Industrial Structures and Dynamics: Evidence, Interpretations and Puzzles

GIOVANNI DOSI, FRANCO MALERBA, ORIETTA MARSILI and LUIGI ORSENIGO

('Dipartimento di Scienze Economiche, Università degli Studi di Roma 'La Sapienza', Via A. Cesalpino 12/14, I-00161 Roma, University of Brescia and CESPRI, Bocconi University, Milan, Department of Public Economics, University of Rome, Italy and SPRU, University of Sussex, Falmer, Brighton, UK and Istituto di Economia Politica, Bocconi University, Via Gobbi 5, 20136 Milan, Italy)

1. Introduction

The idea that industries ought to be analysed in an explicitly dynamic perspective is certainly not new. The interest in the patterns of industrial dynamics and evolution was clearly at the heart of industrial economics in its formative stages. Both Marshall and Schumpeter, in radically different ways, considered the study of industries as primarily a dynamic exercise. Both viewed competition as a dynamic process by which less efficient firms tend to decline in profit and market shares, shrink and eventually exit, while new, more efficient firms enter the industry, grow and gain market shares. Similarly, both suggested (together with Kuznets and Colin Clark among others) that a loosely similar process takes place among industries in the longer run, with some industries growing and others declining.

The subsequent developments in industrial economics (for reasons that cannot be thoroughly discussed here) did not follow these insights. Both the 'structure-conduct-performance' and the 'new industrial organization' (IO) approaches concentrated their attention on essentially static analyses of the structure of industries. Indeed, if anything, things got worse as the theoretical commitments to microeconomic rationality and collective equilibria became heavier. Relatedly, subjects like entry and exit, or the patterns of evolution of market structures over the long term, have not been considered for a long time as primary objects of research.

More precisely, a good deal of investigation during the 1960s focused the...
analysis on 'snapshots' over industrial structures, highlighting such themes as size-related economies of scale and other absolute cost advantages as a major source of asymmetry between incumbents and entrants and among incumbents themselves. Then some direct links were established between structural conditions—so defined—with the degrees to which firms can exercise oligopolistic power and, as a consequence, also with the average price and profitability performance of the various industries (cf. Bain, 1956; Mason, 1957; Sylos-Labini, 1967). The primary question therefore concerned the determinants of industrial performances at any time [even if a few authors within this 'structuralist' tradition are rich of insights, unfortunately neglected far too long, on the endogenous dynamics coupling technical progress, changes in industrial structures and aggregate performances: see in particular Downie (1958), Sylos-Labini (1967) and Steindl (1976); a few more comments are in Dosi (1991)].

As known, the line of attack against this 'old' breed of industrial economics has been primarily directed against the lack of explicit behavioural foundations (taken most often to mean 'rational', maximizing conducts) which would link given technological and market conditions with revealed performances. This is precisely what the 'structure—conduct—performance' paradigm attempted to do (for a synthesis, cf. Scherer, 1980). A general shift in emphasis followed—which became a landslide with the contemporary 'new industrial organization' approaches—toward an almost exclusive focus on 'conducts' and an increasingly casual account of the technological and competitive specificities constraining behaviours and collective outcomes.

At the same time, the commitment to the idea, especially pronounced in the new IOs, that each empirical observation should be rationalized as the equilibrium result of strategic interactions among highly sophisticated forward-looking agents, tended to rule out any appreciation of all those aspects of industrial dynamics involving unexpected discovery, far-from-equilibrium competition between more and less efficient entities, and such rather messy processes as natality and mortality. Nonetheless, these aspects seem indeed to form the picture emerging from the rapidly growing number of empirical contributions over the last 15 years on how industries change over time.

In turn, renewed interest in this set of phenomena is due to few complementary reasons. Firstly, the growing availability of large micro, longitudinal databases at the firm level has allowed analysts to measure and account for phenomena like entry, exit and firm growth.1 Secondly,
developments in computing resources and techniques have made it possible to handle extensive data and the calculations needed for studying them. Thirdly, the dissatisfaction with dominant modes of looking at industrial organizations spurred the production of models, theories and methodologies (both through simulation and analysis) which have industry dynamics as their primary focus.

The evidence that has been produced thus far already challenges in many important ways the standard picture of industries and it is starting to change deeply the conventional way of analyzing industries. As witnessed by the papers forming this Special Issue of *Industrial and Corporate Change*, this evidence is far from homogeneous, presupposes different levels of observation, and is grounded on different methodologies. Nonetheless, it is starting to highlight a few 'stylized facts' and point to challenging interpretative issues.

By way of an introduction to the articles that follow, let us consider some of these issues.²

### 2. Empirical Regularities

Let us begin with the evidence concerning some features of the dynamics in (i) *industrial structures*, broadly understood to cover characteristics of firms such as size, productivity, innovativeness and their intra-industry distributions; (ii) *performances*—including individual profitabilities, growth profiles and survival probabilities, together, again, with their aggregate distributions—and (iii) *patterns of technological and organizational learning*—e.g. modes of innovative search, knowledge bases upon which firms draw, the relative importance of process and product innovations, etc.

