Modelling industrial dynamics with “History-friendly” simulations

Christian Garavaglia\textsuperscript{a, b,*}

\textsuperscript{a} University of Milano-Bicocca, Faculty of Statistics, via Bicocca degli Arcimboldi 8, U7, 20126 Milan, Italy
\textsuperscript{b} KITEs, Bocconi University, via Sarfatti 25, 20136 Milan, Italy

\textbf{Article info}

\textbf{Article history:}
Received May 2007
Received in revised form May 2010
Accepted July 2010
Available online 8 August 2010

\textbf{JEL classification:}
B41
C63
O39
B52

\textbf{Keywords:}
Simulation
Industrial dynamics
Evolutionary economics
“History-Friendly” models
Complexity

\textbf{Abstract}

The use of simulation techniques has increased greatly in recent years. In economics the industrial dynamics approach makes use of simulation techniques to understand the complexity of the industrial process of continuous change. Among these models, a new branch of studies known as “History-friendly” models aims at establishing a close link between formal theory, developing stand-alone theoretical simulation models, and empirical evidence. In this paper, we study “History-friendly” analyses and counterfactuals. Some examples of “History-friendly” models are widely examined. Finally, the paper makes a critical contribution to “History-friendly” methodology and defines the role of “History-friendly” models in the debate on the empirical validation of simulations.

© 2010 Elsevier B.V. All rights reserved.

\section{Introduction}

The need to take into account the continuous processes of change of modern economies has been an important stimulus for evolutionary economics theories and for the industrial dynamics approach to prosper. In particular, a lack of realism in orthodox theories and a growing recognition of the importance of dynamics in modern societies have caused a shift in the interest of many economists to a different approach. Following publication of the seminal work “An Evolutionary Theory of Economic Change” in the early 1980s \cite{Nelson}, many hundred papers have been published, a lot of new specific questions have been raised, and new journals have appeared, specifically focusing on industrial dynamics and evolutionary economics.

Three levels of analysis of industrial dynamics can be found: first, the specific dimensions of industry dynamics (so-called “industrial demography”); second, the dynamics of the main structural variables, jointly considered, that characterise an industry (“structural dynamics”); and third, so-called “structural evolution”, which takes into account a broader analytical view: going beyond the first two levels of analysis, the research agenda focuses on the evolution of the industry as a whole, the emergence of new products and technologies, of new skills and firm competencies, the definition and related changes in firm boundaries, learning processes, diversification and integration strategies, the emergence of networks and the role of public authorities and institutions \cite{Malerba}. Such an attempt brings an obvious higher degree of complex-
ity into the study. Simulation techniques appear to be the best option for the economist who aims at such a complex analysis.

Since the 1970s and 1980s, Nelson and Winter, and then others, have begun to define a secure ground and indicate a trajectory for simulation models (Nelson and Winter, 1977; Silverberg and Verspagen, 1995a; Silverberg, 1997; Andersen and Valente, 1999; Kwasnicki, 2003). Following this tradition, among the numerous significant developments of the evolutionary approach, a “first generation” of simulation models (Windrum, 1999, 2007; Bottazzi et al., 2001) has been increasingly implemented (see among others Silverberg et al., 1988; Chiaramonte and Dosi, 1993; Silverberg and Verspagen, 1994, 1995b; Dosi et al., 1995). More recently, a new family of simulations has tried to develop more complex models in order to better understand the characteristics and complexity of the industrial process of continuous change, and with the explicit purpose of achieving more satisfactory empirical predictions. We claim that this family of models can be considered to have a crucial role in what has been called by Bottazzi et al. (2001) and Windrum (2007) the “second generation” within the expanding literature of evolutionary thinking. With regard to this, it is possible to recognise at least two different complementary methods: on the one hand, there are models that aim at identifying potential invariances and generic basic properties of industrial structures and dynamics without focusing on specific sectoral characteristics (see among others Winter et al., 2000; Bottazzi et al., 2001; Winter et al., 2003; Windrum and Birchenall, 2005). On the other hand, there are the “History-friendly” models, based on the detailed illustration of a specific industry, which add more history-based details to a formal representation.

Simulation methodologies have represented a new way of both studying and of modelling the features of industrial economics in economic thinking. Different directions of research and approaches have been investigated and no common protocols or uniform frameworks for developing and studying simulation models have so far emerged. Many papers dealing with simulations and “Agent-Based Models” (ABM) have already been published. The new contribution of this paper is to put some order into a distinctive class of ABM, the new branch of studies known as “History-friendly” simulation models, emphasising their original aspects and their way of model developing.

This work is structured as follows: the second section examines the main features of simulations and the “History-friendly” approach. The third aims to critically review the basic characteristics of existing “History-friendly” simulations. The fourth section concentrates on some methodological aspects of “History-friendly” simulations. The fifth one focuses on the debate about empirical validation of simulation models. Finally, some conclusions and challenges are presented.

2. General features of “History-friendly” simulation models

Simulation techniques have been more and more widely used during the past years in many fields of the social sciences. In the economics literature, different approaches for simulating economic systems have emerged under different labels, usually grouped as a whole under the name “Agent-Based Models” or “Agent-Based Computational Economics”. The “History-friendly” modelling style belongs to this broader family of models (Dawid, 2006).

2.1. Complexity and simulations

The strongest criticisms of orthodox models have been concerned with the inability to fit well with the complexity of economic systems. One possible explanation for this lack of complexity might be linked to excessively simplistic assumptions needed in order to analytically manage and solve the models. Much complexity is lost in the formal modelling definition of behaviour and interactions of economic agents. The development of simulation models and the inherent formalisation of the relationships, variables and interactions among these variables, force the economist to better understand some concepts, usually disregarded by a pure analytical approach of analysis, and to analytically define a larger number of relations and equations that enable the researcher to capture a representation of the dynamics of the model that is closer to the characteristics of complex systems.

This need for realism has certainly been one of the most important causes of the emergence of evolutionary theories. In this regard, simulation models in general, if built upon a detailed analysis of economic behaviour at a micro level, represent an attempt to study the patterns and processes thought to characterise the real world’s economic environment. In this sense, the ‘demand’ which has recently emerged for a simulating approach comes directly from the need to analyse increasingly dynamic, volatile and changing economic systems.

Complexity is at the heart of modern adaptive and dynamic economic systems (Tesfatsion, 2001a,b, 2006; Pyka and Fagiolo, 2005) and simulations provide researchers with a tool to translate complex relationships, agents and interactions into economic models. In particular, complexity and consideration of the macro (or industrial) level as a result of the repeated interaction of micro heterogeneous agents are two of the core features of evolutionary theories (Marenko and Willinger, 1997). The interest of our analysis and the targets of our studies, then, are not directly and explicitly modelled in our frameworks. Rather, they emerge from the repeated computed interactions of components of the model (see among others Lane, 1995a; Gilbert, 1995; Kirman, 1997; Hodgson, 2000; Ziman, 2000; Tesfatsion, 2006; Richiardi et al., 2007).

