1 Introduction

GIOVANNI DOSI AND
MARIANA MAZZUCATO

THIS book addresses, from a variety of perspectives, the patterns of innovation through the history of the pharmaceutical industry (which we take to include nowadays both pharmaceutical and biotechnology firms) and the ways that knowledge has coevolved with the dynamics of firm growth, industry structure and the broader socio-economic environment.

Ever since it began the pharmaceutical industry has indeed been an example of a "science-based" industry, whereby innovation is driven, to a large extent, by joint advances in pure and applied sciences, together with complementary progress in research technologies – undertaken both within public research institutions and business firms. As such, it is a fascinating industry to study. And it is even more so in the light of the profound changes that have occurred recently in the underlying knowledge bases – associated with the so-called "biotech" revolution – as well as in the broad institutional regimes governing the generation and appropriation of the economic benefit of innovations. The welfare and policy implications are equally paramount, given the socio-economic importance of the industry for health, agriculture and food production.

Let us start with a brief overview of the history of the industry, in order to give some background for the analyses that follow, and then proceed by flagging some of the central issues addressed by the chapters below.

1.1 The evolution of the industry: an overview

The history of the international pharmaceutical industry has already been extensively analyzed by several scholars.[1] Here let us first mention

[1] See Aftalion (1959); Ackerknecht (1973); Arora, Landau, and Rosenberg (1998); Bovet (1988); Chandler (1990); Freeman (1982); Gambardella (1995);

a few major characteristics of innovation processes, competition and industrial structures. The origin of the pharmaceutical industry dates back to the late nineteenth century, and is one of the earliest examples of commercial exploitation in an organized manner of scientific research and discovery, beginning with the emergence of synthetic dye production and the discovery of the therapeutic effects of dyestuff components and other organic chemicals.

Before the 1930s pharmaceutical innovation was almost completely dependent upon a few large, diversified and vertically integrated German and Swiss firms, such as Hoechst, Bayer, Ciba, and Sandoz. These firms entered the industry in the late nineteenth century and manufactured drugs based on synthetic dyes, being able to leverage their scientific and technical competencies in organic chemistry, chemical synthesis, and medicinal chemistry.

The early emergence of a restricted group of firms with large-scale in-house research and development (R&D) capabilities was a consequence of the nature of pharmaceutical R&D in the chemical synthesis paradigm, and, in particular, of the appearance of a dominant "routinized regime of search" – paraphrasing Nelson and Winter (1982) – based on extensive exploration of chemical compounds and on incremental structural modifications of *drug prototypes*, organized around highly structured processes for mass screening programs (see Schwartzman, 1976). Throughout the evolution of the industry, "the organizational capabilities developed to manage the process of drug development and delivery – competencies in the management of large-scale clinical trials, the process of gaining regulatory approval, and marketing and distribution – have also acted as powerful barriers to entry into the industry" (Henderson, Orsenigo, and Pisano, 1999).

World War II induced a massive jump in research and production efforts, sponsored by the US government (especially in the field of antibiotics), which fostered the accumulation of vast search capabilities in US firms and their entry into the international oligopolistic core.

Following World War II and the commercialization of penicillin, pharmaceutical companies embarked on a period of massive investment in R&D and built large-scale internal R&D capabilities (Henderson, Orsenigo, and Pisano, 1999). Also benefiting from the dramatic

Henderson, Orsenigo, and Pisano (1999); Orsenigo (1989); and Pammolli (1996). This overview draws upon Bottazzi et al. (2001).

increase of public support for biomedical research and health care expenditure in the post-war period, the international pharmaceutical industry experienced a significant wave of discovery, with the introduction of a few hundred new chemical entities in the 1950s and 1960s, from hydrocortisone and several other corticoids, to thiazide diuretic drugs, to major and minor tranquilizers, to the initial birth control products (Grabowski and Vernon, 2000).

Throughout its evolution the industry has been characterized by a significant heterogeneity in terms of firms' strategic orientations and innovative capabilities. Competition, in the top segment of the industry, has always centered around new product introductions, often undertaken by the oligopolistic core of the industry, subject both to incremental advances over time and to imitation and generic competition after patent expiration (allowing a large "fringe" of firms to thrive). In a good approximation, the "oligopolistic core" of the industry has been composed of the early innovative entrants, joined after World War II by a few American and British firms. At the same time, until the mid-1970s a relatively small number of new firms entered the industry, and even fewer its "core."

However, things have begun to change since then, with a major transition in the technological paradigm underlying search activities, from one based on pragmatic knowledge and quasi-random screening to one of "guided discovery" (or "discovery by design"). This has been linked with major advances in computational techniques and the biological sciences, including molecular and cell biology, biochemistry, protein and peptide chemistry and physiology.[2]

All in all, the "molecular biology" revolution has had (and is having) major consequences on the patterns of division of "innovative labor" (Arora and Gambardella, 1994; Gambardella, 1995), fostering the emergence of specialized suppliers and pure *search firms*; moreover, the dramatic increase of plausibly explorable disease targets offered novel opportunities of entry for a few new small firms into new product markets.

