, pp. 285–310.

Irvine, J.H. and Martin, B.R. (1984), *Foresight in Science Policy: Picking the Winners*, London, Frances Pinter.

Johnson, C. (1982), *MITI and the Japanese Miracle: The Growth of Industrial Policy, 1925–1975*, Stanford, Stanford University Press.

Jones, D.T. (1985), 'Vehicles', Chapter 2 in C. Freeman (ed.), *Technological Trends and Employment: 4: Engineering and Vehicles*, Aldershot, Gower.

Kodama, F. (1985), 'Mechatronics technology as Japanese innovation: a study of technological fusions', Graduate School for Policy Science, Saitama University, Tokyo.

—— (1986), 'Japanese innovation in mechatronics technology', *Science and Public Policy*, vol. 13, no. 1, pp. 44–51.

Koshiro, K. (1985), 'Japan's industrial policy for new technologies', Centre for International Trade Studies, Yokohama National University, Yokohama.

Kuwahara, Y. (1985), 'Creating new jobs in high technology industries', in OECD (1985), *Employment Growth and Structural Change*, Paris.

McCormick, K. (1985), 'The flexible firm and employment adjustment: fact and fable in the Japanese case', Unit for Comparative Research in Industrial relations, University of Sussex, mimeo.

Melody, W.H. (1986), 'Telecommunication policy: directions for the technology and information services', *Oxford Surveys in Information Technology*, vol. 3, pp. 77–106.

Patel, P. and Pavitt, K. (1987), 'Is Western Europe losing the technological race?', *Research Policy*, vol. 16, nos. 2–3–4, pp. 59–86.

Pavitt, K. (1985), 'Technology transfer among the industrially advanced countries: an overview', in N. Rosenberg and C. Frischtak (eds.), *International Technology Transfer: Concepts, Measures and Comparison*, New York, Praeger, pp. 3–23.

Peck, M.J. (1986), 'Joint R & D: the case of microelectronics and computer technology cooperation', *Research Policy*, vol. 15, no. 5, October, pp. 219–31.

Peck, M.J. and Goto, A. (1981), 'Technology and economic growth: the case of Japan', *Research Policy*, vol. 10, pp. 222–43.

Peck, M.J. and Wilson, R. (1982), 'Innovation, imitation and comparative advantage: the case of the consumer electronics industry', in H. Giersch (ed.), *Emerging Technology: Consequences for Economic Growth, Structural Change and Employment in Advanced Open Economies*, J., Mohr, Tübingen, J. Mohr, pp. 195–212.

Perez, C. (1983), 'Structural change and the assimilation of new technologies in the economic and social system', *Futures*, vol. 15, no. 4, October, pp. 357–75.

—— (1985), 'Micro-electronics, long waves and world structural change: new perspectives of developing countries', *World Development*, vol. 13, no. 3, pp. 441–63.

Prais, S.J. (1987), 'Educating for productivity: comparisons of Japanese and English schooling and vocational preparation', *National Institute Economic Review*, no. 119, February, pp. 40–56.

Sakurai, H. (1986), 'Metamorphosis of the mining company: a case study of Mitsubishi Metal Corporation', International Symposium on Japanese Technology, Saitama University, Tokyo, mimeo.

Schumpeter, J.A. (1928), 'The instability of capitalism', *Economic Journal*, pp. 361–86.

Sciberras, E. (1981), *Technical Innovations and International Competitiveness in the Television Industry*, Omega, pp. 585–596.

Shinohara, M. (1982), *Industrial Growth, Trade and Dynamic Patterns in the Japanese Economy*, Tokyo, University of Tokyo Press.

STA (1972), *Science and Technology Developments up to AD 2000*, Science and Technology Agency, Tokyo, Japan Techno-Economics Society.

Sutton, A.C. (1971), *Western Technology and Soviet Economic Development*, 3 volumes, Hoover Institute, Stanford.

Tamura, S. (1986), 'Reverse engineering as characteristic of Japanese industrial R and D activities', International Symposium on Japanese Technology, Saitama University, Tokyo.

Tanaka, M. (1978), 'Industrialisation on the basis of imported technology: a case study of the Japanese heavy chemical industry, 1870–1930', M.Phil. thesis, Science Policy Research Unit, University of Sussex.

Vogel, E.F. (1980), *Japan as No.1*, Tokyo, Tuttle.

Williamson, O.E. (1975), *Markets and Hierarchies: Analysis and Anti-Trust Implications*, New York, Free Press.

Yakushiji, T. (1986), 'Technological Emulation and Industrial Development', paper presented at the Conference on Innovation Diffusion, Venice, DAEST, March 1986.

Yamauchi, I. (Nomura Research Insitute) (1986), 'Long range R and D', in C. Freeman (ed.), *Design, Innovation and Long Cycles in Economic Development*, London, Frances Pinter, pp. 169–85.

17 Innovation as an interactive process: from user–producer interaction to the national system of innovation

Bengt-Åke Lundvall

Institute for Production, Aalborg University, Aalborg

Introduction

In this chapter the focus is upon the interactive aspects of the process of innovation.[1] The analysis takes as its starting points two important characteristics of an industrial economy: the highly developed vertical division of labour and the ubiquitous and all-pervasive character of innovative activities. It follows that a substantial part of innovative activities takes place in units separated from the potential users of the innovations.[2]

Here we shall argue that the separation of users from producers in the process of innovation, being 'a stylized fact' of a modern industrial society (capitalist or socialist), has important implications for economic theory. When we focus upon innovation as an interactive process, theoretical and practical problems tend to present themselves differently than in mainstream economic theory.

The interactive aspects of the process of innovation can be studied at different levels of aggregation. In the first part of the chapter we discuss 'the microeconomics of interaction'. In the second part we present some preliminary ideas on how a model of a national system of innovation can be developed.

The micro-foundation: interaction between users and producers

In standard microeconomics the agents—firms and consumers—are assumed to behave as maximizers of profits and utility. Perfect competition with numerous buyers and sellers, the flow of information connecting them, encompassing nothing but price signals, is the normative and analytical point of reference of the theory. Monopolistic structures and complex client relationships are regarded as deviations from this normal and ideal state.

The kind of 'microeconomics' to be presented here is quite different. While traditional microeconomics tends to focus upon decisions, made on the basis of a given amount of information, we shall focus upon a *process of learning*, permanently changing the amount and kind of information at the disposal of the actors. While standard economics tends to regard optimality in the allocation of a given set of use values as the economic problem, *par préférence*, we shall focus upon the capability of an economy to produce and diffuse *use values with new characteristics*. And while standard economics takes an atomistic view of the economy, we shall focus upon the *systemic interdependence* between formally independent economic subjects.

Product innovations in a pure market?

In an economy characterized by vertical division of labour and by ubiquitous innovative activities, a substantial part of all innovative activities will be adressed towards users, outside the innovating units. In such an economy successful innovations must be based upon knowledge about the needs of potential users, and this knowledge is as important as knowledge about new technical opportunities (Freeman, 1982, p. 124, *passim*). When an innovation has been developed and introduced, it will diffuse only if information about its use value characteristics are transmitted to the potential users of the innovation. Within organizations and firms, this constitutes an intra-organizational problem, to be solved through interaction and information exchange, involving different individuals and departments belonging to the same organization.

Here, however, the focus will be upon those innovative activities which are oriented towards new products to be presented to a market. For simplicity, we shall label such innovations 'product innovations', keeping in mind that they might constitute new materials and new process equipment, as well as new consumer products. Further, we shall not primarily treat innovations as single events. By using terms such as 'the process of innovation' and 'innovative activities', we indicate that the traditional separation between discovery, invention, innovation and diffusion might be of limited relevance in this specific context.[3]

How can the mutual information problem be solved when the producer and the user are separated by a market? If the market is 'pure', in the neo-classical sense, the problem must remain without a solution. In such a market the only information exchanged relates to products already existing in the market and it contains only quantitative information about price and volume. Anonymous relationships between buyer and seller are assumed. In such a market the innovating units as well as the potential users will operate under extreme uncertainty. Producers have no information about potential user needs and users have no knowledge about the use-value characteristics of new products. If the real economy was constituted by pure markets, product innovations would be haphazard and exceptional.

It is interesting to note that the pure market—hailed by some neo-classical economists for its ability to establish an efficient allocation of resources on the basis of very limited amounts of information—forms an environment hostile to innovative activities, and that product innovations would be all but absent in a capitalist economy characterized by perfect competition. At an abstract level, a socialist economy would be expected

to overcome this crucial information problem more easily through a planning mechanism, taking into account the need for the exchange of qualitative information. According to a recent study of innovations in the Soviet Union, however, the lack of efficient user–producer interaction seems to be a major problem in the 'real existing socialist countries' (Amann and Cooper, 1982).

Anne P. Carter (1986) has recently pointed to the neglect of product innovations in production models as a general and serious weakness. But this neglect might be said to be fully consistent with the microeconomic assumption of pure markets as the norm. In a world where all products were characterized by constant use-value characteristics, pure markets could survive, and those pure markets would tend to reproduce the existing set of use values. Introducing product innovations into economic models cannot but erode the traditional concept of the pure market.

Product innovations and transaction costs

One well-established alternative conception of the process of exchange is the transaction cost approach presented by Oliver E. Williamson (1975). What are the implications of product innovations if we take this approach as our point of departure? According to Williamson, markets characterized by small numbers, uncertainty, limited rationality and opportunistic behaviour will tend to become hierarchies. High transaction costs will induce vertical integration. A market where product innovations were frequent would involve true uncertainty at both sides of the market, the uncertainty emanating not from the external conditions for transaction but from qualitative change in the commodity itself. It would also imply what Williamson calls 'informational impactedness'—an uneven distribution of information. The innovating unit would, typically, have much more, and more certain, information about the use value characteristics of the new product than the potential user.

In the Williamson framework, as in the neo-classical world, we would expect product innovations to be exceptional. They should become internalized and transformed into process innovations through vertical integration.

It is, of course, quite difficult to measure the proportion of innovative activities directed towards product innovations in the sense of the concept used here. One of the few systematic innovation data banks is the one developed at the Science Policy Research Unit, Sussex University. Among the more than 2,000 important post-war innovations reported in Pavitt (1984), more than a half were developed for outside firms (ibid., p. 348). OECD data on the allocation of R & D activities confirm that product innovation is as important a phenomenon as process innovation in the OECD area.

Thus neither standard microeconomics nor the original transaction cost approach are easily reconciled with the stylized facts of a modern industrial economy. In order to explain the actual importance of product innovations we must take a closer look at the (assumed) market–hierarchy dichotomy.

The organized market as a solution?

If all transactions in the real world took place either in 'pure markets' or in 'pure organizations', innovative activities would be less frequent than they are, and they would mainly take the form of process innovations. The fact that product innovations are frequent in the real world demonstrates that most real markets are 'organized markets' rather than pure markets. The actually observed relative efficiency of the capitalist system, in terms of innovative behaviour, can only be explained by the fact that the invisible hand of the pure market economy has been replaced by bastard forms, combining organization elements with market elements.

The organized market is characterized by transactions between formally independent units and by a flow of information on volume and price. But it also involves relationships of an organizational type. Those relationships might involve flows of qualitative information and direct cooperation. They might take a hierarchical form, reflecting that one party dominates the other, by means of financial power or of a superior scientific and technical competence. As we shall see, a purely hierarchical relationship will, however, often prove insufficient. Mutual trust and mutually respected codes of behaviour will normally be necessary in order to overcome the uncertainty involved.[4]

User–producer interaction in the process of innovation

We shall now take a closer look at the specific forms of user–producer interaction in relation to the process of innovation. The producer will have a strong incentive to monitor what is going on in user units. First, process innovations within user units might be appropriated by producers or represent a potential competitive threat. Second, product innovations at the user level may imply new demands for process equipment. Third, the knowledge produced by learning-by-using can only be transformed into new products if the producers have a direct contact to users. Fourth, bottlenecks and technological interdependencies, observed within user units, will represent potential markets for the innovating producer. Finally, the producer might be interested in monitoring the competence and learning potential of users in order to estimate their respective capability to adopt new products.

The user, on the other hand, needs information about new products, and this information involves not only awareness but also quite specific information about how new, use-value characteristics relate to her/his specific needs. When new user needs develop—for example, when bottleneck problems occur—the user might be compelled to involve a producer in the analysis and solution of the problem. This can only be done successfully if the user has a detailed knowledge about the competence and reliability of different producers.

When complex and specialized equipment is developed and sold to users, there will be a need for *direct cooperation* during the process of innovation. The cooperation is not a single act but takes place at different stages of the process (Rothwell and Gardiner, 1985). First, the user may present the producer with specific needs to be fulfilled by the new product. Second, the producer might install it and start it up in cooperation with the user. At this stage, the producer might offer specific training to the user. After the product has been adopted there might follow a period where the producer would have obligations regarding repair and updating of the equipment.

The uncertainty involved in this kind of transaction will be considerable. Not only is the user buying a product with unknown characteristics. He is also buying the cooperation of an external party for a future period. It should be obvious that the room for an opportunistic producer to cheat is considerable. Conversely, this implies that 'trustworthiness' becomes a decisive parameter of competition. If a user has a choice between a producer known for low-price and technically advanced products, but also for having a weak record in terms of moral performance, and one well known for trustworthiness, the first will be passed by. This implies limits to opportunistic behaviour. Those limits are reinforced when users pool their information about the reliability of different producers.

The exchange of information between user and producer also involves uncertainty and room for cheating and disloyal behaviour. The user must disclose her/his needs to the producer in order to get workable solutions. The producer has an interest in disclosing the full capacity of his product and in giving the user insight into his technical competence as a potential cooperator. But in both cases a full disclosure might be abused by the other party. Information might be spilled to competitors and each party may invade the market of the other party. Again, the abuse can only be restrained if codes of behaviour and mutual trust form an element of the relationships. Without any such restraints, transaction costs would become prohibitive and vertical integration would become a necessary outcome.

How strong is the element of organization?