2.2. “History-friendly” simulations

Starting from the definition and formalisation of behavioural routines and attributes of economic agents,
simulation models are able to reach an outcome in terms of time paths of the principal endogenous variables at the system level: the aggregate structure emerges from the organisation of interacting agents (Tesfatsion, 2001a,b).

Along this line, “History-friendly” models, like many other simulation models, could be defined as “a representational mechanism that is distinguished by its capacity to generate relations that are not explicitly encoded” (Rasmussen and Barrett, 1995). If we believe that the whole provides more information and results than the simple sum of its parts, because of the inherent process of change and interaction among them, simulations then give us the possibility to study some phenomena that are generated by the relationships occurring among different elements not inherent in or predictable from a knowledge of their constituent parts (Lane, 1993b; Byrne, 1997; Holland, 1998; Lomi and Larsen, 2001; Dopfer et al., 2004).

It is clear, then, that what we mean by “History-friendly” simulations is not the numerical analysis of complex formal models, whose analytical solution is difficult to derive. On the contrary, “History-friendly” models are stand-alone simulation models (Axtell, 2000), precisely specified and formally structured, which build upon the definition of ‘interacting objects’. Objects represent agents as well as environments of the specified economic system. Objects are defined by a given set of data and variables – i.e. their characteristics – as well as methods and algorithms – i.e. behaviour. In order to do this, we make use of an ‘object-oriented’ language for programming the model.

Appreciative theories bring to light how economies are characterised by non-linearity, stochastic dynamics, heterogeneity, uncertainty, interaction, bounded rationality, path-dependency and coevolution. In addition, most of these aspects seem to be highly industry-specific. A deep understanding of these phenomena, then, requires a close and particular representation of the characteristics of the specific sector and environment under consideration in order to be able to capture the mechanism and processes that govern its evolution. The structure of “History-friendly” models aims at representing the routines, relationships and behaviour of economic agents as indicated by qualitative theories regarding the mechanisms underlying the evolution of given industries. Accordingly, the declared purpose of “History-friendly” models is to “capture the gist of the appreciative theory put forth by analysts of the history of an industry or a technology, and thus enable its logical exploration” (Malerba et al., 1999). Thus, “History-friendly” models want to encourage a dialogue between formal theories and industry studies.

3. Critical discussion of “History-friendly” models in the literature

Various studies on “History-friendly” simulation models have been published in the literature in recent years. In this section we sketch an outline of these models. We believe that the examples we present elucidate, on the one hand, the various opportunities of modelling complexity with a “History-friendly” approach and, on the other hand, the possibilities of investigating, exploring, testing and reasoning the theories. A more detailed and analytical discussion about some of the relevant building blocks of the models (demand, innovation, entry) is provided in Section 3.1.

• Evolution of the computer industry

The seminal study is that on the “History-Friendly” model of the evolution of the computer industry (Malerba et al., 1999, 2001a), discussing some major events that characterise the evolution of this industry: in short, the emergence of a dominant firm in the mainframes market (IBM), the introduction of microprocessors and the subsequent rise of the new market segment of personal computers. The key results of this first model relate to the exploration of determinants of the achievement of market dominance by one firm in mainframes that, in the presence of substantial changes in fundamental technologies, was not able to seize the major share of the new PC market that the new technology opened up. The timing of the introduction of the new technology, the higher consumer lock-in for the mainframe market compared to the PC market, and the greater sensitivity of consumers to the price of computers in the PC market are considered to be key factors in the evolution of the industry.

As an example of theoretical investigation, in a subsequent paper Malerba et al. (2001b) extended the basic model in order to debate the efficacy of economic policies in a complex environment characterised by dynamic increasing returns generated by multiple sources, that is, both on the marketing side, given by a strong lock-in effect of customers, and on the production and technological side, given by the “success-breeds-success” story. The analysis is conducted in a counterfactual style (see the more detailed discussion in Section 4.2). With regard to this it becomes clear how the run of these exercises represents an investigative device, so that we are able to assess the question “What might have happened…”, as we mentioned above, “…if the Antitrust Authority had intervened early or late?”, with the aim of giving insights and to contributing to the discussion about the role and effect of antitrust laws.

This version of the model is particularly interesting because it explicitly considers that firms are embedded in a system where other agents such as public and other private institutions are involved. This consideration adds complexity to analysis of the industry, and our simulation runs represent a way of coping with it.

• Evolution of the pharmaceutical industry

In another model (Malerba and Orsenigo, 2002) the focus of analysis rests on the pharmaceutical industry. The purpose is to capture the main features of the industry evolution, examining the differences between the so-called era of random screening and the age of molecular biology. It is shown that, given the nature of the processes of drug discovery, the characteristics of the competition process, the low degree of cumulativeness and the fragmented nature of the markets, a low level of overall concentration of the industry soon emerges, and the advent of biotechnological firms does not represent a displacement of the incumbent ones.
Again, as an example of theoretical discussion, in an investigative exercise of the simulation model (Garavaglia et al., 2006), the entry process of pharmaceutical firms (ignoring the advent of biotechnology) was examined. Considering the low cumulativeness in the process of drug discovery and a low degree of concentration, the study looks at why entry was not a relevant feature of the pharmaceutical industry. First results show that timing is relevant in determining the effective entry rate of potential entrants: the larger the time span among the timing of different cohorts of entrants, the smaller the probability of successful entry and the achieved market share for the latest cohorts. Second, the study investigates the factors believed to be relevant in affecting or obstructing entry. The simulation results show that higher discovery opportunities make it easier for later potential entry cohorts to prosper, while increasing costs of looking for new compounds hamper the entry rate of later cohorts. The size of the market is also relevant; larger markets may imply a reinforcement of the “success-breeds-success” effect (Nelson and Winter, 1982): early entrants gain larger profits, and, therefore, have higher probabilities of developing new products in the future, at the expense of later entry cohorts. No significant effects in the longer patent protection and higher cumulativeness in the search process are found.

Interpretation of the results shows that the absence of significant entry in the industrial evolution of pharmaceuticals has been mainly attributable to: strong first mover advantages of early entrants (the “success-breeds-success” story);\(^2\) the increasing cost of searching for new compounds (that characterised the pharmaceutical industry for decades, as many scholars have argued – see for example Pammolli, 1996); decreasing market opportunities (until new opportunities opened up with the advent of biotechnology that, in fact, led to a significant wave of new firms entering the market).