More precisely, Gambardella (1995) identifies two subsequent technological paradigms in the pharmaceutical industry. In the first one, dominant before about 1980, the search for innovation was carried out

[2] See Sneader (1996) and Orsenigo, Pammolli, and Riccaboni (2001).

principally through "random screening," driven by relatively tacit search heuristics and involving a great deal of serendipity (e.g. the search for one therapy leading to the unexpected discovery of a different one). Conversely, the paradigm increasingly dominant after 1980, resting on major theoretical advances in molecular biology and, jointly, in biotechnologies and computational chemistry, has made the search process more "guided," also entailing a higher degree of path dependency in the search process (Gambardella, 1995). In a similar vein, Orsenigo, Pammolli, and Riccaboni (2001) argue that the pharmaceutical industry went through a sort of *transitional regime* from the late 1970s until the early 1990s, and has been going through a new one ever since. The former began to define new "biological hypotheses" and new molecular targets, while search relied on heuristic/search methods and co-specialized technologies that tended to be specific to given fields of application. Conversely, the contemporary regime – they argue – is characterized by the emergence of new generic tools (transversal technologies), such as combinatorial chemistry.[3]

Finally, at the institutional level, the "guided search" paradigm has come together (causality is a trickier matter) with major changes in the legal conditions for the appropriation of new knowledge, including prominently in the United States the Bayh–Dole Act (1980), which allowed universities and small businesses to patent discoveries emanating from research sponsored by the National Institutes of Health (NIH), and then to grant exclusive licenses to drug companies. This institutional change led to two phenomena that are fundamental to understanding the recent changes in the structure of the pharmaceutical industry: first, a boom in biotech startups; and, second, a new division of labor between small and large firms. In the new division of labor, dedicated biotechnology firms (DBFs) and publicly funded labs (NIH and universities) typically concentrate on upstream research, while

[3] Moreover, Padua and Orsenigo (2003) claim that, more recently, the division of labor in the pharmaceutical industry has been evolving towards a hierarchical structure, whereby large firms "decompose the search space" by defining general hypotheses of work whereas small firms work on the sub-hypotheses. They also claim that in recent years biotech has been developing an intense net of collaborations between large and small firms whereby, more so than in previous regimes, the pattern of innovative activity follows a process of "creative accumulation," in which the overall coordination takes place at the level of the large firms.

"big pharma" buy from them initial drug compounds and concentrate on bringing the drugs to market through more costly clinical trials and marketing campaigns.

There are several themes running through the interpretation of the history of the pharmaceutical industry, sketched above.

1.2 Mechanisms of knowledge generation

We have made only brief mention of the profound changes that the industry has undergone recently, amounting to what we have called a paradigm change. By that we mean changes in the knowledge bases, know-how, search procedures, and characteristics of physical equipment (Dosi, 1982, 1988; Nelson and Winter, 1982), in turn implying significant discontinuities in the ways that knowledge is generated and economically exploited.

The chapters by Nightingale and Mahdi, and by Mariani, address different properties of such paradigmatic discontinuities. The former (chapter 3) analyzes the extent to which new biotechnologies have indeed affected the search process for new drugs and, in particular, how far the industry has been able to "industrialize" search and experimentation. The answer they offer is that the industry did undergo a series of incremental changes, which also increased economies of scale and scope in R&D, but this fell short of the "revolution" in drug discovery and development heralded by many industrialists and policymakers. In line with the similar argument in Nightingale and Martin (2004), the interpretation is rather "incrementalist," whereby multiple technological bottlenecks make changes in search patterns slow and uneven. The authors also interpret in this light the division of labor between "big pharma" and biotech firms, involving the painstaking absorption by the former of any array of process (and product) technologies.

Mariani's chapter 4 tackles the issue of knowledge accumulation from a complementary perspective, finding that there might indeed be solid evidence for a discontinuity between a "big pharma/chemical mode" and a "biotech mode" of knowledge accumulation. In the former, her evidence shows that innovative capabilities are firm-specific and cumulative within firms, while relatively shielded from the characteristics of the context in which they operate. This appears to apply notwithstanding the fact, noted by Galambos in his comments

(chapter 5), that traditional "big pharma" also continues to be heavily concentrated in a few locations, at least in the case of the United States – a phenomenon that dates back to the beginning of the industry. Conversely, under the "biotech mode," the statistical evidence suggests that the innovative capabilities of individual firms are fueled by the knowledge spillovers from other firms located in the same areas.