The element of organization might be quite weak in certain markets. If the product is simple, its use-value characteristics changing but slowly, and the expenditure for its procurement forms a negligible part of the user's budget, the market might become quite 'pure'. When its use-value characteristics are changing rapidly, are complex and the product is expensive, the element of organization will be strong. The former type of goods will, typically, be developed by the producer alone and bought 'off the shelves', while the latter will be developed in an interaction between the user and producer, and the act of exchange will involve direct cooperation and exchange of qualitative information.

The flow of information

In markets where the element of organization is strong, the flow of information might be analysed in terms parallel to those applied in the theoretical analysis of pure organizations. Here we shall use some elements from a conceptual framework developed by Kenneth Arrow (1974). The flow of information can only take place if there exist *channels of information* through which the message can pass. Further, a *code of information* is necessary in order to make the transmission of messages effective. The establishment of channels of information may, according to Arrow, be regarded as parallel to a process of investment in physical capital. It is a time-consuming process involving costs. The development of a common code is also time-consuming and involves learning. The more the code is used in transmitting information, the more effective it becomes. 'Learning-by-interacting' increases the effectiveness of a given set of channels and codes of information.

The selectivity of user–producer interaction

The organizational element will not link every single producer to every single user—here we disregard pure monopolistic and pure monopsonistic situations. Normally, each producer will have a close interaction with a subset of all potential users and each user will be attached to only one, or a small subset of all potential producers. This selectivity reflects the need to develop non-economic relationships of hierarchy and mutual trust. It also reflects the need to develop effective channels and codes of information.

User–producer relationships in time

It takes time to develop selective relationships involving elements of hierarchy and mutual trust. It also takes time to develop effective channels and codes of information. Once those relationships have become established, it will not be cost-less to sever the connections. Inertia—a general resistance to change and risk aversion—combines with rational motives in reinforcing existing user–producer relationships. *Ceteris paribus*, the user will prefer to trust producers, known from her/his own experience, rather than getting involved with a new producer. The investment in information channels and codes will be lost if the old relationships are severed and new investment in the creation of new relationships will be required. Therefore user–producer relationships will tend to become enduring and resistant to change. Only if the costs of keeping the existing relationships going are apparent, or the economic incentives offered by new relationships are substantial, will a reorganization of the markets take place.

User–producer relationships in space

The user–producer relationship is defined in 'economic space' coupling units, close to each other, in an input–output system. The selective user–producer relationships will involve units more or less distant from each other in geographical and cultural space. The importance of distance will

vary with the type of innovative activity involved. When the technology is standardized and reasonably stable, the information exchanged may be translated into standard codes, and long-distance transmission of information can take place and involve low costs. Here, user–producer relationships involving units located far away from each other might be effective.

When the technology is complex and ever changing, a short distance might be important for the competitiveness of both users and producers. Here, the information codes must be flexible and complex, and a common cultural background might be important in order to establish tacit codes of conduct and to facilitate the decoding of the complex messages exchanged. The need for a short distance will be reinforced when user needs are complex and ever changing.

When the technology changes rapidly and radically—when a new technological paradigm (for a discussion and a definition, see Dosi, 1982) develops—the need for proximity in terms of geography and culture becomes even more important. A new technological paradigm will imply that established norms and standards become obsolete and that old codes of information cannot transmit the characteristics of innovative activities. In the absence of generally accepted standards and codes able to transmit information, face-to-face contact and a common cultural background might become of decisive importance for the information exchange.

Vertical integration as a means of overcoming geographical and cultural distance

The development of transnational capital and of vertically integrated firms operating all over the world reflects that 'organizational proximity' may overcome geographical and cultural distance. But vertical integration may have its price. It tends to exclude integrated units from the interaction with producer units and user units outside the integrated firm. Such independent firms will tend to guard themselves against an open information exchange with a vertically integrated unit. As users, they risk to get less efficient technology than their integrated counterpart and competitor. As producers, they fear that the know-how built into their product innovations will become expropriated by the integrated user and transferred to an integrated competing producer.

Also, the vertically integrated units may prove to be more rigid and less susceptible to new technical opportunities and new user needs than the parties operating in an organized market. The tendency towards vertical integration is strong, but there are also certain counter-tendencies at work. The trade-off between saved transaction costs and the loss in terms of a more narrow interaction with external parties will differ between different parts of the economy. It will, among other things, reflect the state of the technology and the character of the process of innovation.

User and producer characteristics and the innovative potential of interaction

Not all user–producer relationships promote innovative activities. Being closely linked to conservative users having weak technical competence might be a disadvantage for a producer, and vice versa. The innovativeness and the competence of users and producers are important qualities which might stimulate the other party. The degree of standardization among users might also be important. Being dependent upon a set of users with very diversified needs might make it difficult for the producer to accumulate experience and to exploit scale economies.

The effectiveness of the user–producer relationships grows with time. As a subset of users and producers gets more experience from interaction, the elements of hierarchy and mutual trust are strengthened and the exchange of information becomes more open. The code of information becomes more effective in transmitting complex messages related to the process of innovation. As we shall see below, this 'effectiveness' does not, however, guarantee *efficiency* if the criterion is user satisfaction at a low cost. The negative side is inertia and resistance to change.

'Unsatisfactory innovations'

Traditional welfare economics tends to disregard innovative activities. It analyses the allocation of a given set of use values with given characteristics. Nor are the concepts used easily adapted to a normative analysis of the process of innovation. There is no point in asking how actual innovations deviate from 'an optimum'. Innovations not yet conceived are not known to us, and therefore we do not have any well-defined points of reference for such an analysis.

In certain instances it might, however, be possible to demonstrate how innovative activities and technological trajectories deviate systematically from user needs. When deviations cannot be ascribed either to a lack of technical opportunities or to an unwillingness among users to pay the costs for an adaption to the user needs, we might label the innovations 'unsatisfactory'.

When the user–producer relationships are characterized by a strong dominance of producers in terms of financial strength and technical competence, such deviations become more likely. In the field of consumer goods the producer dominance is very accentuated. The producer organizes both the process of innovation and the information exchange with users. In this field we should expect 'unsatisfactory innovations' to be frequent (Freeman, 1982, p. 202ff). A pattern of dominance and hierarchy might be found also when the user is a professional organization. If a few big firms produce scientifically based, complex and systemic products for a great number of small, independent user units—each with a low technical and scientific competence—producers will dominate the process of innovation and the likelihood of unsatisfactory innovations becomes great. In a study of the Danish dairy industry, such a pattern, resulting in 'hyper-automation', was found to characterize the relationships between producers and users of dairy equipment (Lundvall *et al.*, 1983).

In such situations coordination among users might develop and resources might be pooled in order to develop a counter-competence. Such a co-ordination will often be more difficult to make efficient when the users are consumers than when they are professional units. Government regulation or government support to user organizations might be necessary in order to rectify an unsatisfactory trajectory in consumer technology.

Another background for unsatisfactory innovations might be inertia in user–producer relationships and the 'effectiveness' of already established channels and codes of information. In a historical period characterized by the development and introduction of basic radical innovations, the rigidity of the existing set of user–producer relationships tends to become manifest. A basic radical innovation will often be produced by a new sector with weak forward linkages. The potential users of the innovation will be found in most parts of the economy, and those users will have backward linkages to producers, having little experience and competence in relation to the new technology. Existing user–producer networks will prove to be tenacious and it will take considerable time for a new network to become established. During such a period of transition, productivity might be stagnating, while new technological opportunities seem to flourish.

Here, the problem is not only specific unsatisfactory technical innovations, but rather a general 'mismatch' in the whole economy. Christopher Freeman and Carlota Perez (1986) have discussed how a 'technological revolution', based upon information technology, might provoke mismatch problems related not only to capital and labour but also to the existing socio-economic institutional set-up. The rigidity of user–producer relationships might be regarded as one important aspect of this last type of mismatch. It is important because it has its roots in the very core of the market system, in markets producing innovations. Policy strategies, putting all the emphasis upon flexibility through the market mechanism and minimizing the role of government in the process of adjustment, seem to be somewhat off the point when rigidities are produced and reproduced within the markets themselves.

Is innovation induced by supply or by demand?

One of the classical disputes in innovation theory refers to the role of demand and supply in determining the rate and direction of the process of innovations (Mowery and Rosenberg, 1979; Freeman, 1982, p. 211). The user–producer approach puts this question in a new perspective. On the one hand, it demonstrates that demand does play an important role in the process of innovation. On the other hand, it puts the emphasis more upon the *quality of demand* than upon demand as a quantitative variable. The very substantial user expenditure channelled into the demand for private transportation has not resulted in radical product innovations in the automobile industry. Conversely, very competent and demanding users have provoked radical innovations in areas where the volume of expenditure has been miniscule. The role of users in relation to the development of new scientific instruments is illustrating in this respect.

Individual innovations might appear as unrelated to user needs, such as innovations emanating from science. In the second part of this chapter it will be argued that even science has its users and that many innovations, appearing as purely supply-determined, have their roots in a user–producer interaction placed early in the chain of innovation. In this perspective *general* statements about the role of 'demand' and 'supply' do not seem very relevant.

Some implications for industrial and technology policy

The fact that technology is influenced by the demand side has been used to argue for a *laissez-faire* technology policy. If demand is provoking the innovations called for, there is no need for state intervention. Those arguing that the supply side plays the dominating role will often recommend government support to R & D activities and education, combined with an active manpower policy. The implications of a user–producer approach are somewhat more complex.

First, technology policy should take into account not only the competence and innovativeness of units placed early in the chain of innovation. The lack of competence of users and the tendency of producers to dominate the process of innovation might be as serious a problem as a lack of competence on the producer side. Even when the state itself acts as a user, one will often find that the competence will be too weak and this might result in 'unsatisfactory innovations'. Two Danish case studies, looking into the role of local government as user of waste-water technology and office technology, demonstrated how a lack of local user competence had a negative effect upon the systems developed and used (Gregersen, 1984; Brængaard *et al.*, 1984).

Second, government may intervene, directly or indirectly, in relation to the establishment and restructuring of patterns of user–producer relationships. In a period characterized by gradual technical change and incremental innovations, a national government might sustain national and international user–producer linkages which already exist. It might also support the establishment of specific organisations, intermediating between groups of users and groups of producers, pooling information, and thereby stimulating the production and diffusion of innovations.

In a period characterized by radical innovations and a shift in technological paradigm, the task of government becomes vastly more complex and important. In such a period, there is a need for a transformation of the existing network of user–producer relationships. The inertia originating in the organized markets will at the national level be supported by the political power of strong interest groups, closely associated with the prevailing structure. The difficult task for government will be to stimulate the renewal, or severance, of well-established user–producer relationships and the establishment of new relationships.

Standard microeconomics and the user–producer approach

Some of our results can now be confronted with the kind of microeconomic theory presented in standard textbooks. We make the following observations:

- The element of organization will be different, in terms of content and strength, between different markets and it will change over time. Some markets will be more susceptible to an analysis based upon the concepts of optimizing agents acting at arms-length distance than others. This raises some doubt about the intentions to construct one single model of micro-behaviour, assumed to be generally valid for all markets—a problem discussed by Kornai (1971, p. 207ff).
- The standard approach will be most relevant when technological opportunities and user needs remain constant. When product innovations are continuously provoked by changing technological opportunities and users needs, it is no longer meaningful to assume optimizing behaviour. 'Short-run' decisions, by producers to become involved in certain lines of innovating activities, and by users to choose among new products, will be characterized by true uncertainty, as will, *a fortiori*, 'long-run' decisions, referring to the establishment of (and investment in) new relationships and information channels.
- Standard microeconomics regards technical change as an exogenous process and its outcome as technical 'progress', indicating growing efficiency. In organized markets the existing set of user–producer relationships may produce technological trajectories, deviating systematically from what is 'satisfactory', even when users and producers act according to profit motives.
- In standard microeconomics, changes in relative prices will influence the decisions taken by users and producers automatically and instantaneously. A world characterized by organized markets will be sluggish in this respect. The existing set of user–producer relationships and the continuous qualitative change in products will reduce the responsiveness to changes in relative prices.

National systems of innovation

In the first part of this chapter, we found that the microeconomic framework, as presented in standard textbooks, is not easily reconciled with certain stylized facts of the modern economy. A highly developed vertical division of labour, when combined with ubiquitous innovative activities, implies that most markets will be 'organized markets' rather than pure markets. In this second, and final, part we shall sketch some of the implications of our micro-approach for the national and international level. Elements of a model of a national system of innovation will be introduced.

The subdisciplines in economics most relevant in this context are theories of economic growth and international trade. Standard growth models are developed under the assumption of a closed economy. This is a natural assumption in so far as the models regard new technology as falling 'as manna from heaven' and as equally accessible for all actors, sector, regions and nations. Standard foreign-trade theory assumes labour and capital to be perfectly immobile and commodities to be perfectly mobile across national borders. It has the assumption of perfectly free and mobile technology in common with standard growth theory.

This last assumption is at odds with what can be observed in the real world, where some countries establish themselves as technological leaders, generally or in specific technologies, while others tend to lag behind. According to the user–producer approach, geographical and cultural distance is a factor which may impede the interaction between user and producers. This might contribute to an explanation of why different national systems display different patterns of development.

The nation as a framework for user–producer interaction

The tendency towards internationalization of trade, capital and production has been strong during the post-war period. Some would even argue that nations tend to become obsolete as economic subjects. But this process of internationalization has not wiped out idiosyncratic national patterns of specialization in production and international trade. The fact that Denmark is strongly specialized in dairy machinery, Sweden in metal-working and wood-cutting technology, and Norway in fishery technology cannot be explained by the general factor endowments in those countries. Rather, we should look for the explanation in the close interaction between producers of such machinery and a competent and demanding domestic user sector (Andersen *et al.*, 1981).

Interaction between users and producers belonging to the same national system may work more efficiently for several reasons. Short geographical distance is part of the explanation; more important may be a common language and the cultural proximity. It is thus interesting to note that firms in the Nordic countries tend to regard all the Nordic countries as their 'home market'. This might reflect that those nations have very much in common in terms of culture and social organization (Dalum and Fagerberg, 1986).