**• The German synthetic dye industry**

Another interesting model about the pharmaceutical industry is reported in an already-mentioned paper by Brenner and Murmann (2003). Relating to Murmann’s (2003) empirical research on the dominance of the German synthetic dye industry in international competition in the first decades of the twentieth century, the paper explores the role of the scientific capabilities of a large number of organic chemists, and the effect of the university system in determining the success of German firms. The results of the simulations support the thesis that the responsiveness of the German university system to the industrial demand for chemists provided sufficient human capital to firms that have consequently been able to dominate the global market. Moreover, the paper claims that this responsiveness shows decreasing returns; this, in turn, implies that investing in the education of human capital has decreasing returns. The interesting normative message of these results refers to the possibilities for countries to respond effectively in the development stage of an industry in order to support its growth.

**• The random access memory chip industry**

Evolution of the dynamics random access memory chip industry is analysed in a “History-friendly” model by Kim and Lee (2003). The paper aims at explaining the impact of different technological regimes on the entry of firms, on organisational selection among different types of firms and on the evolution of market structure. The early history of the industry at the beginning of the 1970s was characterised by the dominance of small specialised firms. By the mid-1980s, these firms were displaced by large diversified firms. The model takes into account two populations of heterogeneous firms and explains the dominance of the large diversified firms in the later stage with regard to the features of the technological regime in the market. Two aspects, in particular, are considered: the degree of cumulativeness and the impact of process innovation on productivity increase. Simulation results show that the process of entry is easier, ceteris paribus, the lower the degree of cumulativeness and the larger the parameter that indicates the degree of productivity increase related to process innovation. Cumulativeness impacts negatively on entry because it implies that productivity is influenced more by the pre-existing level of productivity, favouring incumbent firms. The productivity increase parameter, on the other hand, determines the pace at which the gap between the average incumbent firm and potential entrants shrinks: the higher it is, the easier the entry process. Historical analysis of this industry shows that: the degree of cumulativeness turned out to be low because the productivity of a given generation of chips is not dependent on cumulative experience with the previous generation chips; the impact of product innovation on productivity is found to be increasing, as the measure of memory capacity across generations of chips shows. These aspects explain why entry has not been disadvantageous in the later stage of the industrial evolution. Moreover, the model explains in the same way why the dominant firms in the industry are the large diversified ones: low cumulativeness, speed of productivity catch-up and larger investments in R&D (and hence process innovation) associated with the larger size of firms worked in favour of entry of large firms (that in the end have been able to dominate the market).

**• Coevolution of semiconductor and computer industries**

Malerba et al. (2006, 2008a,b) explore the coevolution of two vertically related industries in a simulation model about semiconductor and computer firms. The research illustrates the dynamics of the two industries from the
1950s through the 1980s in the American market. The dynamics of the computer industry shows the introduction of mainframes, minicomputers and personal computers during these decades. In the semiconductor industry, three different technologies followed one after another: transistors, integrated circuits and microprocessors. The core claim of the research is that the processes of vertical integration and specialisation of firms over the years are the result of the coevolution of upstream and downstream industries. In the early days of the industry, computer firms that produced mainframes were vertically disintegrated. After a few years, some big firms started to integrate and by the end of the 1950s, IBM came to dominate the market. IBM was vertically integrated and after the introduction of integrated circuits in the semiconductor industry it reinforced its integration strategy. When microprocessors appeared IBM disintegrated. Moreover, the new computer firms producing personal computers were also all specialised.

While the process of vertical integration of firms has been widely discussed in the economic literature with the so-called ‘transaction approach’ and with contract theories, the research claims the need for a new theoretical explanation that might be able to grasp the dynamic aspects of this process. After some simulations that match the evolution of the industries under examination, the model explores different conditions and settings, and shows that three factors influence the evolution of market concentration and vertical structure of firms: the size of external markets for semiconductor producers; the magnitude of technological change; and the lock-in effects in demand. The basic intuition behind the model is that computer firms are able to accumulate competencies over time in the production of both systems and components, which in turn may be brought into the marketplace by semiconductor producers. To put it very briefly, only in the years of microprocessor technology were computer firms almost totally disintegrated. This is explained by the fact that the level of semiconductor quality depends basically on the relevant technology (the microprocessor represented a significant technological advancement), the size of external demand (the microprocessor also had large external applications in telecommunications, consumer electronics and the automobile; this means that the bigger the external demand, the larger the investment in R&D, and the higher the quality of semiconductors) and the size of the producers (only with microprocessor technology did a large firm, i.e. Intel, emerge and was consequently able to invest heavily in R&D and improve the quality of semiconductors). Given the high quality and uncertainty of microprocessors in semiconductor technology, computer firms producing both mainframes and personal computers decided not to integrate.

3.1. Building blocks of “History-friendly” models: analytical framework of demand, innovation and entry

In this section we provide a detailed and analytical discussion about some of the relevant building blocks of the existing above-mentioned “History-friendly” models. This section is self-contained: those readers who are not interested in analytical analyses may skip this part.

3.1.1. Demand

The demand of individual firms plays a key role in all simulation models because in the end it determines the firm’s market share and consequently the evolution of market structure. In some models, demand is more complex while it is less in some others. In Kim and Lee (2003), total consumer demand is exogenous: it is measured in terms of total revenue for the product, it is fixed to $R$ and it contributes to determine the market price as follows:

$$P_t = \frac{R}{Q_t^{\max}}$$

if $R/Q_t < P_{\max}$, where $Q_t$ is the total output of the industry and $P_{\max}$ is a given upper bound value of the price, and otherwise it is equal to $P_{\max}$.

Thus, demand is completely determined by the production decisions of firms. The output of the $i$-th firm at time $t$, $Q_{it}$, depends on capital productivity, $h_{it}$, and capital stock, $K_{it}$, of the firm. In this model, the activities of innovation and imitation are the key determinants of productivity, $h_{it}$, that in turn is the crucial variable in establishing a firm’s individual demand.

Also in Malerba and Orsenigo (2002), demand is exogenous. Each submarket (a therapeutic area, $TA$, in the model) has a given value that resembles the dimension in terms of the patients affected by a disease and the importance of that disease. This value is fixed at the beginning of the simulations. However, the decision to buy a product is well detailed: each company’s product gives to the consumers a utility. When a firm $i$ launches a new drug $j$ onto a given market, the utility of its product, $(U_{jt})$, is a function of the quality of the drug, $PQ_{jt}$, a measure of cheapness, $1/\text{markup}$, the advertising expenditures at the drug launch at time $t$, $A_{jt}$, and the annual marketing expenditures of the firm, $YA_{jt}$:

$$U_{jt} = PQ_{jt}^a \cdot \left(\frac{1}{\text{markup}}\right)^b \cdot A_{jt}^c \cdot YA_{jt}^d$$

where $a$, $b$, $c$ and $d$ are parameters. Consequently the share of this market value that a firm $i$ gains for the product $j$ is:

$$MS_{ij} = U_{ij} / \sum U_{TA},$$

where $U_{TA}$ is the utility of all firms’ products in this market. To conclude, the quality of the product – i.e. the results of a firm’s innovation activity – and advertising strategies are the key determinants of the firm’s relative demand in this model.