1.3 Organizational capabilities, "creative destruction," and "creative accumulation"

How do the processes of search and knowledge accumulation relate to the behaviors and life profiles of individual firms? This represents a crucial link between those investigations addressing the general *characteristics of search and problem-solving procedures*, on the one hand, and those focusing on the patterns of *corporate learning and growth*, on the other (Geroski and Mazzucato, 2002).

This is an absolutely central concern of *knowledge-based* (i.e. *capability-based* or, largely overlapping, *resource-based*) theories of the firm (see Teece, Pisano, and Shuen, 1997; Teece et al., 1994; Montgomery, 1995; Dosi, Nelson, and Winter, 2000; and Dosi and Marengo, 2000). The bottom line of this perspective is that the proximate boundaries of the firms are shaped by the knowledge that they embody and roughly map into the "decompositions" of the problem-solving space over which they operate. What one firm does along the whole chain from raw material to final products is "cut" along the boundaries of the *sub-problems* that a particular firm happens to address. So, for example, some firms might embody specific capabilities concerning the exploration of either gene properties or gene manipulation (in our case, *in primis*, biotech firms), while others might have accumulated abilities in the search for therapeutical benefits (and side effects), together with complementary assets in production and marketing (i.e. typically "big pharma").

The evidence mentioned above does indeed suggest patterns of *complementary knowledge accumulation*. Given that, what is the role, if any, for the "creative destruction" emphasized early in his career by Joseph Schumpeter as the driving force of capitalist innovation?

Lichtenberg (chapter 2) suggests that such a process takes place primarily at the level of *individual products*, and such evidence is corroborated by Bottazzi, Pammolli, and Secchi (chapter 7). More

generally, the history of the pharmaceutical industry can be read – as emphasized in Galambos' comments (chapter 5) – by a multi-level evolutionary process driven by scientific advances and discoveries of new chemical entities (undertaken to a large extent by public, non-profit institutions) coupled with the transformation of the former into new, hopefully effective and safe drugs (most often undertaken by the pharmaceutical companies).

The ensuing market evolution is, in fact, made up of two dynamics, operating on different timescales (see Lichtenberg, chapter 2 of this volume, and Bottazzi et al., 2001). The first one concerns the opening of new markets addressing a new pathology or an already "cured" pathology through different biochemical routes. The second evolutionary process regards competition *stricto sensu* amongst firms within each market on the grounds of similar chemical compounds.

Clearly, the growth of firms depends on both dynamics, with the timing of entry and the types of product introduced being important factors in shaping corporate competitive profiles. Let us turn to these issues.

1.4 Innovation, competition, and firm growth

The pharmaceutical industry offers a privileged point of observation in order to tackle a long-lasting but crucial question in industrial economics: *what determines the observed rates of innovation across firms?*

The evidence from the industry does indeed add to the evidence against the old conjecture, according to which innovation ought to be driven by firm size and by the intensity of market competition. In fact, most empirical studies show that the intensity of R&D spending is not influenced in any statistically significant way by the size of the firm, while at the same time both *ex ante* and *ex post* proxies for "market power" explain very little of the inter-industry differences in the propensity to innovate (see Cohen, Levin, and Mowery, 1987; Cohen and Levin, 1989; and Geroski, 1994).

Rather, there is mounting evidence suggesting that a good deal of the inter-industry differences in the propensity to search and to innovate relate to the characteristics of the underlying technological knowledge – including the opportunities it entails and the mechanisms of appropriability it offers (Dosi, 1988; Levin, Cohen, and Mowery, 1985; and

Klevorick et al., 1995) – which in turn often vary along the life cycle of each industry (Klepper, 1997). In fact, in this emerging tradition, well in tune with the seminal analysis by Pavitt (1984), one has begun to "map" intersectoral differences in the propensity to innovate and the modes through which innovation is pursued into underlying differences in the sources of innovative opportunities themselves (e.g. direct scientific advances, versus equipment-embodied technical advances, versus learning by interacting with suppliers and customers, etc.).

In this respect, the pharmaceutical industry conforms well with such a knowledge-driven view of the preparedness to undertake innovative activities. The steady emergence of new opportunities fueled both by science and by the development of new instruments and techniques supports a quite high *average* R&D propensity and a persistent flow of innovations (with some qualifications, as discussed below).

What about *inter-firm* differences in the revealed rates of innovation? Indeed, as some of our contributors show in another work (Bottazzi et al., 2002), in the international pharmaceutical industry one observes the *persistent* coexistence of two basic types of firms, mapping into distinct technological competencies and competitive strategies. The first group, closely corresponding to the *core*, undertakes what is sometimes called "pioneering R&D" (Grabowski and Vernon, 1987); generates the overwhelming majority of new chemical entities (NCEs); when successful enjoys big, albeit not very long-lasting, first-mover advantages; and charges premium prices. The second group undertakes primarily imitative R&D; generates incremental innovations and more competitively priced "me too" drugs; takes up licenses from the core; and is present to different degrees in the generic markets, after patent expirations.