Another factor of importance is, of course, national government. The role of government in relation to the process of innovation has been seriously underestimated according to recent historical studies (Yakushiji, 1986). Besides more direct interventions in relation to specific innovations, government imposes standards and regulations, making domestic interaction more efficient. In important instances the state intervenes directly in the network and supports existing user–producer relationships.

The fact that national economies have idiosyncratic technological capabilities reflects that international transfers of technology is neither cost-less nor instantaneous. Some parts of knowledge can be embodied in traded commodities, while other parts are embodied in the labour force. The limited mobility of labour across national borders can partly explain why technology is not easily transferred internationally. The structure of the national systems of production and innovation is a product of a historical process and it cannot be transferred as easily as 'factors of production'. It might be here that we find the most fundamental restriction to international learning and international transfer of technology.

The importance of nations as frameworks for user–producer interaction does not rule out transnational interaction, however. In some industries and technologies the required scale of the R & D effort is so enormous that not even the biggest of the transnational firms can afford to go alone when developing a new product. This is the case for civil aircraft, space technology and nuclear power. Here the pattern of user–producer interaction transcends national borders. But even in these areas, national interests related to international competitiveness and military goals put certain limits to the actual cooperation taking place, according to recent case studies (OECD, 1986).

Applying a user–producer perspective to international relations brings forward the structural interdependency, characterizing the process of innovation within and between nations. On this background we shall sketch the outlines of 'a national system of innovation'. Earlier research involving international comparisons of innovative capabilities has demonstrated important international differences at the micro level, in terms of management strategies and firm behaviour, sometimes taking into account differences in the environment of firms, financial institutions and labour relations, for example. Such studies, useful as they are, might underplay the importance of the structure of the full system of innovation, however. When the process of innovation is regarded as the outcome of a complex interaction, it is obvious that the whole system might be more than a sum of its parts.

The concept of the national system of innovation will be developed step by step, using earlier contributions on systems of production and on the division of labour within systems of innovation as some of its elements.

National systems of production

While Anglo-Saxon Industrial Economics tends to regard national economies as 'a bunch of industrial sectors', the French tradition has been more oriented towards the systemic interdependence between different parts of the economy. Verticals of production or 'filières', encompassing all stages of production from raw materials to final products, are important units of analysis in this tradition (de Bandt and Humbert, 1985). A broader concept, also bringing in public agencies and financial institutions, industrial subsystems or 'mésosystèmes industriels', has recently been developed and proposed as the units, most adequate, for industrial policy (de Bandt, 1985).

An even more ambitious approach, presented by some French marxists, and inspired by the work of François Perroux, defines 'the national system of production' as a unit of analysis. The national industrial system is divided into a small number of sections, defined by the economic function of the output and by its sector of use (investment goods, semi-manufactured goods and consumer goods) (GRESI, 1975). Some of the contributions in this tradition assume the section producing investment goods for the production of investment goods to be the strategic one for economic growth and development. National systems, having a strong position in this area, will tend to have a strong international competitiveness and vice versa. The national system of production is thus not assumed to be a closed system. On the contrary, it is the specific degree and form of openness which determines the specific dynamics of each national system of production.

Production and innovation

In order to judge the relevance of this model it is necessary to look into the relationship between the process of production and the process of innovation. These processes differ in important respects but they are also mutually interdependent.

Production is a repetitive process where routines tend to develop. The flows of goods and services between different subsystems can—if use-value characteristics remain constant—easily be quantified in terms of value and volume. The process of innovation might be continuous and cumulative, but it will always have a unique element, stressing the importance of creativity, as opposed to routine decision-making. The flows between the subsystems will be complex and systemic information, difficult to translate into quantitative terms.

The interdependency between production and innovation goes both ways. On the one hand, learning taking place in production—as 'learning-by-doing' or as 'learning-by-using'—forms an important input into the process of innovation. 'Learning-by-interacting' will, typically, take place between parties, linked together by flows of goods and services originating from production (this is a prerequisite for user–producer relationships to become enduring and selective). On the other hand, the process of innovation might be the single most important factor restructuring the system of production, introducing new sectors, breaking down old, and establishing new, linkages in the system of production.

This interdependency between production and innovation makes it legitimate to take the national system of production as a starting point when defining a system of innovation. But the division of labour in the system of innovation is not just a reflection of the division of labour in the system of production. Some parts of the production system will be more productive in terms of innovations while others primarily will be users of innovations developed by others. This is documented in some recent contributions to innovation theory.

The vertical division of labour in the national system of innovation

Most innovation studies, focusing upon vertical interaction, have put the emphasis upon the division of labour in the process of innovation. The pioneering studies of the sector producing scientific instruments, made by von Hippel (1976), demonstrated that process innovations were often developed by the sector itself. Even when independent producers were involved the users played an important active part in the process of innovation.

In Pavitt (1984), a taxonomy, referring to different types of industries according to their respective role in the process of innovation, is presented. Using a data base for important UK innovations, containing information of origin and address of each innovation, three different types of sectors were identified—supplier-dominated, production-intensive and science-based. This taxonomy and the further subdivisions made are extremely useful in defining the division of labour within the national system of innovation.

Flows and stocks in the national system of innovation

Earlier we pointed out that the flows within the system of innovation take the form of complex and systemic information—messages difficult to translate into quantities. This is also true for the stocks of the system. Knowledge, scientific as well as know-how and tacit knowledge, is difficult to measure. Other important 'stocks' may be the inventiveness and creativity of individuals and organizations and those are even more difficult to assess in quantitative terms.

In standard economics there is a strong tendency to define scientific analysis as synonymous with the establishment of quantitative and mathematical models. If we accepted this dictum, important aspects of the national system of innovation would be regarded as being outside the realm of economic science. As pointed out by Georgescu-Roegen (1971, p. 316ff), this ideal of science is not uncontroversial, however. It reflects an epistomology imported from Newtonian physics. Georgescu-Roegen demonstrates that 'dialectical concepts'—along with arithmomorphic concepts—must be a part of any science analysing change.

Further, there have been different attempts to develop a quantitative analysis of the flows within national systems of innovation. As a matter of fact, the already mentioned study by Pavitt (1984) may be regarded as a quantitative approach using the number of 'important UK innovations' as the unit of account. Another interesting contribution in this field is Scherer (1982). Here a detailed input–output matrix for the US industrial system is developed on the basis of information gathered on patenting and R & D activities.

In both of these papers it is the industrial system which is at the centre of the analysis. This is natural in so far as most innovations emanate within this system. But when we look at the system of innovation from a user–producer perspective it becomes interesting to take a closer look at the interfaces between industry and the academic community and at the interfaces between industry and some of the 'final users' of industrial innovations—workers, consumers and the public sector.

In a recent paper by Nelson (1986) the division of labour in, and performance of, the US system of innovation is discussed. It is demonstrated that universities and other public institutions involved in the production of science are important parts of this system, acting in a way which makes them complementary to the innovative activities going on in the private sector. It is obvious that any model of a national system of innovation must take into account the interaction between universities and industry.

Science and technology in a user–producer perspective

In the first part of this chapter we focused mainly upon the interaction between firms producing goods and services. The user–producer perspective might, however, be applied to early stages in the chain of innovation—basic research, applied research and developmental activities. It is almost built into the definition of 'basic research' (as non-applied) that it should take place without any specific purpose or address. This picture is too simple, however. Even pure science, as mathematics and logics, has its users, and the agenda of science will often be determined by users in applied science. Also in this area the innovativeness and competence of users may influence the rate and direction of scientific discovery. In a case study referring to Bell Telephone Laboratories, Nelson (1962) has demonstrated the close interaction between basic and applied research.

What separates pure science from technology is primarily the institutional framework. Science will, typically, be produced in universities according to an academic 'mode of behaviour', while technology primarily will be produced in private firms according to a profit-oriented 'mode of behaviour'. The academic mode will typically be characterized by non-pecuniary incentives—the 'search for excellency' will be a strong motive power (sometimes even combined with an urge to understand what is going on). The output of science will be widely dispersed because the world-wide diffusion of research results is a precondition for recognition of excellency (David, 1984). This mode of behaviour implies a different culture from the one predominating in profit-oriented firms. Norms, values and incentives are different, as well as the language and the codes of information used in the two spheres.

It is not surprising that the link between universities and industry has become a political issue. The growing recognition of the role of science in relation to technology and production has made it a national priority to strengthen this link. The flourishing of 'Silicon Valleys', characterized by a close interaction between 'excellent' universities and high-technology firms in different parts of the world, has given the debate further impetus. In most OECD countries the establishment of 'sicence parks' and 'technopolises' has become a part of industrial policy.

The efforts made to integrate and subordinate academic activities in relation to industry may not be cost-less, however. If the academic mode of production is undermined and replaced by a profit-oriented mode of behaviour, where pecuniary incentives become more important and where secrecy regarding the output becomes more frequent, the academic mode of behaviour may lose one of its principal merits—its tradition for world-wide diffusion of knowledge. In the field of biotechnology this process seems already to have reached a critical level (Chesnais, 1986). National systems of innovation may temporarily become strengthened when universities become subordinated to industry. In the long run, the production and world-wide distribution of knowledge may become weakened.

Introducing the final users of technology into the system

The classical actors in innovation studies are individual entrepreneurs and the R & D laboratories of big firms. Secondary parts may be played by scientists and policy-makers. The user–producer approach points to the fact that 'final users' in terms of workers, consumers and the public sector may have a role to play in relation to innovation.

The fact that workers and consumers tend to be absent from the scene in most innovation studies reflects, to a certain degree, the reality of a modern industrial system. Both in planned and market economies the process of innovation tends to become a professionalized activity and workers and consumers tend to become passive beneficiaries or victims in relation to new technology, rather than subjects taking an active part in the process of innovation. It is, however, not self-evident that such a division of labour is 'natural' and appropriate. Active and competent final users might enhance the innovative capability of a national system of innovation.

Further, the actual participation of 'final users' may be underrated in the literature on innovation. Workers play an important part in the daily learning process taking place in production and many incremental innovations may be the product of skilled workers improving on the process equipment. Where workers are directly involved in the process of innovation, the outcome in terms of productivity and efficiency might be more satisfactory than when they are excluded from this process. Some studies of the Japanese experience seem to point in this direction.

Among consumers we find some interesting examples in the user clubs established in relation to specific brands of personal computers. Here private consumers act as professional users, developing new software in an interaction with producers of hardware and software. But for most consumer goods the interaction is organized exclusively by producers gathering information about, and manipulating, consumer needs. An interesting theoretical contribution giving consumer learning an important role in the overall development of the national economy is made by Pasinetti (1981), who maintains that the learning of new needs are of crucial importance for the maintenance of full employment. When productivity is growing and demand for existing consumer goods becomes satisfied, the learning of new needs by consumers is a necessary condition for avoiding 'technological unemployment'.

We have already pointed out the importance of the public sector as a final user in relation to technology policy. The most comprehensive and important historical example might be the military industrial complexes in the United States and the Soviet Union. In both these cases, the state has acted as a competent and very demanding user on a very big scale. Through long-term contracts radically new and advanced products have been developed. In the Scandinavian countries there is a growing debate on the possibilities of building 'welfare–industrial complexes' oriented towards the fulfillment of social needs in relation to energy, housing, environment, transport and the health service. Such complexes might, if the public sector acts as a competent user with a long-term perspective, be as effective as 'warfare–industrial complexes' in provoking new technology. There is no reason to believe that the positive impact upon the well-being of citizens should be less.

Social innovation as the basis for technical innovation

In a period characterized by radical change in the technological basis of the economy, established organizational and institutional patterns might prove to be important obstacles to the exploitation of the full potential of new technology. In such a period, social innovations might become more important for the wealth of nations than technical innovations. The Gorbachev drive for social change and democratization in Soviet Union might be seen in this light. In the capitalist countries the focus is still narrowly oriented, either towards the manipulation of financial variables or towards an 'acceleration of technological progress'. Institutional change, strengthening the competence and the power of final users, might be one of the social innovations which can give national systems of innovation a stronger position in the world economy. It would also imply that unsatisfactory innovations became less frequent.

The need for social innovations and institutional change is even more urgent at the world level. The enormous and growing gaps between rich and poor countries reflect that the international transmission of knowledge and technology is not working as assumed by standard economy theory. In so far as specific technological capabilities are rooted in national networks of user–producer relationships, 'technology transfer' can only solve part of the problem, however. There is a need for strengthening the whole national system of innovation, including science, industry and final users.

Notes

1. The basic ideas presented in this chapter have many different and heterogeneous sources. They reflect a collective effort among the IKE group, at Aalborg University, where a research team, studying Industrial Development and International Competitiveness, has pursued theoretical and empirical work, based upon a dual inspiration from French industrial economics and British innovation theory. An earlier, but more extensive, presentation of those ideas and their different sources can be found in Lundvall (1985). This booklet was worked out in 1984, during my stay as a visiting fellow at the Science Policy Research Unit, Sussex University, and at the Department of Economics, Stanford University, and financed by a grant from the Social Research Council in Denmark. Christopher Freeman, Carlota Perez, Luc Soete, Keith Pavitt, Kenneth Arrow, Nathan Rosenberg, Paul David and many others at SPRU and in Stanford commented generously upon my work. This version has benefited not only from discussions with the participants at the Lewes and Maastricht meetings but also from comments from my colleagues and friends in Aalborg, Esben Sloth Andersen, Bjørn Johnson, Asger Brændgaard, Bent Dalum, Birgitte Gregersen and Lars Gelsing.
2. Adam Smith recognized the significance of this separation, presenting it as an important source of wealth and productivity growth: 'All the improvements in machinery, however, have by no means been the inventions of those who had occasion to use the machines. Many improvements have been made by the ingenuity of the makers of the machines, when to make them became the business of a peculiar trade' (Smith, 1776, p. 8).
3. We believe, however, that the user–producer perspective might be useful in clarifying how the different stages in the chain of innovation relate to each other in different parts of the economy.
4. It is interesting to note that Williamson, in his most recent work, recognizes that most transactions take place in organized markets. The dichotomy between pure markets and pure hierarchies is substituted by a scale where those two forms represent the extreme points. It is now argued that most transactions take place 'in the middle range' of such a scale (Williamson, 1985, p. 83). But still his analysis tends largely, to neglect the process of innovation *per se* as a factor reinforcing vertical integration and organized markets. Recent contributions by Japanese economists (Imai and Itami, 1984) do take into account technical innovation as a factor affecting the pattern of organized markets, but their focus is primarily management strategies rather than the implications for economic theory.