In the model about the evolution of the computer industry (Malerba et al., 1999, 2001a), demand is more complex: there is heterogeneity among different groups of customers and demand is a function of the quality of the products.

---

3 In more recent development of the model, demand, like other parts of the framework, has been more composite (presentation by Orsenigo at the Conference on the Global Health Economy “The Global Organisation of Biomedical Innovation: Funding, Intellectual Property Rights, Incentives and the Diffusion of New Technology” in Kiel, 4-8 October 2007; presentation by Garavaglia at the FIRB - RISc Conference (“Research and entrepreneurship in the knowledge-based economy”) in Milan, 7-8 September 2009.
according to a utility function, and of the number of producers. More precisely, there are two groups of potential buyers with different preferences for the characteristics of the computers: cheapness, $X_1$, and performance, $X_2$. One group values the characteristic of performance more (this group mimics the behaviour of large firms), while the other places more value on low price (i.e., individuals). The utility, $M$, associated with a computer is given by a Cobb–Douglas function:

$$M = b_0 (X_1 - X_{1\text{min}})^{b_1} (X_2 - X_{2\text{min}})^{b_2}$$

where $X_{1\text{min}}$ and $X_{2\text{min}}$ are the minimum requirements of the two relevant characteristics a computer has to meet in order to be sold on the market. Thus, the quality of a computer (i.e., utility for the customer) is an increasing function of its performance and cheapness. Demand is organised into submarkets, each characterised by independent buying behaviour. In each period of the simulation run, a fraction of the submarkets wants to buy a new computer because they have never bought one or because of computer breakdown. Moreover, the probability that submarkets buy computers is related to the existing supply, i.e. the existing number of firms $n$ selling in the marketplace: if $n \geq \bar{n}$, then submarkets purchase with a probability equal to 1; if $n < \bar{n}$, then the probability decreases proportionally. The final decision to buy a particular computer from the $i$-th producer is a function of its quality $M_i$, the market share of the firm $m_i$, the level of advertising expenditure of the firm $A_i$:

$$P_i = c_0 (M_i)^{c_1} (m_i + d_1)^{c_2} (A_i + d_2)^{c_3}$$

where $c_0, c_1, c_2, c_3, d_1$ and $d_2$ are parameters.\(^4\)

In the evolution of the computer and semiconductor industries (Malerba et al., 2006, 2008a,b) the demand for computers is similar to that discussed above. Each firm’s product is characterised by a level of utility (merit), $M$, that is a function of the position of the firm in the space of characteristics cheapness-performance. The higher the utility $M$, the higher is the probability of the computer being sold. Also, the higher the firm’s market share, the higher is the probability of the computer being sold: there is a brand-loyalty effect in consumer behaviour. The demand for components, faced by component firms, comes from two sources: an exogenous external demand and the demand from specialised computer firms that do not produce their own components internally. Again, the higher the utility $M$ of the component and the market share of the component firm, the higher the probability the component will be selected by a computer firm.

In the synthetic dye industry (Brenner and Murmann, 2003), each $i$-th product is characterised by two values, $k_i$ (the typology of product in the characteristic space) and $a_i$ (the level of technological performance of the product). Each consumer $j$ in this market is also characterised by two values: $k_j$ defines the favourite typology of product in the space and $a_j$ is the minimal level of performance that induces the consumer to switch from buying natural to synthetic dyes. At time $t$, then, some consumers will buy natural dyes while some others buy synthetic dyes, until such time as technological improvement of synthetic dyes replaces natural dyes (at the beginning of the simulations, all consumers buy natural dyes). The number of consumers is assumed to increase linearly. The utility of the product $i$ is given by:

$$a_i - \phi |k_i - k_j|$$

where $\phi$ indicates how likely consumers are to buy a product that differs from their favourite one, resembling in this way a demand à la Hotelling: the higher the difference between the preferred type and the given type of product, the less valued the product is. Given a linearly decreasing demand function, this value determines the potential demand for the producing firm, $d_{pot}$. There is inertia in demand such that it takes time for consumers to change the supplying firm even if there are other better products in the market; accordingly, the sales of firm $i$ at time $t$ are:

$$s_t = (1 - \sigma)s_{t-1} + \sigma d_{pot}$$

where $\sigma$ is a parameter that determines the speed of consumer adaptation.

To sum up, the definition and computation of firm demand is a key building block in all the models we examined because it determines firms’ relative market share and, ultimately, emerging market structure. In all models, the key determinant of demand is investment in innovation: each product is characterised by a level of utility (i.e., merit, performance), or in the Kim and Lee model, each firm is characterised by a level of productivity that is defined by a process of innovation and/or imitation. The behavioural rules of companies in innovative and imitative activities thus represent the primary determinant of the evolution of simulations. Another frequent relevant factor seems to be a sort of lock-in, or brand-loyalty, effect that affects consumers’ behaviour, and that gives advantages to established large firms over newcomers and small producers.

3.1.2. Innovation rules

In this section, when appropriate, we distinguish two levels of analysis: innovation at the technological level, i.e. the dynamics of the technological frontier, and innovation at the firm level, i.e. innovative and imitative strategies of firms.

In all models by Malerba et al. (1999, 2001a, 2006, 2008a,b) and Malerba and Orsenigo (2002) there is one, or more, exogenous shock at the technological level that brings a new technology into existence. In the evolution of the computer industry (Malerba et al., 1999, 2001a) after a given number of periods, firms may decide to switch from using transistor technology to microprocessor technology that has become exogenously available. In the pharmaceutical industry (Malerba and Orsenigo, 2002), a “biotechnological revolution” opens up the opportunity of new techniques for researching and discovering new products. In the semiconductor industry (Malerba et al., 2006, 2008a,b), there are two technological discontinuities...
that represent the transition from transistor to integrated circuits and from integrated circuits to microprocessor technology, at a time when a new cohort of firms enters the market of semiconductor producers.

These exogenous shocks represent an important building block in these models in the sense that they represent new opportunities in the market, affecting entry of new firms, opening up of new markets, innovation strategies, selection among firms, and ultimately driving the technological evolution of the industry.