As a preliminary *caveat* it might be worth drawing a distinction between two different, albeit highly complementary, methodologies in extracting 'stylized facts' from such dynamic evidence—also reflected in the papers that follow. One approach primarily focuses upon invariances and differences across industries in some properties of industry dynamics—e.g. entry and exit, degrees of heterogeneity across firms, patterns of innovation (cf. Jensen and McGuckin, Audretsch, and Malerba and Orsenigo, in this volume). As such, it is a sort of 'cross-sectional' exercise in comparative dynamics. Another approach (reflected in the papers by Carroll, Klepper, and Afuah and Utterback) takes a closer look at the history of particular industries—typically starting from their birth—and tries to detect the microeconomic

² This discussion is partly based on Dosi *et al.* (1995), where one can find also a more extensive review of the literature.
determinants of particular time-series observations in industrial dynamics itself. The fact that the two methodologies carry a number of overlapping conclusions is highly encouraging about the robustness of the results (but it also raises some puzzles, which we shall briefly mention below).

Microeconomic Heterogeneity

Within industries one observes significant heterogeneity in firms’ characteristics, behavior and performances. As Jensen and McGuckin report in this volume, firms differ along practically any observable dimensions, such as size, age, productivity, wage, job creation and destruction, investment patterns and innovative activities. Heterogeneity emerges at any level of aggregation, and not only in cross-section but also over time: during any time interval observed changes among firms in the same industry are uneven and idiosyncratic. Moreover, diversity is not simply associated with traditional ‘explanatory variables’ like location, industry, size, age or capital. Rather, it appears to be associated to a much larger degree with unobserved firm- or business unit-specific factors (cf. Schmalensee, 1985; Nelson, 1991; Rumelt, 1991). Heterogeneity is also the core of both the industry life cycle (Klepper in this volume) and the organizational ecology (Carroll in this volume) approaches. In the former, firms differ in terms of their innovative capabilities and timing of arrival and these differences are dynamically reflected by differential growth rates and firms’ size. In the latter, the unit of analysis is primarily different types of organizations, rather than individual firms as such, and the goal of the research program is precisely to account for their relative diffusion within the industry over time.

Persistence

Equally important, such diversities appear to be persistent. For example, as far as innovation is concerned, high (low) innovators at time $t$ have—ceteris paribus—a higher probability of remaining high (low) innovators at time $t + 1$ (Griliches, 1986; Pavitt and Patel, 1991; Cefis, 1996; Malerba et al., 1997; for a different assessment of the evidence, see Geroski et al., 1996). Similarly, one observes significant persistence in plant productivity and, in general, most of the unobserved specific factors that influence plant and firm heterogeneity appear to be rather persistent over time (Jensen and McGuckin, this volume). These results are consistent with the evidence concerning persistence in profits: firms enjoying higher (lower) profits are expected to earn higher (lower) profits in the
future as well (Geroski and Jacquemin, 1988; Mueller, 1990). That is to say, profits do not seem to converge on a common rate of return.

Persistence is also a major feature highlighted by 'life cycle' studies, the evidence from which suggests (often but not always) some form of dynamic increasing returns underlying first mover advantages (cf. Klepper, this volume). And, of course, persistence of both differential micro-features and performances should be expected to be even higher in the 'organizational ecology' interpretation of the evidence—whereby a strong emphasis is put upon the inertial reproduction of the organizational and sociological birthmarks of different types of firms (cf. Carroll, this volume).

A more subtle issue, however, concerns the distinction between transitory and persistent features of those firm-specific characteristics and performances. One observes, in general, regression-to-the-mean patterns, with long-term means being themselves different across firms, while the speed of adjustment is a subject of debate (on these issues, cf. Baldwin, 1995; Geroski and Jacquemin, 1988; Jensen and McGuckin, this volume).

Turbulence

Most industries are characterized by high degrees of turbulence, due to entry, exit and changes in market shares, as discussed in the paper by David Audretsch in this volume. Birth rates are high, even in industries characterized by high barriers to entry (e.g. as proxied by capital intensity). Exit rates tend to be high too, however, resulting in high turbulence and smaller net entry as compared to gross entry. In general, a positive correlation is observed between rates of entry and rates of exit across industries (Dunne et al., 1988). Also Malerba and Orsenigo (this volume) found significant degrees of turbulence among innovators in certain sectors and technologies (what they call call ‘Schumpeter Mark I’ sectors).

Interestingly, together with the phenomenon of high degrees of turbulence, one observes that most entrants are small firms, far below any measure of a minimum efficient scale, and that a large percentage of entrants exit the industry within a few years after entry. The probability of survival appears to increase with size and age. Surviving entrants show either a higher initial size or higher growth rates, and a negative correlation is observed between these two latter variables: bigger initial size often implies lower growth.