References

Amann, R. and Cooper, J. (eds) (1982), *Industrial Innovation in the Soviet Union*, New Haven, Conn., Yale University Press.

Andersen, E.S., Dalum, B. and Villumsen, G. (1981), 'The importance of the home market for the technological development and the export specialization of manufacturing industry', in *Technical Innovation and National Economic Performance*, IKE seminar, Aalborg, Aalborg University Press.

Arrow, K.J. (1974), *The limits of organization*, New York, W.W. Norton & Company.

Bandt, J. de (1985), 'French industrial policies: successes and failures', in *A Competitive Future for Europe?*, congress report, Rotterdam.

Bandt, J. de and Humbert, M. (1985), 'La mésodynamique industrielle', in *Cahiers du CERNEA*, Nanterre.

Brændgaard, A., Gregersen, B., Lundvall, B.-Å., Maarbjerg-Olesen, N. and Aaen, I (1984), *Besparelser eller beskæftigelse: en undersvegelse af danske kommuners anvendelse af EDB og ETB*, Aalborg, Aalborg University Press, 1984.

Carter, A.P. (1986), 'Diffusion from an input–output perspective', paper presented at the Conference on Innovation Diffusion, Venice, March 1986.

Chesnais, F. (1986), 'Some notes on technological cumulativeness: the appropriation of technology and technological progressiveness in concentrated market structures', paper presented at the Conference on Innovation Diffusion, Venice, March 1986.

Dalum, B. and Fagerberg, J. (1986), 'Diffusion of technology, economic growth and intra-industry trade: the case of the Nordic Countries', paper presented at the Second Knoellinger Seminar, Aabo, April 1986, mimeo.

David, P. (1984), 'On the perilous economics of modern science', paper presented at the TIP Workshop, Stanford University, August 1984, mimeo.

Dosi, G. (1982), 'Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change', *Research Policy*, vol. 11, no. 3, 1982.

Freeman, C. (1982), *The Economics of Industrial Innovation*, London, Frances Pinter.

Freeman, C. and Perez, C. (1986), 'The diffusion of technological innovations and changes of techno-economic paradigm', paper presented at the Conference on Innovation Diffusion, Venice, March 1986.

Georgescu-Roegen, N. (1971), *The Entropy Law and the Economic Process*, Cambridge, Mass., Harvard University Press, 1971.

Gregersen, B. (1984), *Det miljøindustrielle kompleks: teknologispredning og Beskæftigelse*, Aalborg, Aalborg University Press.

GRESI (Groupe de réflexions pour les stratégies industrielles) (1975), *La Division Internationale de travail*, Paris.

Hippel, E. von (1976), 'The dominant role of users in the scientific instruments innovation process', *Research Policy*, no. 5.

Kornai, J. (1971), *Anti-equilibrium*, Amsterdam, North-Holland.

Lundvall, B.-Å. (1985), *Product Innovation and User–Producer Interaction*, Aalborg, Aalborg University Press.

Lundvall, B.-Å., Maarbjerg-Olesen, N. and Aaen, I. (1983), *Det lanbrugsindustrielle kompleks: teknologiudvikling, konkurrenceevne og beskæftigelse*, Aalborg, Aalborg University Press.

Mowery, D. and Rosenberg, N. (1979), 'The influence of market demand upon innovation: a critical review of some recent empirical studies', *Research Policy*, no. 8, pp. 102–153.

Nelson, R.R. (1962), 'The link between science and invention: the case of the transistor', in National Bureau of Economic Research, *The Rate and Direction of Inventive Activity*, Princeton University Press.

Nelson, R.R. (1986), 'The generation and utilization of technology: a cross-industry analysis', paper presented at the Conference on Innovation Diffusion, Venice, March 1986.

OECD, (1983), 'Report on the results of the workshop on research, technology and regional policy', DSTI/SPR/83.117.

——, (1986), 'Technical cooperation agreements between firms: some initial data and analysis', DSTI/SPR/86.20, May.

Pasinetti, L. (1981), *Structural Change and Economic Growth*, Cambridge, Cambridge University Press, 1981.

Pavitt, K. (1984), 'Sectoral patterns of technical change: towards a taxonomy and a theory', *Research Policy*, vol. 13.

Rosenberg, N. (1976), *Perspectives on Technology*, Cambridge, Cambridge University Press.

—— (1982), *Inside the Black Box: Technology and Economics*, Cambridge, Cambridge University Press.

Rosenberg, N. and Mowery, D. (1978), 'The influence of market demand upon innovation: a critical review of some recent empirical studies', *Research Policy*, no. 8.

Rothwell, R. and Gardiner, P. (1985), 'Invention, innovation, re-innovation and the role of the user: a case study of British Hovercraft development', *Technovation*, no. 3.

Scherer, F.M. (1982), 'Inter-industry technology flows in the United States', *Research Policy*, no. 4, 1982.

Smith, A. (1776), *An Inquiry into the Nature and Causes of the Wealth of Nations*, Dent edn. (1910).

Williamson, O.E. (1975), *Markets and Hierarchies: Analysis and Antitrust Implications*, New York, Free Press.

—— (1985), *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*, New York, Praeger.

Yakushiji, T. (1986), 'Technological emulation and industrial development', paper presented to the Conference on Innovation Diffusion, Venice, March.

18 Can the imperfect innovation systems of capitalism be outperformed?*

Pavel Pelikan

The Industrial Institute for Economic and Social Research, Stockholm

Introduction

In this chapter, capitalism will be considered in a somewhat broader perspective than in the rest of this volume. Rather than examining how technical innovations are handled within any given capitalist system, I propose to use a comparative approach. I will consider the entire class of capitalist systems, compare its potential with the potential of some non-capitalist systems, and then search to identify in more detail those systems where technical innovations would be handled relatively best.

To justify the usefulness of such a comparative approach and to formulate my main questions, Chapter 15 by Nelson is a convenient point of departure. Focusing on the US capitalist system, Nelson examines its ways of handling the great variety of kinds and stages of the innovation process. He points to the advantages as well as to the imperfections of this system, and suggests that it could be improved by changes in its institutional design. He leaves open the question, however, of how far and in which direction such changes should go.

Clearly, this question cannot be given a solid answer without comparative analysis. As most likely no economic system is perfect, one cannot judge what should be done with a given imperfect system without comparing its imperfections with the imperfections of alternatives. Two errors are to be avoided: the one of defending a given system if at least one alternative were better, and the opposite error of rejecting it if all alternatives were even worse.[1]

Given an imperfect capitalist system where technical innovations are handled in a suboptimal way, there are two specific questions which I wish to examine here:

– Is a superior system more likely to be found within the class of capitalist systems or outside it?

* The financial support of the Marianne and Marcus Wallenburg Foundation is gratefully acknowledged. I thank Piet-Hein Admiraal, Leszek Balcerowicz, Bo Carlsson, Pierre-André Chiappori, Gunnar Eliasson, Ken Hansen, Ronald Heiner, Albert Hirschman, Richard Nelson, Keith Pavitt, Tomas Pousette, Nils-Henrik Schager, Nick von Tunzelmann, Stephen Turner, Oliver Williamson, Sidney Winter and Bengt-Christer Ysander for valuable comments on earlier drafts. None of them is responsible for my conclusions and remaining errors.

– What properties should such a superior system have; in particular, what role in promoting technical progress should it assign to government?

Comparing economic systems for their technical innovativeness is, however, not easy. The problem is that a suitable approach must be not only comparative, but moreover dynamic. As Nelson and Winter (1982) emphasize, technical progress is an evolutionary process for which static analysis is insufficient. But with a few exceptions, the existing dynamic approaches have not been comparative, while the comparative ones have been largely static. Surprisingly enough, this is even true of Schumpeter (1942), who did discuss both capitalism and socialism, but when he came to his famous dynamic problem of 'creative destruction', it was only capitalism he examined with care.

Among the exceptions, the most valuable ones probably are the comparative study of organizational adaptability and technical innovativeness by Balcerowicz (1986), and the more empirically oriented survey of technical progress in East and West by Hanson and Pavitt (1986). In a somewhat parallel way, I have made my own attempts to develop a dynamic comparative approach to economic systems in Pelikan (1985, 1986, 1987).

But none of them provides a ready-to-use method for the present purposes. This means that this inquiry must pursue two objectives: besides comparing different economic systems for their technical innovativeness, it must also search for a theoretical method which would make such a comparison fruitful.

The context and the scope of the inquiry

There are many different questions which the two objectives involve. Given the limited space of one chapter, I can examine only a few of them here. To avoid misunderstandings, it is useful to present a brief survey of all the questions involved, situating the ones which will be examined in a broader context. This will also allow me to point to the limitations of the inquiry and to warn in advance about some of its unusual steps.

The questions involved in any comparative study can be divided into three main areas: *terminology*, *values* and *analysis*. As the central question usually is of the kind, 'is one economic system better than another system?', the three areas can be exemplified by the following questions: What is an 'economic system'? What is to be 'better'? How to find out which 'system' is 'better'?

The questions of terminology cannot be neglected, for the clarity and the productivity of the entire inquiry depends on the care with which its terms are selected and defined. As the present topic is broader than what the well-defined terms of neo-classical economics can handle, the use of some non-standard terms—in particular 'system', 'arrangement', 'structure', 'organization' and 'institution'—will be inevitable. The greatest difficulty with these terms is that they sound treacherously familar, but lack an operationally clear and generally accepted meaning.

As the following section will explain in more detail, one of the unusual steps I propose to take—and which I claim is essential for all dynamic comparative approaches—is to depict an economic system in a dual way, by what I term 'the regime–structure framework'.

In the short run, an economic system is depicted in the usual way, by the economy's 'structure'—e.g. by the specific mixture of markets and/or private and/or public hierarchies into which the economy's agents are arranged.

In the long run, when such a structure itself changes—e.g. as markets and hierarchies form, expand, contract, take over each other, or dissolve—the system is depicted by the set of the prevailing institutional rules, referred to as the economy's 'regime'. By itself, this is not so unusual either, for institutional rules are often discussed in modern economic literature.

What is less usual is to consider a regime together with the corresponding structure as a couple, putting them into a well-defined, dynamically interesting relationship. Intuitively, the regime can be thought of as the rules of a game, and the structure as the configuration of the players actually playing the game. The basis of the relationship is that each regime is largely responsible for the formation and the development of the corresponding structure, which in turn is directly responsible for the economy's performance, including technical innovativeness. (For the reader familiar with biology, a useful formal analogy is the modern dual view of a living organism—as a genotype and as a phenotype.)

When the meaning of the term 'economic system' is clarified, the next terminological question is, of course, to classify and provide with names the great variety of forms which economic systems can assume. But here is the first limitation of the present inquiry. Without giving this question any detailed satisfactory answer—such an answer would probably require a new Linné—most of the time I will limit myself to two rough classification principles. Following Williamson (1975), I will classify structures into markets, hierarchies and mixtures of the two. As to regimes, the basic classification will be into two large classes—capitalist and socialist. Following the marxist tradition, it is private ownership of capital, transferable through capital markets, which will be regarded as the main distinguishing feature of the capitalist regimes. Towards the end of the inquiry, however, some finer distinctions will also be made.

Two points should be emphasized. One is the difference between the two classifications. Under the influence of neo-classical economics, which is limited to studies of given, constant and simplified structures, this difference has been effaced and economic systems have been reduced to such structures. Typically, a capitalist system has been reduced to a set of markets and a socialist system to a hierarchy of central planning. Here, in

contrast, capitalism is not identified with markets, nor socialism with hierarchies. The present view is that, in general, *the structures of both capitalist and socialist systems are variable mixtures of both markets and hierarchies*. Clearly—given the extensive use of markets in modern socialist economies and the presence of large hierarchies in modern capitalist economies—this view must be recognized as far more realistic. Only one kind of market is not allowed, by definition, to develop under a socialist regime—the market for capital.

Consequently, the technical innovativeness of capitalism is definitely not to be judged from the technical innovativeness provided for by pure markets. It will be fully recognized that hierarchies can often outperform markets—as witnessed by the example of large and successful capitalist firms, and theoretically exposed by Williamson (1975). What I will argue is that different regimes should be judged according to how conducive they are to the formation and the development of suitable structures, which in any modern economy will most likely contain both markets and hierarchies. The unusual problem to which I will then call attention is that *markets as well as hierarchies can be of very different qualities*, and that *some regimes may be conducive to the formation of markets and hierarchies of better qualities than other regimes*.

The second point of emphasis concerns the meaning of the terms 'capitalism' and 'capitalist regimes'. As already indicated, they do not refer to any specific regime, either real or idealized, but to an entire *class of regimes*. This class should be understood as containing a very large number of specific regimes, both real and idealized, with possible wide differences in performance capacities in general, and in technical innovativeness in particular. Their only common feature is that they all allow for private ownership of capital, transferable through capital markets. As regimes are only sets of rules, capital markets need not even actually exist; the only condition a regime must fulfil in order to be classified as capitalist is that its rules (in particular property rights) allow such markets to form and to develop, if entrepreneurship is supplied. This means that this class is far from limited to *laissez-faire* regimes. It also contains regimes which allow government to play a more or less significant role—to be discussed in the last section—provided that the above feature is maintained.