More interesting is the analysis of innovative and imitative behavioural rules of firms. On the one hand, all models present a similarity: uncertainty is associated with R&D (innovation) activity, such that typically R&D activities are presented in the form of probability functions. On the other hand, there is heterogeneity among models: each model has its own specificity and routines that we analyse below.

In Malerba et al. (1999, 2001a) innovation is incremental. Improvements in the relevant characteristics cheapness, $X_1$, and performance, $X_2$ are given by:

$$\Delta X_i = a_0(R_j)^{a_1}(T_j)^{a_2}(L_i - X_i)^{a_3} e$$

where $i=1, 2$. Thus, innovation depends on: R&D expenditures, $R$; the experience of working with a particular technology $j$, i.e. the number of periods a firm has the technology, $T$; the distance of the achieved level of the characteristic $i$ to the frontier, $(L_i - X_i)$, thus resembling diminishing returns of progress; a random element, $e$, that recalls the uncertainty associated with innovation activity. $a_0, a_1, a_2, a_3$ are parameters.

Adoption of the new technology in this model is an interesting two-step process. The advent of the microprocessor is assumed to be competence-destroying. First of all, incumbent firms need to perceive the existence of the new technology. This is modelled with a stochastic process that is a function of the position of the firm in the technological frontier in transistors, $z_i$, and the position of the best-practice firm that has already adopted microprocessors, $z_{mp}$. Firms need to invest a fixed amount of money, $F$, and a variable part that is a function of the accumulated budget, thus capturing the re-training of engineers. The competence-destroying feature is captured by the reduction of the experience $T$, accumulated on the transistor technology, for the new adopter. No imitative rules are developed in this model.

In Malerba et al. (2006, 2008a,b), a firm’s routines of innovation are more structured, although similar to those presented above. Incremental innovation is modelled with a “draw scheme” procedure (Nelson and Winter, 1982). The process is subject to uncertainty. The probability of improving the utility of the product depends on a firm’s R&D expenditures, the level of publicly available knowledge and the number of periods since a new technology has been introduced. Moreover, firms which vertically integrate in the production of both systems and components for computers benefit from coordination advantages through a spillover effect that improves R&D activity. Again, no explicit behavioural rules of imitation are present in this model.

In the model of the evolution of the pharmaceutical industry, both innovation and imitation are developed. The interesting feature of Malerba and Orsenigo’s (2002) model is that in the first era of the industry’s evolution, innovation is a purely random activity, thus resembling the research regime in the so-called age of “random screening” in pharmaceuticals. With regard to this, innovative activity is a two-stage process. In the first stage (called “search”) firms screen a number of potential products (molecules in the model) by randomly picking up some points in the product space (most of which are zero-valued, thus representing the degree of opportunity for developing a new drug). The number of trials is determined as an increasing function of the amount of money invested in “search” activities with decreasing returns. In the second stage (called “research”), after a potential product with a positive quality has been found, firms patent the potential product (that then becomes unavailable for other firms) and start to invest in “research” activity in order to rise to its unknown potential (i.e. quality) (Fig. 1).

Firms are heterogeneous according to the amount of budget invested in their “search” and “research” activities, and to the speed of the researching process. After this development process is completed, firms may commercialise their product only if its quality is superior to the minimum level required by the external agency of regulation.

In the second era of the industry’s evolution, after the biotechnological revolution has occurred, new firms enter the market with a new method of “search”. The model sets a sort of knowledge-based activity and learning process in the new routines for screening the product space. In some specific markets, these new biotech firms are able to look at an interval in the product space and detect, in an imperfect way, the quality of their potential products. The width of the interval is a firm-specific parameter that represents firm’s capabilities. The larger the width of the interval, the higher is the probability of finding more potential products, but obviously there is a higher risk of imperfection in the detection of the products’ true quality. Economies of scope and an experience effect are introduced in these routines. Incumbent pharmaceutical firms may also adopt the new mode of searching by paying a fraction of the accumulated budget for reorganising the company.

In this model firms may also imitate other firms’ products whose patent has expired.

The most interesting feature of this model is the richness of the routines that represent the research activity in general. In addition, in this model we are able to recognise both incremental and radical innovations: the former is represented as a new product that enters an already discovered submarket (TA); the latter may be associated with the very
first product that abruptly opens up a new submarket for curing a previously uncured disease.

Brenner and Murmann (2003) work out firm behaviour so that they can obtain new products through both innovation and imitation about both process and product. Successful process innovation leads to a higher performance for a given product. The probability of this innovation is a function of a basic innovation rate, an innovation rate that increases with the size of the firm and the number of researchers (chemists in their model) employed by the firm. Firms may also imitate a product that is similar (i.e. whose position in the space of characteristics is close; see Section 3.1.1 for more details) but more technologically advanced than their own: if imitation is successful, the product gets the level of performance of the imitated one distorted by a random term.

Product innovation is structured according to a satisfactory behavioural rule. Each firm has a firm-specific aspiration level of growth g_a. In each period, firms check the growth of their sales g and compare it to g_a. If g > g_a, then the firm is satisfied, whereas if g < g_a, the firm’s strategy value v(t) reaches value 1. When v(t) > 0 firms try to invent new products, whose probability of success depends on the size of the firm and number of researchers. The characteristics of the new product are determined randomly, and its level of performance is set similar to the closest existing product in the market. The process of product innovation is similar to process innovation.

Kim and Lee (2003) model innovation and imitation activities as an improvement in the firm’s capital productivity, h_{it}. The probability of successful innovation or imitation is proportional to the amount of R&D expenditures, determined as a ratio to capital stock (r_{it} for innovation and r_{it}^m for imitation). Interestingly, in this model firms pursue both innovation and imitation at the same time. When innovation is successful, productivity increases by a factor \rho (the growth rate of productivity over the period) as follows:

\[ h_{it}^n = (1 + \rho) h_{it} \]

When imitation is successful, the new level of productivity is a weighted average of the productivity of the imitated firm, h_{it}^m, and its own productivity:

\[ h_{it}^m = \rho h_{it} + (1 - \rho) h_{it}^m \]

Firm i productivity at time t+1, h_{it+1}, is the maximum among \[ h_{it}, h_{it}^n, h_{it}^m. \]

Similarly to Brenner and Murmann, the model defines a behavioural rule that leads to a change in the actual expenditures on innovation and imitation according to a satisfying criterion. When firms realise that their “performance indicator” is below the industrial average profit rate, they change their expenditures on innovation and imitation, through a stochastic change in r_{it} and r_{it}^m.

In the above-mentioned models, innovation routines represent the real engine of simulation dynamics. There is a certain degree of heterogeneity among the routines that, on the one hand, shows the richness of modelling but, on the other hand, works against comparability of the models.