Notice also that industry profitability does not seem to have any significant

---

3 However, entrants include firms operating in different industries or countries engaged in a process of diversification. The role of this second type of entry is small in terms of number of firms, but much more important in terms of share of output or employment.
effect on entry and exit, which are instead positively correlated with industry growth. Similarly, entry barriers (e.g. capital intensity) do not appear to affect entry but rather survival probabilities (cf. Audretsch, this volume).

Turbulence is clearly also a major feature of the industry life cycle interpretations (see Klepper, this volume). In the early stages, entry tends to be high and higher than exit, so that the number of producers increases. At a later stage, the number of firms in the industry experiences a shake-out, due both to a decline in the rates of entry and an increase in the rate of exit. In the (partly overlapping) interpretation based on ‘organizational ecology’ approaches (see Carroll, this volume), the prediction, borne by a few sectoral studies, is that the dynamics of entry and exit of specific types of organizations is a (non-linear) function of the density of the latter. Entry rates are expected to rise and then attenuate while exit rates are expected to fall and then climb, depending on absolute and relative frequencies.

Mechanisms of Productivity Growth

An overwhelming amount of evidence converges on the description of the typical dynamic patterns of industries as involving continuous (albeit not time-invariant) entry, exit and expansion/contraction of incumbents. However, there is much less agreement about the relative importance of these processes for the productivity growth of industries. According to certain studies, entry and exit are relatively unimportant in aggregate productivity growth, despite the large turnover and the long period of time under consideration (Baily et al., 1992; Jensen and McGuckin, this volume). Other studies suggest instead that turbulence is a major determinant of the dynamics of industry productivity, at least when the cumulated growth of entrants and exiters over longer time spans is properly accounted for (Baldwin, 1995). To some extent, these different empirical assessments might simply reveal different methodologies for computing the phenomena under consideration, or reflect country-specific factors. In any case, it is an issue worth further exploration, since it also carries deep implications for understanding the nature of the competitive process and its efficiency properties.4

Invariant Structural Patterns

Jointly with heterogeneity and turbulence, one also observes some striking

4 As an additional example consider the empirical studies concerning persistence of profits. According to some studies, persistence is mainly the outcome of some degree of sluggishness of the competitive process. According to other studies, the adjustment to firm-specific “permanent” values is quite rapid. See Mueller (1990).
One of them is certainly the persistence over time (at least since one has the data) of a skewed distribution of both firm and plant size—approximating a Pareto distribution—that is rather similar in manufacturing industry over time and across developed industrial countries. The explanation of this phenomenon has usually been sought in some version of Gibrat's Law of proportionate growth. In its simplest formulation, the Law states that firms' growth rates are i.i.d. random variables independent of size. This process generates a log-normal distribution (or, under suitable modifications, Yule or Pareto distributions) of firm size (Simon and Bonini, 1958; Ijiri and Simon, 1977). Evidence, however, is mixed: many analyses tend either to reject this law or to restrict its applicability to particular size classes. While some recent studies show that both firms' growth rates and their variance tend to fall with size and age (Evans, 1987a,b; Hall, 1987), others (Acs and Audretsch, 1990) corroborate Gibrat's Law, when the samples observed included firms that were exiting the industry. Nevertheless, Gibrat's Law seems better suited to describe the growth process of relatively big firms (Geroski and Machin, 1992). We shall come back to the interpretations of growth processes.

Industry-specific Features

Significant industry-specific differences emerge from the data. The observation that variables like capital intensity, advertising intensity, R&D intensity—along with structural measures like concentration and performance measures like profitability—differ widely across sectors is at the very origin of the birth of industrial economics as a discipline. Jensen and McGuckin (this volume) find that industry-specific effects also significantly influence firms' heterogeneity, even if most of the observed variance in plants and firms characteristics is mainly within industries. Thus, it should not come as too big a surprise that phenomena like entry, exit and survival, persistence in firms attributes and performances, and innovative activities and firms' growth also exhibit significant inter-industry variability. In this volume David Audretsch reports on the relationships between entry, exit and survival of entrants on the one hand, and industry characteristics like the rate of innovation and capital intensity on the other. The evidence suggests, in particular, that survival is easier in those industries in which small firms are important sources of

1 Other studies (for example, Mueller 1990 and Geroski and Jacquemin 1988) showed that the persistence of profits also appears to depend on industry-specific characteristics as well as on firm-specific characteristics. In particular, industry-specific features such as the intensity of advertising and R&D appear to be highly correlated with the persistence of higher than average profits.
innovation and that new surviving firms tend to grow faster in innovative industries and as a function of the gap between minimum efficient scale of output and firm size. At the same time, however, the likelihood of survival decreases as a function of that gap. The same happens in terms of innovation rates. Thus, factors that promote growth reduce survival and vice versa. And even the life cycle patterns, as Steven Klepper discusses in this volume, show some important differences across industries.