Regarding the questions of values—to decide what is 'better' and what is 'worse'—they are often considered to be the stumbling-block of all comparative studies. It is often claimed that such a study must depend more on the subjective values held by the student than on any objective analysis. For instance, any verdict in favor of a capitalist system is claimed to require liberal, individualistic or even egoistic values, whereas to value above all equity, solidarity and altruism is expected to yield a verdict in favor of a socialist system

Fortunately and somewhat surprisingly, however, this is not quite true, and especially not when different systems are to be compared for technical innovativeness. The general idea of how to avoid values is due to Nelson

(1981) and—in what may be regarded as another unusual step—I will apply it here. To begin with, each economic system is divided into two interrelated but separable parts:

- the system of final consumption, generating final demands;
- the system of production, determining the ways to meet these demands.

The focus is then reduced to comparing different *systems of production*, regarded as instruments for meeting some given final demands. These may be mostly demands for private goods, as generated by an individualistic consumption system with high income inequality, or contain high demands for subsidized merit goods—such as day care, education, public transport, medical care, health insurance and pension plans—as generated by a social welfare system with low income inequality. *If some systems of production can be shown to outperform other systems of production, regardless of what the final demands are, they can be said to be 'better' in a value-free fashion.*

As to innovation systems, they are simply regarded as parts of the production systems. Clearly, much of the abilities of any production system to adjust to and meet any final demands will depend on the abilities of its innovation system to supply it with suitable product and production process innovations.

The limitation here is that only those values which can be expressed in terms of final demands are respected—and avoided—in this way. Although 'final demands' can be given a very broad interpretation, even embracing items not usually regarded as goods to be produced—such as high employment, social security, and protection of nature and culture—some values may nevertheless not qualify. In particular, these are the values which one may have about systems of production *per se*—e.g. appreciating the freedom and the challenge of private enterprise or, on the contrary, the less free but possibly more stable and reassuring atmosphere of central planning.

Let me emphasize, however, that even if such values remain unsettled, they need not disturb the inquiry. The reason is that one can take them into consideration *ex post*, while fully recognizing any results which the inquiry might reach without them. For instance, if different systems of production were ranked according to 'pure' technical innovativeness, this ranking could easily be adjusted to any such values *ex post* by a suitable trade-off: the ranking of a highly innovative but little valued system would simply be somewhat lowered, and vice versa.

Finally, let me turn to the analysis employed. It has two features which should be noted and justified in advance. One is its roundaboutness. Although the main topic is technical innovativeness, not much will in fact be said about new products and production technologies. Instead, most attention will be paid to organizational and institutional problems.

To justify such a shift of focus, recall that, since Schumpeter, technical and organizational changes have been recognized as closely tied to each

other. The regime–structure framework enriches this picture by showing that organizational changes moreover strongly depend on the prevailing institutional rules. The inquiry must thus consider a long loop of causes and consequences, leading from the rules ('regime') through changing organizations ('structure') to the evolution of products and production technologies, and back to the organizations—and, in the long run, to the rules themselves.

The evolution of products and production technologies is thus only one link in a long chain, and not even the most important one for the present purposes. As Nelson and Winter (1982) point out, this evolution is not entirely natural ('darwinian'), but more or less directed. What is most important here are the different ways in which the structures themselves evolve under different regimes. And in these areas, organizational and institutional problems clearly dominate.

The other notable feature of the present analysis is the lack of mathematical models. Although all the terms used have been defined with sufficient precision, I have not yet found suitable mathematical methods to embrace the complex relationship between regimes and technical innovativeness in its entirety, without assuming away some of its essential parts. Consequently, the analysis is purely qualitative, and its results only approximative. But hopefully it does throw some new light on the question of how an innovation system could, and how it could not, be improved upon.

The regime–structure framework and organizational dynamics

To begin with a well-known picture, recall that neo-classical microeconomics depicts a capitalist economy as a collection of maximizing private producers and consumers, linked together by a set of markets. In the same spirit, comparative economics often depicts a socialist economy as a collection of maximizing socialist producers and consumers, linked together by a hierarchy of central planning.

Generalizing slightly, I define *structure*—meaning 'organizational structure'—by three groups of parameters:

- a *collection* of economic agents (e.g. firms, agencies or individuals);
- their *behavior* (e.g. maximizing or satisficing);
- the (organizational) *arrangement* which links them together (e.g. a certain mixture of markets and hierarchies).

A structure can be visualized as an active device ('mechanism') or ('organism'), which *functions* in a certain specific way. The agents inform and motivate each other through exchanges (transactions) of signals and resources, either among themselves or between themselves and environments (e.g. nature, other economies). What the agents do is determined in part by the agents' own behavior and in part by the arrangement which links them together. As will be explained below, each such arrangement contains rules which constrain, but usually do not uniquely determine, the behavior of its participants.

Globally, the functioning of a structure can be measured and evaluated by various *performance indicators*, such as aggregate output, productivity, efficiency, transaction costs, or—what is of particular interest here—technical innovativeness.

Note the difference between a structure and its arrangement. An arrangement only describes the agents' interrelations, e.g. the markets and/or the hierarchies which link them together. A structure moreover specifies the agents' actual behavior—e.g. their response functions or routines—by which they have adapted to the constraints of the arrangement. Whereas *each structure implies a certain way of functioning, and certain performance abilities*, this need not be true of arrangements. An arrangement only has a certain potential to perform, but its actual performance also depends on the behavior of the participating agents. Only in the special case of identical agents (e.g. if all were equally perfect optimizers) would an arrangement determine performance. On the other hand, if the agents are not identical and their true behavioral characteristics are difficult to observe—and this case will be of much importance later in my argument—*identical arrangements may be observed to have different performance abilities*.

Nelson's chapter provides important examples of the potential of different arrangements in the area of innovation. Extending Williamson's (1975) comparative analysis of markets and hierarchies, he shows that different arrangements are differently advantageous in providing for different innovation activities. For instance, activities with high costs of technology transfer and easy-to-appropriate returns are shown to be best conducted by profit-seeking firms on markets, whereas various non-market arrangements are shown more suitable in opposite cases.

For the present argument, the most important lesson one can learn from these examples is that *there is no single-type arrangement—neither markets, nor hierarchies—which would be universally optimal for all innovation activities*. This extends into the area of innovation the well-known result of Williamson's that neither market nor non-market arrangements are universally superior, but have different comparative advantages and disadvantages for different kinds of transactions. The general conclusion here is that *technical innovativeness requires a variegated structure of different kinds of arrangements*—e.g. a mixture of several kinds of markets and several kinds of hierarchies—which could successfully handle the entire spectrum of innovation activities.

The greatest advantage of studying structures is that they are directly responsible for the economy's function and performance. In the long term, however, they are unsuitable to represent an economic system. Whereas 'system' should refer to some relatively stable parameters, structures rarely remain stable for a long time. For instance, every time a market forms or

dissolves, or a firm enters, exits or merges with another firm, the structure changes.

The search for a relatively more stable set of parameters naturally leads to the set of institutional rules which, like the rules of a game, constrain the behavior of the economic agents involved. While the agents can enter, exit or change their arrangements in various by the rules permitted ways, the rules can stay put. Examples of such rules are property rights in the sense of Demsetz (1967), or the economic constitution in the sense of Buchanan (1975). The presently used term 'regime' is due to Hurwicz (1971) who formally defines it as the set of institutional constraints on the decision spaces of the agents within an economy.[2]

Note first that *each arrangement implies a regime*. For instance, each market implies certain transferable property rights and certain rules of signalling and contracting. Similarly, each hierarchy implies certain rules determining the rights and the obligations for each of its members. On the other hand, *each regime typically allows several alternative arrangements to form*. For instance, under the same property rights, differently competitive markets may form—including no market at all, if the agents do not supply enough entrepreneurship.

For a national economy, the regime is usually quite complex, including all the institutional rules which pertain to all parts of the economy's arrangement—such as labor law, corporate law, patent law, antitrust law, and also various customary and ethical norms, stemming from the underlying culture.

An important part of each such regime are the rules which specify the economic role of government. At one extreme, a *laissez-faire* regime may prohibit government from playing any significant economic role, while, at the other extreme, a socialist command regime may require government to organize and run the entire production, according to the rules of a certain planning procedure. Postponing the discussion of some more interesting intermediate cases to the last section, let me now only emphasize that the *economic* role of government under the given rules of a given regime is to be strictly distinguished from the *institutional* role of government in making and changing the rules and thus the regimes (e.g. through legislation). This corresponds to Hayek's (1967) classification of government activities into 'particular measures' and 'general rules'. Since the present inquiry is limited to comparing given regimes, without examining how they are, or could be, made or changed, only the former—the economic, or 'particular measures', role—will be discussed here.

The advantages and disadvantages of studying regimes are complementary to those of studying structures. Whereas a regime is relatively more stable, it is, on the other hand, less directly related to the economy's performance. The problem is that a regime does not actively perform, but only influences the performance of an interposed structure. What a regime needs, in order to show analysable effects, are some active, interacting agents whose behavior it would channel, through the specific constraints of its rules, towards certain actions rather than others. All known economic analysis of institutional rules refers indeed to an assumed structure—such as a set of perfectly competitive markets populated by perfectly rational agents, who always take maximum advantage of whatever institutional rules happen to exist, which is explicitly or implicitly referred to in most of the property rights literature.

The question then is what can be gained by considering regimes if structures are needed anyway. My guess is that not much, as long as one limits attention, as neo-classical analysis does, to initially postulated constant structures. But the situation becomes quite different when structures are regarded as variable, and their formation and development submitted to analysis. In such an organizationally dynamic view, regimes and structures play distinct and complementary roles, forming a promising theoretical framework. It is this view which Schumpeter (1942) advocated by saying: 'The problem that is usually being visualized is how capitalism administers existing structures, whereas the relevant problem is how it creates and destroys them.' This view will also be adopted here, with the notable difference from Schumpeter that other regimes than capitalism will be considered, too.

To see what organizational dynamics is about, consider first that in real economies both regimes and structures change and develop. The rules of a regime may change through legislation and/or a spontaneous evolution of custom. Structures may change, as already noted, through organization and reorganization of markets and/or hierarchies. The important point is that the two kinds of changes need not go together. In particular, a *regime need not change every time the corresponding structure changes*. For instance, a market, a firm or an entire industry may appear or disappear, while the prevailing institutional rules may stay put.

Consequently, the dynamics of economic systems can be divided into two relatively independent branches: *institutional* dynamics, studying changes of regimes, and *organizational* dynamics, studying changes of structures under a given regime. The former, which is about the political, legislative and cultural processes through which institutional rules form and reform, will not be considered here. It is the latter on which the present inquiry focuses.[3]

The basic principle of organizational dynamics is that each given regime, through the constraints of its rules, channels in certain specific ways not only the functioning of an existing structure (as standard analysis has studied), but also the formation and development of such a structure (as Schumpeter urges us to study).

This principle indicates the strategy for the present inquiry. If different regimes are to be compared for their technical innovativeness, and if this is part of the performance of the corresponding structures, the crucial problem is *which structures, of which performance, can form and develop under different regimes*.

One subtle point about the regime–structure framework should be

noted. Although the principle is simple enough, a closer look discovers a complication. When using this framework, one cannot ignore the fact that *the structure of a real economy may involve several significant levels of organization*. This may seem upsetting for the theoretical economist who is used to dealing only with one such level at a time, e.g. only with firms but not with their internal organization, as the older microeconomics used to do, or with individuals but not with firms, as the more recent transactional analysis proposes to do. But the complication is not as serious as it may seem. The framework, if slightly adjusted, may depict any multi-level organization by its recurrent application.

To see the main idea, consider the following two levels of organization, which will be sufficient for the present inquiry:

- the internal organization of multi-personal agents, such as the internal hierarchies of firms and government agencies;
- the overall organization of all agents into a national economy, such as a set of markets or a hierarchy of central planning.

Clearly, both levels can be depicted by the regime–structure framework. For instance, a firm can be said to have a certain internal structure (including a certain internal arrangement) and a certain internal regime (the written and unwritten rules of conduct for its members); and the entire economy can be said to have a certain overall structure (including a certain overall arrangement) and a certain overall regime. Whenever necessary, to avoid confusion, the adjectives 'internal' and 'national' or 'overall' will denote the two levels of the same concepts.

Structures and arrangements of different levels are easy to relate to each other. One can simply say that lower levels add details to higher levels. For instance, the internal structures of firms and agencies add details to the overall structure of an economy, displaying some of its agents as arranged collections of smaller agents (e.g. plants, departments, and ultimately individuals). The overall structure (arrangement) can then be seen as a structure of structures (an arrangement of arrangements).

The relationships between regime of different levels is more subtle. In general, a higher-level regime contains rules which constrain the design of lower-level regimes. For instance, corporation law and labor law are rules of the national regime which constrain the design of the internal regimes of firms. Different national regimes may be differently restrictive, allowing for more or less variety of internal regimes. For instance, in many socialist economies, the national regime is so restrictive that it also determines most features of the internal regimes of all firms.

For the present purposes, the essential difference between a national regime and the internal regimes of multi-personal agents is in their origins. Only national regimes originate in the above-mentioned political, legislative and cultural processes. In contrast, internal regimes are designed and redesigned, under the constraints of the prevailing national regime, within the agents themselves—e.g. by the owners or top managers within a firm, possibly after voluntary or compulsory negotiations with the firm's employees. It is therefore natural to regard the formation and development of internal regimes as part of the organizational dynamics which is to be examined here. Institutional dynamics, from which inquiry abstracts, is thus limited to national regimes. This means that only national regimes, and not the internal regimes of firms and agencies, will be assumed given.

Economic self-organization

In the context of capitalist economies, the processes by which structures form and develop—the subject of organizational dynamics—have been studied under several names. For instance, besides 'creative destruction' used by Schumpeter (1942), Alchian (1950) and Nelson and Winter (1982) speak of 'selection', Eliasson (1984) of 'structural adaptation', and Marris and Mueller (1980), in a survey of earlier studies of these processes, of 'self-organization'.