3.1.3. Entry dynamics

Entry and the conditions under which entry occurs are fundamental blocks for the dynamics in the models. In this section we discuss the way entry is modelled in the simulations.

The three Malerba et al. models develop the entry process in a rather exogenous way. However, in each model we recognise some interesting features. In the model about the computer industry (Malerba et al., 1999, 2001a), at the beginning of the simulation, some firms enter the market: firms are born with different initial budgets, used to finance R&D, and distinctive technological capabilities. Only some of these firms will be able to achieve the minimum requirements (see Section 3.1.1 for more details) of quality of their product in order to commercialise it and then enter the market. A similar story applies both to the evolution of the computer firms in the model about the coevolution of computer and semiconductors (Malerba et al., 2006, 2008a,b) and to the pharmaceutical firms in Malerba and Orsenigo (2002).

In this context we would suggest it is better to interpret this process as a pool of potential entrants, some of which succeed while others, the less efficient or rich, fail. This interpretation is better suited to the logic of the model, in which only those firms that are able to commercialise a product in the marketplace are considered actual producers. Moreover, this interpretation is more endogenous: given that the conditions of the market and other firms change over time, and only those potential entrants that are able to fit to the prevailing market conditions effectively enter. Thus, market conditions and the degree of competition among firms act as endogenous forces in inducing or obstructing entry. Further improvement of the model has been made in recent developments of the entry routines: firms evaluate the degree of competitiveness and the value of the potential market to gain with the launch of a new product before deciding to enter.\(^{\text{5}}\) In fact, potential entrants are allowed to try and enter only at the beginning of the simulation (or at the beginning of an exogenous technological shock): after all potential entrants have either successfully entered the market or failed, no more entry may occur in the model. It would be more accurate to have a set of potential entrants that try and enter throughout the whole simulation time span. With this change, Garavaglia et al. (2006) partly overcome these limitations further developing Malerba and Orsenigo (2002). The model is structured so that different cohorts of potential entrants are allowed to enter the market at different times (see the discussion of Model b in Section 3 of the present paper).

There are some other interesting features of these models with respect to entry. In Malerba et al. (1999, 2001a), there is a second cohort of entrants after the introduction of the technological shock (microprocessor technology). These firms face barriers to entry in the mainframe market, given the existence of experienced firms producing transistor-based mainframes in this market that present greater value for the characteristics of their products compared to a newcomer. In addition, a lock-in effect in demand

\(^{\text{5}}\) See presentations by Orsenigo and Garavaglia as reported in Note 3.
generates an advantage for established firms. Moreover, these firms have invested in advertising, thus generating another disadvantage for new microprocessor-based firms.

Another interesting feature is related to the competence-destroying nature of exogenous technological shocks in semiconductors. Firms adopting the new technology perform better than previous firms. It would be worthwhile examining this issue more endogenously.

Brener and Murmann (2003) and Kim and Lee (2003) present more endogenous routines for modelling the entry rule. In Brener and Murmann there is a pool of researchers (chemists), only some of whom are employed in some firms whereas others remain unemployed. The probability of registering a new entrant in a given market is proportional to the number of unemployed researchers and the size of the market being entered. Before entry occurs, the simulation routine controls whether, given the characteristics of the potential product, there is any demand for that product and that no patent is infringed. In Kim and Lee, potential entrants may enter by imitating an incumbent firm. The model sets an amount of external money \( E_{in} \) to be devoted to imitation trials for potential entrants. Entry is modelled as a stochastic process that follows a Poisson distribution with an expectation that is a function of \( E_{in} \). If entry occurs, firms are endowed with a productivity that is calculated as the weighted average of the basic productivity of the industry and the productivity of the imitated incumbent firm. Given this level of productivity, the \( i \)-th potential entrant is able to calculate its expected average total costs, \( \text{ATC}_i \). If the following conditions are satisfied, then entry occurs:

\[
\text{ATC}_i < \bar{\text{ATC}}, \quad \pi_i > 0 \quad \text{and} \quad K_i > K^e_{\text{min}} \]

where \( \bar{\text{ATC}} \) are the average total costs in the industry, \( \pi_i \) are the expected profits after entry, and \( K^e_{\text{min}} \) is the minimum size of capital required to enter. New entrants are endowed with a financial premium in order to stay long enough to try to compete and survive at least some period of time after entry even if they suffer negative profits.

In this model then, entry is endogenously determined: the average productivity (through \( \bar{\text{ATC}} \)) is an endogenous variable and entry crucially depends on its value.

Table 1 summarises the discussion.

4. Analysis and counterfactual analysis in “History-friendly” models

“History-friendly” models do not aim to produce predictive results. The goal of this approach is twofold. On the one hand, the purpose is analysis and description (Cohen and Cyert, 1961). In this sense “History-friendly” models try to mimic the evolution of the industry of interest, with the purpose of describing its main dynamics and relations and understanding what the factors and fundamental processes are that make the model behave as it does.

On the other hand, the purpose is prescription. From this point of view, “History-friendly” simulations are interestingly implemented both to understand what circumstances cause a given outcome and to study counterfactual analyses.

If the former purpose aims at answering to the question “What has happened? And why?”, the latter focuses on the questions “What would happen?” and “What might have happened?”, stressing what conditions we need in order to obtain a given result and pointing out what can be made to happen.

4.1. Calibration, development and analysis

Formal evolutionary models strongly contributed to the understanding and exploration of the logic of economic evolutionary processes. Some of them paid attention to the consistency of their logical explanation and the empirical stylised facts to be explained. On the other hand, some represented a more abstract approach, and less attention was paid to the empirical phenomena they claimed to model. Despite enormous success, the “first generation” of evolutionary models revealed a rather abstract and simple structure and, when present, their empirical basis was given by very broad phenomena. The “History-friendly” approach tries to overcome these limitations by enriching the internal structure of the models and imposing a stronger empirical discipline on the formal structures (Orsenigo, 2005). “History-friendly” models provide a theoretical tool that can engage in dialogue with the logical examination and causal explanations presented by empirical studies. The core characteristic of “History-friendly” simulation models lies in the attempt to put together a deep understanding of the historical events of an industry’s evolution and a formalisation of the relationships that govern the behavioural routines. More precisely, “History-friendly” models are, on the one hand, formal models, a well-suited tool for logical explorations and causal argumentation, making explicit the logic that guides the model results. On the other hand, they follow an appreciative theorising, being empirically oriented and relating to verbal argument put forth by many evolutionary scholars.