Patterns of Technical Change and Industrial Dynamics

So far as innovation and diffusion are concerned, it is now very well known that technological change proceeds along quite distinct patterns, according to the different industries and technologies—in terms of rates, directions, sources of innovative opportunities, and mechanisms by which these opportunities are tapped and exploited (for discussions of the evidence, see Rosenberg, 1978, 1982; Freeman, 1982; Dosi, 1988). A similar diversity characterizes the organization of innovative activities and the organizational competences involved in the innovation process [following the seminal interpretation by Nelson and Winter (1982), see, among others, Levin et al. (1995), Teece et al. (1992, 1994).]

Nonetheless, a major question that is only beginning to be explored concerns relationships, if any, between (i) the patterns of technical change; (ii) the characteristics of the firms undertaking it; and (iii) the ensuing patterns of industrial evolution. For example, Pavitt's widely used taxonomy (Pavitt, 1984) is an early attempt to classify groups of sectors according to some shared features of the innovation process and, together, identify the modal characteristics of the innovating actors (e.g. small/big, specialized/diversified firms, etc.).

In this volume, Malerba and Orsenigo interpret the data on innovative patterns—as proxied by patenting—in terms of two distinct archetypes. In some industries, the patterns of innovative activities resemble the model described by Schumpeter in the Theory of Economic Development (1912), characterized by technological ease of entry, a major role played by new firms in innovative activities, and a continuous erosion of the competitive and technological advantages of the established firms in the industry. In other industries, technological change is more similar to the description provided by Schumpeter in Capitalism, Socialism and Democracy (1942), with strong barriers to entry for new innovators, the prevalence of large established firms in innovative activities and the dominance of a few firms which are continuously innovative through the accumulation over time of technological
capabilities. What is more interesting is that such inter-industry differences are similar across countries. In particular, the patterns of innovative activities tend to be remarkably similar across countries in the same technological classes (Malerba and Orsenigo, this volume). Note also that highly concentrated sectors tend to be the same in all developed countries (Sutton, 1991). These observations together suggest that some 'structural' factors exist that are rather invariant across countries within the same industries and shape in similar ways the patterns of industry dynamics (see also below).

3. Puzzles and Questions

The evidence discussed in this volume certainly provides new insights into old questions but opens also many new challenges.

For example, as forcefully stated by Jensen and McGuckin, given the evidence, the notion of a 'representative agent' can hardly be a building block for any empirically grounded theory of industrial dynamics. Equally, any account of entry processes as 'equilibrating phenomena' is seriously challenged by the evidence discussed in Audretsch's paper. Moreover, one is beginning to achieve a sharper understanding of the factors affecting discontinuities in industrial structures over the evolution of industries (cf. the papers by Klepper and by Carroll).

Together, this same evidence raises fascinating questions. For example:

Why do so many heterogeneous plants operating at a sub-optimal scale exist and how are they able to co-exist within the same industry? Why do so many entrepreneurs continue to enter, even in industries characterized by substantial barriers to entry, despite their low probability of survival? And how is it that profitability does not appear as a significant variable in the determination of entry rates? Or does perception of some unexploited opportunity of innovation drive entry?

What can explain the persistence of profits and other dimensions of firms' performance? Does it reflect the sluggishness of imitation and competitive processes or the persistence of differential technological capabilities and productive efficiencies which are not eroded away by the competitive process? And in this latter case, what are the sources of such differentials?

What are the mechanisms that generate the observed regularities in the long-run patterns of evolution of industries (as in the industry life cycle approach)?

These questions appear even more challenging when these pieces of evidence are considered jointly. For instance, what are the relationships between the evidence presented by Klepper in this volume and the 'aggregate'
statistics concerning entry and exit as measured and discussed by Audretsch in this volume? Or, how does one reconcile the persistent diversity among firms and the relationships between firms’ size, age, survival and growth discussed previously, with that of relatively stable Pareto-type size distributions on the other?

In general, what is particularly intriguing is the coexistence of turbulence and change on the one hand, with persistence and regularities at different levels of observation—from individual firms’ characteristics to industrial aggregates—on the other. Industrial dynamics and evolution appear neither to be simply characterized by random disorder nor by perfectly self-regulating, equilibrium processes that quickly wipe away differences across firms. Rather, the evidence accumulated so far seems to suggest a subtle and intricate blend of these two elements.

In fact, the underlying analytical issues have implications well beyond industrial economics. Consider, for example, the following:

The ‘driving forces’ of firms’ entry and growth. It is perhaps an additional stylized fact (as it appears from the papers in this issue) that technological change is now widely recognized to constitute a primary source of industrial dynamics. But how should technological change be characterized? As a process of convergence to some constant (exogenously given) optimal technique? Or as a continuous, endogenous process of generation of better techniques? More generally, are innovation and the other forces that drive industrial change characterized simply by small, multiple and temporary shock? Or is it possible to identify deeper, more persistent factors, perhaps not directly measured by and therefore non-observable in the available data?

Micro-heterogeneity and invariances in aggregate statistics. What kind of processes are able to generate such aggregate regularities (as those mentioned above) on the basis of such messy microdynamics?