For a comparative approach, I believe the latter term most suitable. It is more comprehensive than 'selection' or 'adaptation', and it also has the advantage of pointing to an interesting literature outside economics where helpful cues for understanding these processes can be found. This is the recent mathematically or biologically oriented literature about strange loops and self-organization. In the present volume, more on self-organization, including further references, can be found in the chapters by Allen and Silverberg, who address this problem in a broad methodological perspective.

In contrast to such broad approaches, the present discussion of self-organization is quite narrow and pedestrian. The processes by which the structure of an economy organizes and reorganizes under a given regime constitute only one limited stage of the processes by which the entire society organizes and reorganizes. In order to mark the limits, I will speak of *economic* self-organization.

This means that I leave aside several other stages, assuming that they have already done their work, somehow. These include the formation of regimes—possibly called 'institutional self-organization'—and the closely related formation of languages, values and customs—possibly called 'cultural self-organization'. And one should not forget, as social scientists tend to do, the historically preceding 'biological self-organization', which has formed the genetic potential of human brains to create and learn languages, cultures and institutions, on which the entire self-organization of societies ultimately reposes.

To narrow the focus on economic self-organization clearly involves the risk of missing some possibly significant feedbacks through which it relates

to other stages of self-organization of societies. An important example is the feedback which, in somewhat different terms, the marxists like to emphasize: maladapted structures of economies produce crises which provoke institutional and cultural changes, resulting in a more or less different regime for the next round of economic self-organization. Admitting that such feedbacks may be significant, I nevertheless contend that a good understanding of how economic self-organization works under different given regimes is essential. Although it may not be sufficient, it is certainly a necessary basis for any more ambitious study where also the evolution of regimes is to be examined.[4]

To visualize economic self-organization in more concrete terms, recall the example of markets and hierarchies which organize, reorganize, expand, take over each other, contract or dissolve. In general, economic self-organization is made of processes which change at least one of the three components of a structure:

- the collection of economic agents (e.g. through entries, exits, takeovers or divestitures);
- their behavior (e.g. through internal self-organization of firms, or learning of individuals, conceivable as internal self-organization of brains);
- their arrangement (e.g. through formation, modification or dissolution of various communication and motivation channels of which different types of markets and hierarchies are made).

It is possible to provide for a relatively simple microeconomic model by abstracting from demographic changes—that is, by assuming a constant set of individuals of given learning potential (the competence to acquire competence). In this case, economic self-organization can be depicted as a game played by these individuals, under the rules of the prevailing regime. Such a game includes the following kinds of moves:

- designing and redesigning various (multi-level) arrangements;
- assigning and reassigning positions within these arrangements to specific individuals;
- learning of new individual behavior (new competence) within the given learning potentials.

These kinds of moves may be taken separately or simultaneously. For instance, an entrepreneur may design a firm, assign himself a certain position within it, and learn a new competence while doing so. As a result of these three kinds of moves, the given individuals keep organizing and reorganizing themselves into a series of structures.

For the present purposes I need not elaborate this model in detail, but only outline how it relates to, and differs from, standard analysis. Obviously, the main difference is that the formation (self-organization) of structures is considered at all. It is only when a structure is formed that standard analysis can start its work—to determine how such a structure functions and performs. This difference then entails several other differences.

First, the usual view of economic behavior must be enlarged by a new dimension. Traditionally, economic agents have been examined for their allocative behavior, that is, for their ways of transacting signals and resources within some already organized structures. To some degree, their learning behavior has also been studied, e.g. in the theory of learning by doing. The new dimension—which I propose to call *associative*—is the behavior by which the agents form, modify or dissolve the various inter-agent links of which structures are made, such as lasting contacts with business partners, long-term employment contracts, and the control of firms or agencies.[5]

Associative behavior involves its specific *associative constraints*—such as limited span of control, limited precision of languages, and limited trust—and *associative preferences*—such as favoritism, nepotism, and likings for rituals, status and control ('power'). Such constraints and preferences influence the behavior of economic agents side by side with the usually considered resource constraints and consumer preferences. To the surprise of conventional analysis, they may push economic self-organization towards structures which are far from allocatively efficient.

All economic agents are thus recognized as associatively active and selective. They all contribute to economic self-organization by influencing at least some links of the structures in which they become involved. But their contributions are likely to be asymmetric, in particular in complex structures. Such structures, in order to begin to form, usually need an entrepreneur–organizer—private or public—to provide an initial design and to trigger the formation. The model thus throws a new light on the role of entrepreneurs, making it comparable to the role of catalysts in the formation of chemical compounds.[6]

There is another additional problem that the model must solve. In the excitement over the new problem of self-organization, one must not forget (as some students of self-organization tend to do) that the old problem of resource-allocation does not disappear. The additional problem, then, is how these two problems relate to each other. Although formal modelling is difficult, the general principle is simple. To recall, economic self-organization forms organizational structures which determine how resources will be allocated. But since economic self-organization, in turn, needs resources—e.g. the capital which a firm needs for entering, expanding, taking over another firm, or simply surviving—the resulting allocation of resources becomes an important constraint on further self-organization.

An organizationally dynamic comparative analysis thus slowly begins to take shape. Although it agrees with traditional analysis that structures determine performance, it does not compare them directly. Instead, the focus of comparison is shifted to regimes. These are compared for their capacities to channel, by the constraints of their rules, economic self-organization towards structures of some desirable performance, which, in the present case, is a high technical innovativeness.

The strange loop of economic competence

The difficulty with this strategy is that it substantially prolongs analysis. It preserves the old question of how structures perform, and raises the new question of how structures self-organize. The strategy could hardly be fruitful, unless we can find a substantial short cut from properties of self-organization to the performance of structures. What we need is a characteristic feature of structures which is crucial for their performance, and at the same time easy to identify as a product of their self-organization. I now wish to argue that such a feature exists, giving it the name 'economic competence'.

My starting point is the concept of competence as introduced by Heiner (1983).[7] In essence, competence measures the capacities of an agent to solve difficult problems. Each level of competence implies a maximum difficulty of problems which can be solved optimally. If more difficult problems must be solved, their solutions are likely to contain costly errors, and thus be suboptimal. To refer to such a situation, Heiner coins the term 'competence-difficulty gap'.

As Heiner points out, limited competence is caused by imperfect ability to use information, which is to be distinguished from the usually considered case of imperfect information. One consequence of this distinction is that one can finally recognize in theory what has been commonplace in practice: the possibility of different results when the same information is used by agents of different competence. In order to refer to such a situation in terms close to the usual economic jargon, I shall say that competence is *scarce* and *asymmetric*.[8]

For the present purposes, it is essential to distinguish *economic competence* from competence in other fields, such as technology and politics. Such a distinction can be found already in Knight (1921), who insisted on the difference between economic problems, the subject proper of economic analysis, and technical problems, which call for the competence of natural scientists and engineers.

Economic competence can be regarded as a mixture of three basic components, corresponding to the above-mentioned three dimensions of economic behavior: *allocative* competence—e.g. the competence for deciding on the quantities and/or prices of inputs and outputs, or for choosing production techniques; *associative* competence—e.g. the competence for designing, joining, modifying or leaving organizations; and *learning* competence, with the meaning of 'economic or business talents', as the competence to learn these two kinds of competence. In contrast, technical competence is the competence for designing products and production processes in terms of physical variables, and includes also the competence to learn such competence, or 'technical talents'.

To be sure, technical competence can also be of much concern for an economist, in particular in studies of technical innovations. Moreover, it is often intimately interwoven with economic competence. Typically, the solutions of technical problems require economic evaluation, while the solutions of economic problems are constrained by available technologies. It may even be the same person—such as the Schumpeterian entrepreneur — who uses both. But a difference nevertheless exists and, as will become clear shortly, is of great importance for economic theorising. It is one thing to design a product or a production process in terms of physical parameters, and another thing to evaluate the private and/or social costs and benefits of such a design, in order to decide for or against its use in production. Whether the technical design and the economic evaluation are made by different persons or by the same person is clearly immaterial for the validity of the distinction.

The reason why this distinction is so important is that the two kinds of competence raise substantially different problems for economic theory. Paradoxically enough, it is technological competence which is easier to handle. As the well-known literature on human capital, learning by doing, and job assignment amply illustrates—and it is perhaps useful to emphasize that this literature is about technological and not economic competence—neo-classical theory has no difficulties in recognising that technological competence may be scarce and asymmetric. This kind of competence can simply be regarded as a property of human factors of production, and its production and allocation treated in a formally similar way as the production and allocation of any other capital good.

But the apparently innocent step from scarce and asymmetric technical competence to scarce and asymmetric economic competence demands a real somersault from economic theory. When acting as workers, engineers or scientists, people can be regarded as factors of production. But when acting as traders, investors, managers, policy-makers or planners, they must be recognized as economic agents. To admit that their competence may be scarce and asymmetric even for these roles undermines the entire neo-classical theorizing. It contradicts the fundamental neo-classical axiom that all economic agents are perfectly rational optimizers—that is, of equally abundant competence for solving economic problems.

To show how much the beautiful axiomatic building of neo-classical economics is damaged, let me elaborate. If economic competence is scarce and unequally distributed, it becomes a scarce resource and the problem of its allocation must be raised. But this problem is fundamentally different from all allocation problems for which neo-classical theory has been built. Whereas all other scarce resources are merely *objects* being allocated, economic competence determines at the same time the very *method of economic calculus* by which resource-allocation is governed.

The crucial role of the rationality axiom in neo-classical theorizing is thus exposed from a somewhat unusual angle. This axiom is needed to separate the objects from the method; what it implies is that all agents have abundant economic competence for which no allocation problem ever arises. But if economic competence is scarce and itself in need of allocation, this separation is destroyed and a strange loop, from the family of paradoxes which have scourged axiomatic building of modern mathematics, appears in full beauty. Economic competence spreads on both sides of the fence: *the already allocated economic competence determines the method by which further allocation of economic competence is governed.*[9]

This step, besides being theoretically disturbing, points to two important problems. The first is the possibility of path-dependency in economic self-organization, in particular in the formation of hierarchies. Initial accidents in the allocation of competence among the founders of a hierarchy may gradually amplify, possibly causing the entire hierarchy to become pervaded by exceptional competence, or exceptional incompetence.[10]

The second problem is the possibility of important failures of economic systems which have been well known in practice, but thus far neglected in theory. According to neo-classical analysis, all welfare losses in all systems must ultimately be ascribed to improper motivation of *perfectly competent egoists*. Even the losses due to imperfect or assymetric information must ultimately be ascribed to improper motivation of the agents who have the right information but do not communicate it, or who could obtain it but do not search for it—for instance, because of differences between private and social costs and benefits. But when this step is taken, we can moreover see the losses caused by possibly well-motivated but *not so competent egoists or altruists*. This means that economic systems can now be assessed not only according to how well they can cope with egoism, but also according to how well they can cope with incompetence. Since lack of competence can harm technical innovativeness at least as much as lack of motivation, this problem is of high relevance here.

The dependence of technical innovativeness on economic competence can now be summarized as follows. Although the competence at the fighting line of technical progress is technical, it is on economic competence that its production, recognition and deployment depends. Therefore, *the technical innovativeness of an economic system ultimately depends on its abilities to allocate efficiently economic competence*. In particular, this is the competence of entrepreneurs, managers, investors, policy-makers or planners to read and interpret economic signals, to estimate future supply and demand, to evaluate the probability of success of different research and production projects, to design contracts and organizations, and, last but by no means least, to estimate the competence limits of others and oneself.

Economic competence as an outcome of economic self-organization

To complete the short cut, I now need to establish that the use of economic competence in an economy is determined by economic self-organization. An essential element of my argument is the concept of 'tacit knowledge', as introduced by Polanyi (1967) and discussed in the context of evolutionary economics by Nelson and Winter (1982). In essence, this is knowledge which can be freely used by its owners, but cannot be expressed and communicated to anyone else. The point I now wish to make is that economic competence is tacit in this sense.

The best theoretical justification that some knowledge (information) must always be tacit can probably be found in computer theory. It clearly shows that in order to observe, interpret, act upon or communicate any information, some information must always pre-exist, such as working knowledge of concepts, codes (languages) and logic. Although some of such information might have been communicated on an earlier occasion, that communication inevitably required some pre-existing information, too. The upshot is that at least some of the information on which all communication and decision-making ultimately repose must be tacit, that is, inherent to the systems involved. As an example, think of a computer, where the entire hierarchy of treatment of software, communicable through its inputs and outputs, ultimately reposes on its hardware, inherent to its construction.

The claim that economic competence is tacit obviously depends on how this is defined. To choose the right definition, let me examine a little closer Heiner's distinction between information and ways of using information. Strictly speaking, a way of using information is also information (e.g. a program or a routine), and there may be ways of using this information (e.g. programming programs), and so on *ad infinitum*.[11] This discloses that the distinction is not uniquely determined, but offers a number of options. To fit my claim, I make the definition coincide with the distinction between communicable and tacit information. This means that not only data (e.g. prices and quantities), but also all communicable ways of using data as well as all communicable ways of using such ways (e.g. instruction books, computer programs), are regarded as (communicable) economic information. This means that I will use 'economic competence' to refer to *the (residual) ways of using information in economic decision-making which cannot be communicated, but are inherent to each economic agent*—such as the competence to understand and suitably apply instruction books and computer programs, but which cannot itself be put in an instruction book or a computer program.

The upshot of this definition is that the only way by which individuals can acquire economic competence is their own learning, based on their own experience, formal education, and the innate (and/or by early education determined) competence to learn economic competence 'economic talents'.

The next step I need to take is to define economic competence as a property of not only individuals, but structures in general. This will make it possible to speak, for instance, of firms, agencies and entire economies as being more or less competent.

Let me now define *the competence of a structure as the allocation of the individual competence involved in the structure*. This means that a structure's competence is made up of all the individual competence involved, but without being a simple sum of individual contributions. What also counts is the structure's arrangement, and the allocation of specific individuals over this

arrangement. Clearly, the competence employed for top economic decisions—such as those of entrepreneurs, managers, investors, policy-makers and planners—must weigh more than the competence of the rank and file.