The basic difference between “History-friendly” models and other evolutionary simulation models rests in the field and scope of analyses: while “History-friendly” models are industry-specific and history-based, other general simulation models investigate collective invariances, generic properties of industry structures and dynamics that apply to many industries, purposely neglecting historical and industry-specific features. For example, Bottazzi et al. (2001) develop a reduced form model that investigates the emergent properties of industrial structures characterised by different innovative opportunities, different degree of cumulativeness in innovation, different technological learning, and different entry conditions. The model identifies some robust interactions among the learning processes, entry, selection regimes and the competitive dynamics. Winter et al. (2000, 2003) present a model that simulates generic properties of industry dynamics related to the number, size and age of firms, in the presence of firm heterogeneity and continuing stochastic innovative entry. Silverberg and Verspagen (1994) analyse a general model that derives basic common mechanisms that relate market structure dynamics to endogenous technological change, R&D investments and economic growth. In a similar fashion, Aversi et al. (1999) investigate a model with bounded rational, heterogeneous agents, which generates aggregate
Table 1
Demand, innovation and entry in History-friendly models.

<table>
<thead>
<tr>
<th>Demand</th>
<th>Demand Malerba et al. (1999, 2001a,b, 2006, 2008a,b)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Innovation at technological levels</th>
<th>Malerba et al. (1999, 2001a,b, 2006, 2008a,b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exogenous shock introduces new technologies (integrated circuits and microprocessors) and new markets (PC). Kim and Lee (2003) Endogenously determined by productivity of firms, given an exogenous productivity growth parameter and the exogenous basic productivity of the industry.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Innovation at firm level</th>
<th>Malerba et al. (1999, 2001a,b, 2006, 2008a,b)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Entry</th>
<th>Malerba et al. (1999, 2001a,b, 2006, 2008a,b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brenner and Murmann (2003) Endogenous entry determined by the number of unemployed researchers and the size of the market.</td>
<td></td>
</tr>
</tbody>
</table>
dynamics of demand in line with some general empirical stylised facts about consumption.

On the contrary, with regard to the analysis of results in “History-friendly” models, we have clearly stated that the goal of this approach is not the reconstruction of general properties of socio-economic processes but the investigation of historical “stylised facts”. This purpose is “industry-specific”, aimed at resembling in a stylised and formal way the functioning of an industry. The structure of “History-friendly” simulation models rests in reproducing in the laboratory an artificial industry, made by artificial agents and environments, in order to investigate and explain some social and economic phenomena, which are contextualised in a given industry. The essence of this analysis consists of testing the functioning of the model in order to understand what the factors and fundamental processes are that make the model behave as it does. Given a real world data-generation process, the goal of the researcher is to develop a model data-generation process that is able to represent a good approximation of the real one. In this way, the modeller aims at gaining some knowledge on the underlying causal structure of the real world under analysis (Windrum et al., 2007).

However, “History-friendly” simulations do not aim at reproducing results that empirically match the quantitative specific values of available historical records closely. The match between the results of the model and the evidence about the industry is understood in a qualitative way. What really matters here is to construct a consistent and logically coherent model in order to be able to explore and investigate which features are responsible for the evolution and functioning of the complex system under study. In this sense, we follow an “analytical sociology” approach and a mechanism-based causal explanation (Barbera, 2004; Boero and Squazzoni, 2005): the target to be explained is a social macro regularity both spatially and temporally organised; agents, their actions and their interactions are the objects we model; theories are the framework of the causal mechanisms and generative relationships in the model. In this context, “History-friendly” simulations are seen as generative tools that aim at formalising a representation of the micro-macro mechanism believed to be at the basis of the social regularity we observe. In terms of Fagiolo et al. (2007a) and Windrum et al. (2007) taxonomy, “History-friendly” models investigate the micro aspects of transient multiple phenomena, both from a quantitative and a qualitative perspective.

The debate about the contestability of history and the theoretical models guided by history raises methodological issues. “History-friendly” simulations have been considered to aim at replicating the “fall of one particular leaf from a tree”, quoting a parallel by Silverberg6 (see also Windrum et al., 2007) and thus criticised because they aim at describing a “typical” history, that may not exist in reality: it may simply be the result of a casual event, the unique roll of the dice, it may only be one of a set of possible worlds (Windrum, 2007; Windrum et al., 2007). With regard to this point, the present paper considers the historical event under study not as the outcome of a sequence of random occurrences but as the result of some causal and systematic explanations. The historical evolution of industries presents phenomena generating other specific phenomena: “History-friendly” simulations try to build mechanisms and causal forces that are able to explain the generation of these phenomena, with regard to a specific industry. The actual history we observe displays deterministic attributes that give birth to given trajectories and phenomena. Thus, “History-friendly” models are interested in:

- explaining how it is possible that we are observing now these phenomena;
- studying the conditions and rules that make an event happen under some circumstances;
- explaining the conditions and the deterministic explanations behind the realisation of that “typical” history that occurred.

Instead of following the “History-friendly” model approach, Windrum et al. (2007) suggest proceeding by trying to develop more general models that explain the “fall of many leaves from many trees”. This is an interesting issue that we believe is still open in “History-friendly” modelling at the moment, linked to the declared purpose of generalisation. Understanding the basic processes of different given industries might enable us to find generalisations, invariances, similarities or distinctions among these industries, creating interesting inductive exercises to push the reasoning back to think about what factors and processes are responsible for industry evolution. With regard to this, at present the performance of the “History-friendly” modelling style is lacking and further research needs to be done.

Analyses in published “History-friendly” models are usually presented in at least four steps: a description of the historical phenomena and stylised facts to be explained; the relevant theories believed to be crucial for explaining the historical phenomena; the structure of the model; and the results of the simulations.

1. The first section is important for both the modeller and the reader: to give the reader a clear idea of the phenomena that the simulation model will study; and to help the modeller gain a deep knowledge of the environment he/she is modelling, giving the assumptions a greater degree of realism and thus giving the model framework more credibility and acceptability.

2. The second step is also fundamental for both the modeller and the reader: for the modeller, in order to represent the basic relationships and structure of the economic environment under study, thus incorporating the theoretical background into the programming language of the model; for the reader, in order to comprehend what theoretical explanations are under investigation. At this stage, repeated processes of internal verification of the correctness of the computer implementation of the model are crucial. Typically, an attempt to develop a model with detailed and rich assumptions and structure would bring deep accuracy in the description and correspondence of the model

---

framework to the system being studied at the expense of a less transparent representation of the causal relations between structure and implications. At this stage, then, the modeller faces an obvious trade-off between descriptive accuracy of the model and explanatory power (Pyka and Fagiolo, 2005; Windrum et al., 2007).

3. The third step outlines the most relevant features of the model in order to understand its structure and functioning.

4. After the calibration process that requires a repeated change in the parameters and methods to obtain a satisfactory specification of the model, the researcher presents in the fourth step the outcome of the simulations and discusses the results.