Cross-sectional variety in industrial structures and dynamic patterns. What determines these observed differences? Can one find some underlying variables—relatively invariant across countries—which shape different dynamic patterns? And what is the degree of detail in the description of industries (agents, technology, demand) necessary for providing a satisfactory account of both similarities and differences across sectors?

4. Evidence and Theoretical Interpretations: An Uneasy but Fruitful Tension

Taken together, the pieces of empirical evidence discussed in this special issue of ICC appear to vindicate the relevance of an explicitly dynamic approach to
the study of industrial organization. However, the theoretical interpretations of the 'stylized facts' reviewed above and, in general, of the underlying processes of 'dynamic competition' widely differ.

Without even attempting to provide a 'taxonomy' of the current, rapidly expanding theoretical literature, let us just point out a few crucial distinctions that might help in identifying some major conceptual issues and possible research agendas.6

Characteristics of Agents: Rationality, Learning and the Sources of Heterogeneity

A first fundamental distinction concerns the representation of agents. Agents (firms) can be characterized by different cognitive capabilities, ranging from hyper-rationality (as in Lucas, 1978) to rather inertial behavioural repertoires (as in the models of organizational ecology, where processes of adaptation by agents are generally not explicitly considered).

The hypothesis of hyper-rationality is sometimes weakened by the introduction of imperfect information about some relevant variables. For example, Jovanovic (1982) assumes that agents have rational expectations about technology, but are uncertain about their own production efficiency; and Ericsson and Pakes (1996) model the relative efficiency of each firm as the stochastic outcome of the investment in research by the firm itself and by all the competitors. However, even when various forms of technological uncertainty and idiosyncratic efficiency shocks are admitted, the 'rationalist' approach [certainly comprising the works just mentioned, Hopenhein (1992) and others] still places an enormous weight upon the knowledge and computing abilities of agents who must perfectly understand the 'model of the world' in which they operate (including the stochastic process driving technological shocks).7

Conversely, in evolutionary interpretations (Nelson and Winter, 1982; Winter, 1984) 'bounded rationality' also takes the form of limited understanding by the agents of the causal structure of the environment in which they are embedded, a much lower ability to think through future contingencies, and behavioural patterns often describable in terms of relatively invariant routines. On the other hand, in this approach agents are capable of learning and thus improve their performance (typically, technology) over time.

6 A more detailed discussion of modelling issues is in Dosi et al. (1993).

7 Klepper's model (Klepper, 1996) of the industry life cycle takes an intermediate position, with rational (i.e. maximizing) but myopic agents.
The issue of the cognitive abilities of the agents purported by various theories partly overlaps with two major issues, namely the sources of firms' heterogeneity and their persistence, and the characteristics of organizational learning processes.

In some models (e.g. Lucas, 1978; Jovanovic, 1982), heterogeneity rests on some intrinsic characteristics (called somewhat metaphorically 'talent' or 'ability'). Agents have them from the start and keep them throughout [except that in Jovanovic (1982), at the beginning agents do not know what they are, and thereafter are subject to firm-specific random shocks].

The idea that idiosyncratic shocks make for part or all of the heterogeneity across firms is common to a few other models. For example, Ericsson and Pakes (1996) and Hopenhain (1992) assume that 'states of success' or productivities follow a stationary Markov process. And, of course, in older Gibrat-type interpretations, the whole story of why firms differ in size is (more parsimoniously) derived from small idiosyncratic shocks—upon otherwise identical agents—whose effects, however, cumulate over time (cf. Ijiri and Simon, 1977).

The fact that firm-specific (or even intra-firm, line-of-business-specific) sources of heterogeneity are a crucial part of the understanding of the competition process is well borne out by the micro-evidence and taken on board across different modelling approaches. However, the details of how one formally captures the phenomenon often hide major disagreements on the underlying conjectures regarding why firms differ, if and how they are able to learn, and how persistent heterogeneity is.

Keeping the technicalities aside as much as possible, it is intuitive, for example, that heterogeneity tends to disappear in the long run when 'shocks' (i) are idiosyncratic (i.e. firm specific), but (ii) are generated by the same distribution which holds across firms and (iii) affect levels rather than rates of change in the relevant variables (whether these be productivity or any other performance indicator). Conversely, it is equally intuitive that if one mainly attributes the sources of heterogeneity to some time-of-birth feature of individual agents (as organizational ecology models tend to do), the implied conjecture is that persistence of the relevant micro-trait is quite high.