Two implications are of particular importance. First—and this is only another way of expressing conventional wisdom—the same individuals can form structures of different competence, if organized into different arrangements. Second—and this is a less usual point for economic theory—*the same arrangement can result in different competence if it involves different, or differently permuted, individuals*.

Economic competence has already been related to individual rationality. It is now moreover possible to relate it to *x-efficiency* of firms and allocative efficiency of economies. The three traditionally separate concepts—rationality, *x-efficiency* and allocative efficiency—are thus provided with a deep common meaning. At different levels of organization, they all refer to the economic competence of structures, determining the structures' performance abilities.

The conclusion that economic competence is allocated by economic self-organization is now easy to draw. That economic competence cannot be allocated through the usually studied inter-agent transactions obviously follows from its tacitness. And that it is precisely economic self-organization which must assume this task is nearly as obvious. As an inherent property of structures, economic competence must be allocated by the same processes by which structures are formed. And, by definition, these are the processes which constitute economic self-organization.

Economic self-organization under different regimes

What has thus been established is that technical progress depends on economic competence, which is allocated by economic self-organization. The next step is to examine how this is channelled by different regimes.

Let me begin with the behavior in a similar way as the rules of a game constrain the behavior of the players. Following the distinction between allocative and associative behavior, the rules of a regime can be divided into two corresponding categories:

- *allocative rules*, constraining the agents in their allocating of resources (e.g. in investment, R & D, production and trade);
- *associative rules*, constraining the agents in their associating and dissociating (e.g. in entries, exits, cooperation agreements, take-overs and divestitures).

As associating and allocating are interrelated, both categories influence economic self-organization. To the extent that resources are required, the allocative rules—such as property rights—play an important role. But the distinguishing feature of self-organization is that it not only uses resources, but moreover changes structures for the next round of resource-allocation.

It is for this additional area that the associative rules are specialized. They can be exemplified by antitrust law, corporate law, the rules regulating entry and exit, and the rules regulating the labor and capital markets—where most of the associating and dissociating of individual employees, managers and owners is conducted under a capitalist regime.

A regime thus influences economic self-organization in a double way. Its allocative rules do so indirectly, via their responsibility for economic results, determining which structural changes become economically feasible. Its associative rules do so directly, by determining which of the economically feasible changes are moreover institutionally permissible.

Although when new, a regime must begin with the structure inherited from its predecessor, its double influence on economic self-organization makes it increasingly responsible for the subsequent states and performance abilities of the structure—much as the genetic message of an organism is responsible for the development of the organism's anatomy and behavior. This discloses as illegitimate the neo-classical habit to assign an arbitrarily postulated structure to a given regime—such as a set of perfectly competitive markets to capitalism, or a hierarchy of optimal planning to socialism—without verifying whether the regimes in question are actually capable of engendering or at least preserving such structures.

To determine how economic self-organization would actually unfold under a single given regime is not easy: a complex simulation model would probably be required. But fortunately, and somewhat surprisingly, the question of how different regimes compare with each other in channelling economic self-organization can be given an approximative but meaningful answer by relatively simple means.

The key idea is to focus on failures of economic self-organization, and to assess different regimes for their abilities to resist them. Since economic self-organization is modelled here as (several levels of) selective associating and dissociating of economic agents, let me denote such failures as *associative failures*.

In general, to speak of failures of economic systems requires the choice of certain values, or performance criteria (e.g. Pareto optimality, or a more specific social welfare function), in order to determine what is a failure and what is a success (or an optimum). But here it is possible to avoid the question of values in the way described above in the second section. Economic self-organization can be assessed for successes and failures in adapting production structures to some given final demands, regardless of what these actually are.

Associative failures can be divided into two basic categories:

- *surviving errors*, consisting of mistakenly formed and afterwards neither corrected nor dissolved maladapted structures—such as an inefficient market, a poorly organized firm, or an incompetent policy-making or planning agency which survive for long periods of time;
- *absent successes*, consisting of potentially successful structures which, although feasible given available competence, failed to form—such as new firms promoting new technologies, or otherwise superior to incumbent firms, whose entry has been hindered or not sufficiently encouraged.

Referring to these two categories, a simple method for comparing regimes can be devised. The main idea is to compare their resistance to associative failures, that is, their intolerance to surviving errors and their openness to associative trials. If regime A proves to be *relatively* more resistant to associative failures than regime B, the conclusion will be that the structures formed under A are likely to become better adapted, and therefore perform *relatively* better than the structures formed under B—and this regardless of final demands, and also regardless of how poorly adapted the structures under A might appear according to some absolute ('nirvana') criteria. Consequently, if it is technical progress which is demanded, it will also be better promoted under A than under B—in the sense that a suitable variegated structure, containing the right mixture of markets and hierarchies of the right qualities, is also more likely to form under A than under B.

Capitalism is necessary for superior technical innovativeness

After the somewhat long, but I believe inevitable, theoretical detour, it is now possible to outline the answer to the initially stated question. Let me expose the main points by way of justifying the following proposition: *the superior regimes, promoting technical progress better than all other regimes, belong to the class of capitalist regimes—that is, regimes allowing for private ownership of capital, transferable through capital markets*.

To avoid misunderstanding, let me emphasize what this proposition does not imply. First, it does not imply that technical progress would be successfully promoted by all capitalist regimes; the possibility that some of them may perform poorly is not at all excluded. Second, the proposition does not imply that a superior regime should be of the *laissez-faire* kind, excluding all active role of government. Private ownership of capital, transferable through capital markets, is claimed to be a *necessary, but possibly not sufficient*, condition for superiority.

This proposition clearly contradicts neo-classical analysis which formally proves—by constructing various methods of optimal socialist planning[13]—that *some* socialist systems can perform at least as well as *the best* capitalist systems. To justify the proposition I need to show why this proof no longer holds when economic self-organization enters the picture.

Recall that the two crucial assumptions for this proof are: the stocks of all resources can be measured, at least by their users; and all production units as well as the planning agency are perfectly competent optimizers. But neither holds for economic self-organization, specialized in allocating tacit, scarce and asymmetric economic competence.

Because of its tacitness, economic competence is a resource whose stocks cannot be directly measured, not even by their users, as the frequent cases of overestimation or underestimation of one's own competence amply demonstrate. Only indirect measuring, via actual performance, is possible. Such measuring requires competition, in the sense of contests or tournaments, conducted in the same field as the competence to be measured. Hence for measuring economic competence, the competition must also be economic, and not political or rhetorical.

Because of its scarcity and asymmetry, economic competence cannot be *a priori* assumed to be allocated in any favorable way. In particular, one cannot simply assume that socialist firms and planning agencies are competent at will. Their competence must be put in question and examined as a result of economic self-organization under a socialist regime.

But if the two assumptions do not hold, the following proposition does: *economic self-organization cannot be optimally planned in advance, but must involve experimentation through associative trials and errors.* The reason is easy to see. If economic competence, which economic self-organization is to allocate, cannot be directly measured, there is no reliable information base, neither centralized nor decentralized, for an optimal planning of its allocation. Moreover—and this is how the strange loop of economic competence manifests itself here—unless economic competence is optimally allocated already at the outset, the planning agency itself may be far from assembling the best available competence for this task. Consequently, any regime which is not sufficiently open to associative trials and ready to cope with associative errors *even at the highest organizational level* can easily cause the structure of the entire economy to become one huge surviving error.

This proposition thus exposes the crucial importance, *under any regime*, of the generation of associative trials and the elimination of subsequent errors or, alternatively, the selection of successes. A too lax selection will cause surviving errors, whereas a too constrained trial-generation will cause absent successes. Moreover, absent successes may also be caused by a severe but misdirected selection which prematurely eliminates future successes in temporary difficulties.

It is now easy to see why private ownership of capital and capital markets are so important. Consider the two basic alternatives for ruling them out, to which I refer as 'government socialism' and 'cooperative socialism'. Each alternative refers to a large class of regimes, both real and idealized, which may differ in many other rules, but have similar property rights for capital. Since much of economic self-organization is shaped by these rights, significant global conclusions can be drawn for each class, regardless of what the other rules are.

In all regimes of the government socialist class, capital is formally owned by central political authorities. Even if the decisions on its current use can largely be delegated to lower levels, economic self-organization must remain largely centralized. Balcerowicz (1985) speaks of 'centralized organizational rights', and Hanson and Pavitt (1986) describe this as a situation where any organizational reshaping requires prior approval from the central authorities. Note that such regimes may, but need not, require central planning of resource-allocation. This class is thus much larger than the usually studied socialist planning, for it also includes quite decentralized socialist regimes which allow for extensive use of product and labor markets, such as in Hungary.

There are several joint reasons why this way of channelling economic self-organization is inferior. First, it assigns to the central authorities a dominant role, for which they are unlikely to assemble the best available competence, given their origins in political and not economic competition. And even if the most competent entrepreneurs, managers and investors of the old regime were initially selected, such a selection would soon become obsolete in a dynamic world where new types of competence, including competence for judging competence, may continuously be required.

Second—and much of this is in fact a consequence of the first—economic competence will likely be misallocated also at all lower levels. The entire structure of socialist firms and their hierarchical and/or market arrangements is likely to contain more of both absent successes and surviving errors, because of a too constrained trial-generation and a too lax error-elimination. The most competent trials are likely to be prevented for lack of the necessary approval from the probably less competent central authorities. And too many errors ('lame ducks') are likely to keep surviving because of their monopolistic privileges and/or generous subsidies—or 'soft budgetary constraints', to use Kornai's (1980) term. The empirical findings of Hanson and Pavitt (1986) are in very good agreement with this theoretical argument.

Note that it is the low expected competence of the central authorities which is seen here as the main reason why soft budgetary constraints are unlikely to be used in a more clever way. To be sure, highly competent investors might be able to perceive the fine differences between 'permanent lame ducks' amd the recuperable ones, which can be transformed into future successes. They can then outperform the short-sighted product markets by helping the latter with selective investment as well as restructuring, often requiring important changes in the top management. But—and this is the essential point—*investors with such competence are unlikely to be found and kept without continuous economic competition on capital markets*.

Let me now turn to the regimes of the class of cooperative socialism, where the social ownership of capital is decentralized, each firm being owned by the collective of its employees. No central planning is required, and product and labor markets can extensively be used. But capital markets must be limited to credit markets, with no real stock exchange.

Following Ward (1958), neo-classical analysis sees the main problem with cooperative socialism in the perverse responses of employee-owned firms to changes in demand and profit. But in broader discussions, such as in Vanek (1970), this problem has not been recognized as decisive. Several ways to alleviate it have been proposed, and the principles of employee participation and profit-sharing have been claimed to more than compensate for it by other social and economic advantages.[14]

Let me therefore emphasize that the present argument is of quite a different kind. It does not put in question any of the claimed advantages. It admits that successful cooperatives may exist, and that many private firms might benefit from applying some of these principles within their internal structures. What the argument claims is that cooperative socialism *as a regime* is conducive to structures of lower technical innovativeness than what at least some capitalist regimes can achieve.

A scrutiny of associative failures can again justify this claim. At first sight, it seems that surviving errors need not worry cooperative socialism more than capitalism. Since product and labor markets can be used fully, competition and hard budgetary constraints seem able to keep eliminating lame ducks as rigorously as under the best capitalist regime. But this is not quite true. In order to see why, let me begin with absent successes which constitute a more obvious drawback of cooperative socialism.

Collective decision-making, as implied by the cooperative ownership of capital, acts as a constraint which discourages or prevents some new firms from entering and some small successful firms from expanding. On top of the problem of perverse incentives for growth of firms, as exposed by Ward, the present argument adds the problem of scarce, tacit and asymmetric competence. Successful entry and expansion of firms are often based on exceptional competence of innovators, entrepreneurs and investors, which, just because of their exceptionality, will often be misunderstood in any larger collective. That this problem will tax new industries and new technologies with particular severity is obvious.

A related problem arises with the supply of capital, in particular risk capital, on which associative trials often strongly depend. As is well explained by Neuberger and Duffy (1976), cooperative socialism not only precludes stock markets, but strongly constrains the entire banking sector. Only government, and under certain restrictions existing production cooperatives, can be allowed to enter. As a consequence, this sector is likely to suffer not only from absent successes, but also from surviving errors, to the detriment of its resulting competence. It is not only that no potentially competent investors from outside government and existing production cooperatives can ever try, but also that government banks may grow bureaucracy, lose competence and yet keep allocating much of the scarce capital.

With the low expected competence of the banking sector in mind, it is easy to see the high probability of surviving errors also in production. The survival of a firm does not depend only on its customers, but also, and sometimes above all, on its investors. Unless these are extremely competent, they will more often than not fail to recognize lame ducks from future successes in need of capital. Consequently, the production structure will likely contain more of the former and less of the latter, in comparison

with regimes which allow more competent investment structures to self-organize. And it is precisely here that private ownership of capital and room for fully fledged capital markets prove essential.

Nelson's qualified praise of capitalism is thus strengthened in a comparative context. It is not markets, but the potential for efficacious experimentation with both market and non-market structures at *all* levels of economic organization which is shown to be the crucial comparative advantage of capitalism *as a class of regimes*.

Note that this argument differs from the usual pro-capitalist arguments of the public choice or the neo-austrian varieties. As to markets, they are not claimed to be always superior to non-market arrangements; the risk of market failures in resource-allocation as well as in self-organization is recognized as real. As to government, it is not *a priori* regarded as the villain of the piece. In contrast to Public Choice, government is here accorded the benefit of the doubt as to its intentions, and only its economic competence is, in a probabilistic way, put in question. Moreover, the discussion is limited to the role of government in production, in particular in R & D, and in the corresponding investment. The areas of macroeconomic policies and policies concerning income transfers and consumption, both public and private, are left aside. A wide variety of policies in these areas—ranging from what may be called 'conservative capitalism' to 'advanced welfare society'—may thus be fully compatible with the present argument.