Ultimately, what one sees is a *long-term ecology* of the industry relying on competition, but also the complementarity, between two organizational populations, the relative sizes of which are shaped by diverse competencies in accessing innovative opportunities (and, to some extent, also by intellectual property right (IPR) regimes, influencing the span and length of legal protection for temporary monopolies on innovation).

In fact, *heterogeneity* in the abilities to innovate across firms, *even within the same lines of business*, is an increasingly accepted "stylized fact." And the pharmaceutical industry stands as a sound corroboration of a more general phenomenon. In turn, such persistent

heterogeneity is circumstantial but robust evidence for *idiosyncratic differences* in technological capabilities across firms that persistently hold over time (see Teece, Pisano, and Shuen, 1997, and Nelson, 1991, and – on the empirical side – Geroski, van Reenen, and Walters, 1997, and Cefis, 2003).

1.5 Firm growth patterns

Granted all that, what are the effects of different corporate innovative capabilities upon the growth profiles of individual firms? In order to start answering the question, one should look at the statistical properties of firm growth profiles themselves. Here, a useful, even if biased, yardstick is the so-called "Gibrat law" (see Ijiri and Simon, 1977) of growth. In shorthand, the "law" in its *weak* version simply states that growth rates are uncorrelated with initial sizes (no matter where the "initial" time is set). Or, putting it another way, there are no systematic scale advantages or disadvantages in the growth process. A *strong* version of the same conjecture fully endorses the former proposition, and adds further that corporate growth is generally driven by small, uncorrelated, independently distributed random events.

For the purpose of the investigation of the growth properties of the pharmaceutical industry, one ought primarily to investigate whether the *strong* version of the "law" holds. After all, concerning the *weak* version, the overwhelming evidence does not bend either towards an unlimited tendency to monopoly or, conversely, towards some mythical "optimal scale." So, even when one finds – as one often does – a unit root growth process, a serious issue remains concerning the properties of growth rate distributions and their temporal profiles.

Suggestive evidence shows that more volatile, largely intertemporally uncorrelated growth rates characterize in particular the early phases of the life cycles of (micro-) industries and markets (Geroski and Mazzucato, 2002; Mazzucato, 2002, 2003). However, in the drugs industry each firm holds portfolios of products that happen to be at different stages of their "life cycle." Hence one should ideally distinguish between the properties of growth at the level of single products/markets, on the one hand, and growth patterns of the firm as a whole. In fact, Bottazzi et al. (2001), as well as Bottazzi, Pammolli and Secchi (chapter 7 in this volume), show that growth dynamics in the world pharmaceutical industry display a significant autocorrelation structure

both at the level of single therapeutical markets and at the level of the firm. Moreover, one observes the rather ubiquitous property of growth rates to be exponentially distributed. As discussed at greater length in Bottazzi et al. (2001), an implication of such evidence is that one can hardly attribute the growth dynamics to the addition of small, uncorrelated events. Rather, the frequent appearance of "big" events implied by an exponential distribution is well in tune with the notions that, first, discrete (possibly innovation-driven) events are at the root of the growth process, and, second, the very nature of the competition process is likely to reinforce the correlation structure underlying "big" positive and negative changes in the market share of the various firms. (The simple intuition here is that the very competitive process implies correlation in firm growth: a big rise in market advantage for one player necessarily implies a corresponding fall for the competitors.)

In this book, the contribution by Cefis, Ciccarelli, and Orsenigo (chapter 6) adds other important pieces of evidence regarding the structure of corporate growth, highlighting, again, *idiosyncratic* firm-specific (but *not* size-specific) growth profiles whereby (i) differences in growth rates do not seem to disappear over time, (ii) the notion of mean reversion finds only weak corroboration (i.e. firms with a higher initial size do not necessarily grow more slowly than smaller firms), and (iii) estimated steady states of firm sizes and growth rates do not converge to a common limit distribution (i.e. heterogeneity is a long-run phenomenon).

Chapter 7, by Bottazzi, Pammolli, and Secchi, also addresses the relationship between the variance in growth rates and firm sizes, identifying a negative correlation quite similar to that often found by other scholars in connection with different industries. What accounts for such a relation? The authors put forward an interpretation in terms of diversification patterns. In fact, the evidence suggests that there is a log-linear relation between firm size and the number of "micro-markets" in which a firm operates. Such a relation, the authors show, is fully able to account for the observed scaling of growth variance versus size. In turn, as discussed in Bottazzi et al. (2001) and Bottazzi (2001), the observed properties of diversification profiles can be well explained by a *branching process*. Such a dynamic finds intuitive roots in the underlying processes of *capability accumulation*, whereby knowledge is augmented incrementally and put to use in interrelated