The degree to which agents can change through time and the ways they do so link, of course, with the nature of learning processes occurring within individual organizations. The issue has been addressed in particular by models of evolutionary inspiration. The basic ideas here are that (i) initial
micro-conditions certainly count (e.g. 'skills', organizational routines, 'clever' strategies), but also (ii) path-dependent accumulation of further competences may well take place (notwithstanding the possibility of systematic mistakes). Here is also where the evolutionary approaches tend to link up with valuable insights from the business and strategic management literature. Both identify persistent sources of heterogeneity in firm-specific learning paths and differential learning capabilities (Winter 1988, 1995; Teece and Pisano, 1994). Hence, agents are heterogeneous also because they accumulate differential problem-solving capabilities which unfold over time, have access to different technologies, and these differences are amplified in the process.9

Nature and Properties of Dynamic Processes

The symmetric complement of the assumptions on what agents know, learn, do concerns what markets (and other interaction environments) do. Observed industrial dynamics are obviously the joint outcome of both. But it makes a lot of difference (except for some rather peculiar circumstances), in terms of the properties of the dynamics themselves, whether and to what extent individual entities can figure out, so to speak 'in their heads', ex ante, what is going to happen to them, at least in probability, because they also know (and possibly collectively share) a common 'model' of their environment and shape their decision accordingly.

In the extreme view whereby everyone knows ex ante everything that is relevant to know—about, for example, technologies, distribution of 'talents' or other causes of heterogeneity across the population of agents, strategies, etc.—markets operate essentially as collective arrangements setting incentive-compatible schemes. But since agents 'work it out' beforehand, not much happens through the markets themselves, the consistency of individual plans (as Frank Hahn would put it) being guaranteed by the (certainly 'hyper-rational') assumptions on micro-knowledge.

Of course, this view implies also that empirical observations—such as those richly presented in this volume of ICC—should in principle be interpreted as sequences of equilibrium outcomes, nested into collectively consistent, highly sophisticated, plans of intertemporally maximizing agents [and this is indeed the spirit, if we understand it correctly, by which Lucas (1978) or Hopenhein

9 More technically, in some fully fledged evolutionary model (still to come), (i) attainable states, at each time, are dependent upon past realizations up to that time, but notionally, opportunities of discovery are unlimited as times goes to infinity; and (ii) transition probabilities are themselves firm-specific since they depend on 'idiosyncratic' accumulation of learning competences.
(1992), for example, try to account for the evidence on skewed distributions of firms’ sizes, positive rates of entry and exit, etc.]

At the opposite extreme, suppose that agents do not have a clue about what is going to happen to them (or to the same effect, that they hold a rather wild distribution of beliefs largely uncorrelated with what is actually going to occur to them). Under these circumstances, markets operate first of all as selection devices, determining, ex post, profitabilities, survival probabilities and expected rates of growth. Under these micro-assumptions, the same empirical observations need to be interpreted as the collective properties of some interactive dynamics, given observationally justified generalizations concerning, for example, populations’ characteristics and their distributions, criteria and speed of selection.\(^{10}\) Outside these two extremes, the challenge is to understand how joint processes of micro-learning and collective selection support the dynamic patterns highlighted by the evidence. And, indeed, the papers in this volume largely point in this analytical direction.\(^{11}\)

Having acknowledged that, however, is only the beginning of the analytical story.

How much interpretative weight should one put upon ‘rational’ forward-looking adjustments, given an existing landscape shaped by unchanged ‘fundamentals’, versus innovative learning and discovery, versus market selection among agents who might well get it systematically wrong?

In this respect, consider, first, the story formalized in Jovanovic (1982). Notwithstanding heterogeneity, a good part of the explanation of the empirical phenomena addressed (e.g. regularities in size distributions) rests on the purported sophisticated algorithms of adaptation of the agents. Market selection is in some sense ‘anticipated’ by each agent, who accordingly plans for equilibrium production schedules. Moreover, the universe of available techniques is given from the start: hence, ‘learning’, in the sense of discovering something genuinely new, is ruled out. New firms continuously enter the industry, with rational technological expectations about the population-level distribution of some technological variable but uncertain about their own production efficiency. By producing they acquire noisy information about their productivity. Those firms who discover that their efficiency exceeds their expectations expand their scale of output, while those

\(^{10}\) Interpretations based on ‘pure selection’ and ‘pure ex ante rationality’ happen to be equivalent whenever the underlying equilibria coincide, and, together, each empirical observation might be understood to be a rather close approximation to the ‘limit’ (in a mathematical sense) of some adjustment process operating at a time scale of order of magnitude faster than that at which empirical observations themselves are collected. Frankly, we find this possibility, at best, rather awkward as a general interpretative framework.

\(^{11}\) See also Baldwin (1995) for an empirical analysis of the relative role of processes of learning and selection in industrial dynamics.
who receive unfavourable signals contract or even exit from the industry. Thus, industry evolution is driven by noisy selection which fosters the growth of efficient firms and the decline of inefficient ones. The observational predictions of the model (on size distribution, entry and exit) hold as limit properties of the processes: hence, each empirical observation—taking the model at its face value—should be understood as a (nearly) equilibrium outcome of a much faster adaptation dynamics.