How to improve upon capitalist regimes?

Since the claimed advantage of capitalism is only comparative and concerns the potential of only a subclass of capitalist regimes, it is fully legitimate to suspect any given capitalist regime—as Nelson does with the US capitalism—of leaving room for improvements. To conclude, let me briefly address the question which Nelson notes but leaves open—of where to search, and where not to search, for such improvements.[15]

Two kinds of institutional rules are central to such a search: those about economic competition, and those about the economic role of government. According to the present argument, the main task of competition is to select and promote persons and multi-personal structures of the best available competence, or at least to demote the persons and to dissolve the structures of insufficient competence. As has been shown, the selection of highly competent investors and, with their help, of highly competent hierarchies is crucial. Whereas highly competent hierarchies can much improve upon the short-sighted selection of producers and innovators by product markets, mediocre hierarchies can, on the contrary, do much worse than these markets.

The implication is, in essence, that economic competition should be modelled after tournaments in organized sports, in order to discover and promote specific competence (rather than general ruthlessness).

The old intuition of the US legislators is thus given a somewhat unusual theoretical support. The main point—which is simple in principle but involves a host of subtle problems in practice—is to keep the entry to and the exit from all markets, including capital markets, reasonably open, and the competition itself reasonably fair-play. The search for improvements is thus directed in part to removing institutional barriers to entry and exit, and in part to prohibiting predatory (strategic) behavior of incumbent competitors, e.g. by suitable legislation on antitrust and fair business practices. Another task is to neutralize perverse incentives to associating, such as the likings of managers for corporate control *per se*, or the incentives of stockbrokers to push for any mergers, whether efficient or not.

As to government, the present argument exposes its low expected economic competence, regardless of its intentions.[16] The general implication is that government should be institutionally prevented from intervening by selective measures in production, R & D and the corresponding investment, *whenever it is possible* to organize economic competition in such a way that more competent private agents for taking such measures (e.g. sponsoring and coordinating research or redressing failing firms) are likely to emerge.

The emphasized clause is of much significance. It may justify a non-negligible agenda of selective measures (which would not be welcome by either the public choice or the neo-austrian approaches). It calls attention to the possibly important category of such measure for which private agents are unlikely to emerge, *and which are better taken with relatively low expected competence than not taken at all*.

One example is the application of antitrust to particular cases. Even if the government agencies in charge are of imperfect competence for this task, their intervention can be justified on similar grounds on which imperfect umpires are preferred to no umpires at all in organized sports.

Another example is government entrepreneurship in some socially important areas where private entrepreneurs are slow in appearing. As the supply of private entrepreneurship is, at least in part, culturally conditioned, such areas may be of more importance in some cultures than in others. The crucial, but often violated, requirement is that the entry to such areas remain open and the government initiated units be exposed to competition, on comparable terms, from potential private entrants. The society will then gain in one of two ways: such a unit may succeed—which the present argument does not exclude but only shows as somewhat unlikely—or provoke, by its poor performance, more competent private entrepreneurs to enter and take over such a previously neglected area.

The coordination and sponsoring of research, in particular generic (or basic) research, is probably one such area in any culture. To be sure, not even here should private entrepreneurship be underestimated; in Chapter 15, Nelson gives several examples of private foundations supporting basic research, as well as of privately organized cooperative agreements among firms and universities for various R & D ventures. But private entrepreneurship here is likely to be insufficient—in some cultures more than in others—which means that some government policies may help, *even when conducted by poorly motivated government agencies of relatively low expected competence*.

Such policies include—and many theoretical economists should acknowledge this case personally—government subsidies to basic research. One may very well admit that the subsidies are likely to be misallocated—e.g. because of favoritism and/or a lack of highly competent foresight, more subsidies may go to conventional lines of research, yielding 'lower marginal contribution', than to emerging scientific innovators, capable of producing 'higher marginal contribution'. Nevertheless, even the disappointed innovators will probably agree that this is a better solution than if no basic research were subsidized at all.[17]

Another promising candidate for government policy is the choice of technical norms, especially when it matters less *which* norm is chosen than that a norm *is* chosen.[18]

In general, the list of candidates worth examining is quite long. France and Japan provide perhaps the best-known examples of capitalist systems where government engages, with variable success, in a particularly long list of policies intended to promote technical innovativeness. As I cannot examine such a list in detail here, let me briefly summarize the main principle implied by the present approach:[19]

- Policies with high coordination effects and low competence requirements—such as the choice of a technical norm among equally good alternative norms—imply high social gains at a low risk, and can be safely recommended.
- The more the effects of a policy—in terms of social gains or losses—depend on the competence employed, the riskier it is to allow a government agency to conduct it.[20]

Paradoxically enough—and this is perhaps the only definite conclusion I can draw here—many such risks can be reduced only in capitalism. It is only there that successful innovations can also be supported by independent private investors, in spite of possible policy errors.

Notes

1. The second error is the one Demsetz (1969) warns against in his discussion of what he calls 'the nirvana approach' and 'the grass is always greener fallacy'. In part, this chapter can be seen as corroborating, extending and qualifying Demsetz's argument.
2. Institutional rules, as any rules of a game, raise the problem of their observance. Here I abstract from this problem by assuming for each of the regimes considered that the agents involved effectively observe, and expect each other to observe, all its rules. Since such rules originate partly in written law and partly in unwritten custom (ethics), one can imagine that their observance is

achieved by a mixture of formal law enforcement, informal social sanctions, and internalized ethical norms.

3. Referring to the above-mentioned analogy with the genotype–phenotype framework of modern biology, organizational dynamics can be considered analogues to ontogeny, and institutional dynamics to phylogeny. The differences in time-scales are, of course, not the same. Phylogeny is so much slower than ontogeny that the phenotype of an organism can usually form and fully develop under a constant phenotype. Although the present simplifying assumption of given regimes corresponds to this case, in real economies the two dynamics are often interwoven (cf. also the discussion on economic v. institutional self-organization below).
4. Although the evolution of regimes may also depend on many non-economic factors (e.g. religious, ideological or cultural), the evolutionary potential of each regime is strongly constrained—and the marxists should be first to agree—by its economic performance. But as this performance depends on the structures formed, it is the economic self-organization under each regime that will eventually determine much of the regime's fate in the broader process of institutional self-organization.
5. The failure to see associative behavior as a separate dimension of economic behavior seems to be the main reason why theory has made so little progress in studies of economic self-organization. In economic literature, the closest topics probably are coalition formation, long-term employment contracts, and the issue exit v. voice as examined by Hirschman (1970). Balcerowicz (1985) has a similar concept in mind when he speaks of 'organizational actions'.
6. It is instructive to note that such an enlarged view of economic behavior can no longer refer to the paradigm of mechanics, on which neo-classical economics has been built, but is closer to that of chemistry or biochemistry. Economic agents can no longer be regarded as passively accepting their roles in a given 'mechanism', but must be recognized as actively and selectively 'reacting' with each other (cf. the affinities of atoms and molecules, and the role of catalysts). They must be recognized as themselves forming and reforming the 'mechanism'—and one should now rather say 'organism'—of which they are parts.
7. See also Heiner's chapter in the present volume.
8. To term unequally distributed competence 'asymmetric' has been suggested to me by Heiner in a personal communication. 'Asymmetric competence' thus nicely complements the familiar 'asymmetric information'.
9. The most inspiring reference about strange loops is probably Hofstadter (1979).
10. The chapter on 'Injelitis' in Parkinson (1957) provides an excellent example of a path-dependent process through which an entire hierarchy can become pervaded by incompetence.
11. Winter (1971) was probably the first economist who exposed such an infinite regression in economic decision-making. This problem has been recently elaborated by Mongin and Walliser (1987).
12. Whereas many students of self-organization use this new field to combat reductionism and methodological individualism, I believe, and hopefully demonstrate by the present discussion, that the two principles can fruitfully be used even in this field.
13. See Heal (1973) for a pedagogically excellent survey of these methods. The paradox that it is neo-classical analysis which provides such a strong defense of socialism is pointed out in Nelson (1981) and elaborated in Pelikan (1985).

14. During the discussions on the economic reform in Czechoslovakia, I helped to elaborate one of such ways in Kocanda and Pelikan (1967).
15. This question has in fact two components: the theoretical knowledge of such improvements, and the political means for implementing them. I will briefly address only the former, fully aware of the fact that there is no direct way from theoretical knowledge to practical implementation.
16. The conjecture that government lacks this type of competence has often been made (see, e.g. Eliasson, 1984), but the present argument seems to be first to provide it with theoretical justification.
17. As Cazes (1986) points out in his revealing comparison of Tocqueville, Cournot, and Schumpeter, it was already Tocqueville who advocated government support to basic research as a necessary condition for avoiding decadence of a democratic society.
18. See Arthur's chapter in this volume for a more detailed discussion.
19. From a somewhat different point of view, the case of Japan is examined in detail by Freeman in Chapter 16.
20. This proposition corroborates and extends in an organizationally dynamic context the conclusion about the limits of government policy-making reached by Heiner (1983).

References

Alchian, A.A. (1950), 'Uncertainty, evolution and economic theory', *Journal of Political Economy*, vol. 58, pp. 211–22.

Berlcerowicz, L. (1986), 'Enterprises and economic systems: organizational adaptability and technical innovativeness', in Leipold H. and A. Schiller (eds.), *Zur Interdependenz von Unternehmens-und Wirtschaftsordnung*, Schriften zum Vergleich von Wirtschaftsordnungen, Band 38, Stuttgart, G. Fisher Verlag.

Buchanan, J.M. (1975), *The Limits of Liberty*, Chicago, University of Chicago Press.

Cazes, B. (1986), *Histoire des futurs*, Paris, Seghers.

Day, R.H. and Eliasson, G. (eds.) (1986), *The Dynamics of Market Economies*, Amsterdam, New York, Oxford, North Holland.

Demsetz, H. (1967), 'Toward a theory of property rights', *American Economic Review*, vol. 57, Proceedings 347–59.

—— (1969), 'Information and efficiency: another viewpoint', *Journal of Law and Economics*, vol. 12, pp. 1–22.

Eliasson, G. (1984), 'The micro-foundation of industrial policy', in A. Jacquemin (ed.), *European Industry: Public Policy and Corporate Strategy*, Oxford, Oxford University Press.

Hanson, P. and Pavitt, K. (1986), The comparative economics of research development and innovation in East and West: a survey, Science Policy Research Unit, University of Sussex, mimeo.

Hayek, F.A. (1967), *Studies in Philosophy, Politics, and Economics*, Chicago, University of Chicago Press, and London, Routledge & Kegan Paul.

Heal, G. (1973), *The Theory of Economic Planning*, Amsterdam, New York, Oxford, North Holland.

Heiner, R.A. (1983), 'The origin of predictable behavior', *American Economic Review*, vol. 83, pp. 560–95.

Hirschman, A.O. (1970), *Exit, Voice and Loyalty: Responses to Decline in Firms, Organizations and States*, Cambridge, Mass., Harvard University Press.

Hofstadter, D.R. (1979), *Gödel, Escher, Bach*, New York, Basic Books.

Hurwicz, L. (1971), 'Centralization and decentralization in economic processes', in A. Eckstein (ed.), *Comparison of Economic Systems*, Berkely, University of California Press.

Knight, F. (1921), *Risk, Uncertainty and Profit*, Boston, Houghton Mifflin.

Kocanda, R. and Pelikan, P. (1967), 'The socialist enterprise as a participant in the market', *Czechoslovak Economic Papers*, vol. 9.

Kornai, J. (1980), *Economics of Shortage*, Amsterdam, North Holland.

Marris, R. and Mueller, D.C. (1980), 'The corporation, competition and the invisible hand', *Journal of Economic Literature*, vol. 18, pp. 32–63.

Mongin, P. and Walliser, B. (1987), 'Infinite regressions in the optimizing theory of decision', in B. Munier (ed.), *Risk, Rationality and Decision*, Dordrecht, D. Reidel.

Nelson, R.R. (1981), 'Assessing private enterprise: an exegesis of tangled doctrine', *The Bell Journal of Economics*, vol. 12, pp. 93–111.

Nelson, R.R. and Winter, S.G. (1982), *An Evolutionary Theory of Economic Change*, Cambridge, Mass., Harvard University Press.

Neuberger, E. and Duffy, W. (1976), *Comparative Economic Systems: A Decision-Making Approach*, Boston, Allyn & Bacon.

Parkinson, C.N. (1957), *Parkinson's Law and Other Studies in Administration*, New York, Ballantine Books.

Pelikan, P. (1985), 'Private enterprise v. government control: an organizationally dynamic comparison', Working Paper 137, Industriens Utredningsinstitut, Stockholm.

—— (1986), 'Institutions, self-organization, and adaptive efficiency: a dynamic assessment of private enterprise', Working paper 158, Industriens Utredningsinstitut, Stockholm.

—— (1987), 'The formation of incentive mechanisms in different economic systems', in S. Hedlund (ed.), *Incentives and Economic Systems*, London, Sidney, Croom Helm.

Polanyi, M. (1967), *The Tacit Dimension*, Garden City, NY, Doubleday-Anchor.

Schumpeter, J.A. (1934), *The Theory of Economic Development*, Cambridge, Mass., Harvard University Press.

—— (1942), *Capitalism, Socialism, and Democracy*, New York, Harper & Row.

Simon, H.A. (1955), 'A behavioral model of rational choice', *Quarterly Journal of Economics*, vol. 69, pp. 99–118.

Vanek, J. (1970), *The General Theory of Labor-Managed Market Economies*, Ithaca, NY, Cornell University Press.

Ward, B. (1958), 'The firm in Illyria: market syndicalism', *American Economic Review*, vol. 48, pp. 566–89.

Williamson, O.E. (1975), *Markets and Hierarchies: Analysis and Antitrust Implications*, New York, Free Press.

Winter, S.G. (1971), 'Satisficing, selection and the innovative remnant', *Quarterly Journal of Economics*; vol. 85, pp. 237–61.