Compare that view with evolutionary interpretations, broadly inspired by Nelson and Winter (1982), such as Dosi et al (1995). There, as already mentioned, the commitment to individual rationality is much lower and, symmetrically, the explanatory burden placed upon market selection is correspondingly higher. An explicit market dynamics is assumed. Change is not driven by (possibly noisy) adjustments to given 'fundamentals' (at least with regard to best-practice techniques). Rather, innovation is the main engine of dynamics and evolution. As biologists would say, the 'evolutionary landscape' upon which evolution occurs is not fixed, but is continuously deformed by the endogenous learning activities of agents. Relatedly, the thrust of the explanation of empirical phenomena also differs and one tends to interpret the aggregate regularities that are observed in the data as emerging from disequilibrium interactions among heterogeneous agents on the basis of some well-specified dynamic process.

**Aggregation and Emergent Regularities**

In the wealth of 'stylized facts', there is a problem of interpretation of statistical data and in particular of aggregation. The data that are available and the regularities that are observed refer actually to entities and phenomena that are defined at different levels of aggregation and at different time scales. For example, in the life cycle approach, the emphasis is sometimes on narrowly defined products and other times on 'industries'. But how would this approach consider, for example, mainframes and personal computers? Are they to be taken as different products with distinct life cycles or as variants of the same basic product and of the same life cycle? Probably the interpretation of patterns of evolution would turn out to be quite different in the two cases (Bresnahan and Malerba, 1996; Malerba and Orsenigo, 1996). Similarly, at what level of aggregation are the regularities on the distribution of firms' sizes to be interpreted? And how does one reconcile this evidence with the observation that this distribution typically changes over the industry life cycle? Or, finally, at what level of disaggregation should the notion of technological regimes be defined?
This issue is likely to go beyond the conventional problem of definition of the relevant unit of analysis and of the proper way of identifying an 'industry'. For example, it may well be that certain regularities that are observed are essentially the outcome of the aggregation of heterogenous dynamic processes occurring simultaneously at different levels. Dosi et al. (1995) show via simulations that a persistent skewed distribution of firms' sizes and the relationships between size, age and firms' growth that are observed in the 'real' data can actually be obtained as the aggregation of different 'sectors', each characterized by different dynamic processes (learning regimes) that do not exhibit any of the properties that emerge at the aggregate level. Thus, more substantially, this problem touches upon the more general and fundamental question that was raised above of how 'aggregate regularities' or some kind of 'order' can be generated by heterogenous unit of analysis (firms, products, industries) in a dynamic process and at which level of observation one detects them.

5. Modelling Inter-sectoral Differences in Structures and Dynamic Patterns

An additional challenge for any interpretation of the dynamics of industries is given by the need to account for the inter-sectoral differences in the patterns of industrial change. That is to say, one has to ask whether in different industries the same dynamic processes are able to account for the observed differences (with modifications of some selected parameters) or different dynamic processes are at work. Moreover, it may well be that the behaviour of certain variables exhibits systematic differences not only across sectors, but also and perhaps primarily at different stages of the dynamic process.

As a typical illustration of this kind of problem, one can consider the explanations of the patterns of entry, exit, survival and innovative activities that are provided in this volume by Audretsch, and Malerba and Orsenigo on the one hand, and by Klepper, Carroll, and Afuah and Utterback on the other. Carroll, Klepper, and Afuah and Utterback interpret differences in the patterns of industrial dynamics as a consequence of the fact that industries are at different stages of their life cycle. Conversely, Audretsch and Malerba and Orsenigo claim that inter-sectoral differences depend mainly on some industry-specific variables—in particular, on the nature of the underlying technological regime, i.e. on the properties of the processes through which firms innovate. [The notion dates back to Nelson and Winter (1982) and Winter (1984), with some refinements thereafter.] Audretsch classifies technological regimes as 'entrepreneurial' and 'routinized', in terms of accessibility to
technological knowledge by firms external to the industry as opposed to incumbents. He presents evidence showing that turbulence is higher and that the survival rate of young firms is lower in industries characterized by an entrepreneurial technological regime. Malerba & Orsenigo propose that a technological regime is a particular combination of opportunity and appropriability conditions, degrees of cumulativeness of technological knowledge and characteristics of the relevant knowledge base. They show that these variables are important in explaining the patterns of innovation across technologies and industries. Afuah and Utterback suggest an interpretation, rich on micro (behavioural and 'strategic') evidence, which begins to merge the two points of view.

Strategies, Learning Regimes and History

There is a much larger issue here concerning why industries in general differ, first, in their structures and performances [as analysed at length with mixed results by traditional industrial economics: cf. the review in Schmalensee (1989)], and secondly in their dynamic profiles [as discussed by the works which follow and elsewhere, e.g. Geroski (1994) and Baldwin (1995)].

At one extreme, the answer to this question is that these differences can be rationalized as different equilibrium outcomes of some underlying strategic interaction. However, even accepting the idea that empirical observations should be understood as the equilibrium outcome of something or other (which a few analysts are not prepared to do: see above), the point remains almost entirely tautological and void of any empirical content as long as one does not specify which games people are playing and why these games differ in different sectors.

At the other extreme, interpretations leaning toward an evolutionary approach emphasize sector/technology-specific patterns and invariances in the ways agents learn (i.e. learning regimes or whatever other name is chosen) and market-specific invariances in the ways agents interact (i.e. selection regimes) as determinants of both structures and dynamics.

Note that in the former approach—not far from the spirit of a good deal of IO—the 'strategic aspect' (in a strict game-theoretic sense) is what drives the explanation. Conversely, evolutionary approaches go a long way in 'de-strategizing' the analysis. This does not mean, of course, that they do not acknowledge that agents try to be 'strategic' (albeit in a sense nearer to the common language or business management use of the word: something like having more or less precise forward-looking plans, which might go right or wrong). However, the conjecture is that what agents do maps only very
imperfectly into the 'true' nature of the technological and market problems at hand. Rather, as a first approximation, what accounts much more for what goes on in a particular industry are, first, modal ways by which knowledge is accumulated (affecting also the propensity to invest in R&D, the probabilities of drawing a successful innovation for a given research effort, etc.) and, secondly, the ways markets deliver reward and penalties and how fast they do so.

In considering the 'evolutionary' and the 'game-theoretic' views, scholars like John Sutton (1995, 1996) share with the former the requirement of identifying 'a small number of mechanisms that appear to operate in a systematic way across the general run of industries [and whose] operation . . . induces a number of bounds on the set of outcomes that we expect to observe in empirical data' (Sutton, 1996, p. 5). At the same time, Sutton shares with the 'strategizing' view the idea that the bottom-line mechanisms have to do with a criterion of rationality (entrants in each product line cover their costs) and some non-arbitrage principle (saying more or less that all profit opportunities are exploited). Indeed, both assumptions are probably insufficient steps in the direction of rationality and equilibrium for hard core IO theorists and far too much for evolutionist ones.

Although these different approaches have to do also with rather deep 'philosophical' commitments, a more systematic analysis of the micro, longitudinal and cross-sectional evidence, as the papers comprising this issue undertake, will help in detecting the relative merits of different interpretations.

Having said that, the complementary question is how far can one go in understanding the 'stylized facts' on industrial dynamics—including those presented below—by simply 'modulating' and reparametrizing some basic common mechanisms. Or, putting it the other way around, how much history is required for a satisfactory interpretation of observed structures and dynamics? What is the appropriate detail in the specification of the nature of the agents, of markets, etc., that is necessary to account for the evolution of particular industries? For example, what would remain unchanged if one could 'twice re-run the tape' of the history of the computer industry, this time without IBM? Would concentration, productivity growth, rates of entry and survival probabilities remain more or less the same?

Certainly the gap between the richness of the histories and of their 'appreciative' interpretations on the one hand and the theoretical models that are used to 'explain' them on the other still appears quite large. Even the more complex models consider a very restricted set of agents (individuals and firms) and the representation of their internal structure is highly simplified. The only
interactions that are taken into account are market interactions (and even then in a highly stylized fashion). Yet, the historical and empirical accounts suggest that other agents such as universities and scientific institutions, political bodies and financial institutions often play a crucial role. Moreover, the internal structure of agents is an important determinant of behaviour and a variable that changes over time. Finally, these 'empirical' and 'appreciative' accounts are almost always told as histories, i.e. as sequences of events, in which the precise sequence is a crucial phenomenon to be explained and often a major determinant of later events.

Given all that, efforts to build 'history-friendly' styles of modelling (as they are called in Malerba et al., 1996) indeed go in the direction of forming a bridge between phenomenological reconstructions of 'what happened and why' on the one hand and more sweeping theories on the other. But, in turn, what more should be demanded from historically richer models than is found in more parsimonious representations?

6. Conclusion

We began these introductory notes by recalling the explicit dynamic emphasis with which early scholars, such as, in different ways, Marshall and Schumpeter, approached industrial analysis. However, such ideas as the 'trees and the forest' or 'creative destruction' retained until recently the status of suggestive metaphors without either precise empirical backing or formal representations. Things are changing deeply in both respects. Regarding the former, dynamic analyses of data on firms, technologies and industries open new avenues of interpretation of the determinants of industrial structures, performances and their changes—this Special Issue of ICC being an important contribution to this endeavour. And, regarding the theory, as we briefly discussed above, this same evidence raises fundamental issues which modellers are only starting to tackle in different analytical perspectives.

In fact, this is probably one of the few areas of socio-economic studies where there are many more facts than theories, and therefore one might have some hope of comparing the power and robustness of different analytical points of view.

We do not think it an exaggeration to say that studies like those which follow are valuable contributions at last to the general understanding of competition processes in decentralized economies and of the ways they are nested into knowledge accumulation and organizational change.
References


Teece, D., G. Pisano and A. Shuen (1990), 'Firm Capabilities, Resources and the Concept of Strategy,' CCC Working Paper no. 90-8, Center for Research in Management, University of California at Berkeley.