Table 2 summarises the discussion.

The discussion of the results in the published papers typically presents a section of analysis of the standard simulation runs and a section in which the analysis investigates different settings of the parameters and structure, as well as counterfactual analyses. The results are the average outcomes of a number of simulation runs, thus averaging out outlying values. Given a setting up of the parameters and methods, the model run produces simulated data on which some relevant (for the purposes of the analysis) statistics are computed. A process of sensitivity analysis is implemented in order to check the robustness of the model outputs by determining how variations in the values of the parameters generate changes in the dynamic behaviour of the system. A meticulous sensitivity analysis controls how the outcomes vary when the level of aggregation of the variables in the model, the timing structure and the distribution of the random elements vary. We believe some progress has been made in “History-friendly” modelling along these lines since the seminal paper was published. However, we suggest there should be a more intense discussion of sensitivity analyses in future papers, controlling the robustness of the results against initial parameters, methods and variability due to the presence of stochastic components.

Consider, for example, the first and the last published papers that present a “History-friendly” model. In the seminal paper (Malerba et al., 1999), the authors do not report the parametrisation of the model and they analyse the results in much less detail compared to their most recently published paper (Malerba et al., 2008b). In this last paper, the reader can find a list of the main parameters and a more complete discussion about the sensitivity analysis carried out for building and testing the model results.7 Extensive Monte Carlo exercises have been conducted in order to check the robustness of the model under different parameter settings, whereas in the seminal paper, only a few words were devoted to both parametrisation and simulation checks. Between these two papers, the same authors published other works that report continuous, more meticulous descriptions of sensitivity analyses, Monte Carlo experiments, and tests of the results. My experience (after 10 years of collaboration with Malerba, Nelson, Orsenigo and Winter) is that increasing attention has been paid to the processes of internal validation and calibration of “History-friendly” simulation models. After development of the initial model, a long process of checks for possible ‘bugs’ is conducted. After that, the basic hypotheses of the models are explored under a range of parameter settings in order to test if the model performs as expected. Several ‘unit tests’ are conducted (i.e. the run of the simulation after every modification in the program in order to check that a ‘bug’ has not been brought in). Often, the simulations are run under very simple and extreme scenarios, where the results of the model are straightforwardly predictable. Also, in many cases, the models are divided into subsets that can be analysed separately. Finally, a process of calibration involves stages of iterative analyses of the results of the simulations under different versions of methods and parameter settings, where the parameters have been previously determined as a defined value or a range of values.

A clear example of a highly detailed process of parametrisation and analysis can be found in Brenner and Murmann (2003). The authors initially define the values of the parameters that can be observed with precision.

---

7 See pages 227 and 229 in the paper.
Then, for those parameters whose value ranges in predefined intervals, the simulation process randomly draws the values of each parameter: Each draw makes up the support of a separate group of simulations that represent a set of possible worlds. Within each group of simulations, the independent variables are varied and the results in terms of the dependent variable are recorded. This process, repeated for each group, generates datasets of independent and dependent variables that are analysed with statistical regressions. If the causal relationship between the independent and dependent variables demonstrates the same functional form across all possible worlds, the authors argue that this causal relationship represents a “true description of the real world” (Brenner and Murmann, 2003). The accuracy of this process has to face at least two drawbacks: for the readers, the dense and technical writing may turn out not to be immediately comprehensible; for the authors, the usual limits imposed on the number of words for a paper when publishing in journals may make it impossible to give a complete and clear description of the models and implemented simulations.

4.2. Counterfactual analysis

Counterfactual analyses (known also as if-worlds, all-history, cliometrics, uchronia) consist in the falsification of an antecedent (X) in order to investigate the consequent (Y). Analysing alternative versions of the past is a useful exercise for debating about the functioning and dynamics of the system under study and evaluating how some changes would have influenced the course of actions. Thus, counterfactual comparisons may be able to indicate causal conclusions and prescriptions that facilitate future improvements. In this sense “History-friendly” models may represent a normative tool to investigate different institutional, socio-economic settings, industrial public policies as a support for decision makers.

Many economists think counterfactually without admitting they are doing so. This happens in particular in policy evaluations, in which comparisons imply the prediction or imagination of alternative states. This process may be referred to as evaluation of the impact of past policy measures, or it may be used for formulation of future proposals. More generally, economists often reason by comparisons that imply counterfactuals. Von Mises (1949) wrote that the method of economics is the method of imaginary constructions and “the main formula for designing imaginary constructions is to abstract from the operation of some conditions present in the actual action. Then we are in a position to grasp the hypothetical consequences of the absence of these conditions and to conceive the effects of their existence”: a process that closely recalls counterfactual thinking.

The debate on the possibility of conducting counterfactual analysis and its practical value is lively. Some researchers argue that valuable counterfactuals are difficult to perform and therefore infrequently useful. Some other scholars look more favourably at the validity of counterfactual experiments and apply them more broadly. Good counterfactual exercises might be informative and beneficial to the economic investigation because counterfactual reasoning helps the researcher and policy maker to be more receptive towards the role of contingency in key decisions and events (Lebow, 2007). Also, counterfactual simulations help in escaping from the cultural bias that offers the beliefs that structure the understanding of social reality, by raising different and divergent research questions: this exercise, even if visibly provocative, might raise issues otherwise disregarded, or investigated from different points of view.

In the present paper we argue to develop counterfactual discussions following basic principles from an appreciative theorising perspective. Recognising the historical nature of evolutionary economics, we symbolise history as a tree (Cowan and Foray, 2002; Cameron, 2006). If we accept the branching view of history, counterfactual analyses mean to move back down the tree in order to study a branch that has not been run and analyse the conditions that may lead the course of actions along a different branch (Figure 2). Thus, in order to understand the present through the past, counterfactuals help in “understanding the major events and chains of decisions which, coupled with processes that magnified rather than damped their effects, can be considered as having played a role in disconnecting some sub-regions of the tree from the branches followed by actual history” (Cowan and Foray, 2002).

We argue it is crucial that “History-friendly” models develop counterfactual reasoning by linking the antecedent X and the consequent Y by considerations of rationality and clear causal connections (Bunzl, 2004). There are no laws of nature in the processes of historical and social sciences like economics. The researcher must then discuss the steps linking X to Y, raising indirect evidence
in support of them, discussing the causality of the principles, rules and/or laws that connect X to Y, and examining the reasonableness of the inferences made. Certainly, theories that explicate restrictive causal forces and that impose rational and logical constraints to reasoning limit the alternatives we can analyse and what “our” history can generate but the analysis becomes more compelling and agreeable.

Moreover, good counterfactuals need to keep the analysis as simple as possible so that the less strong the impact of the changes in the antecedent X, the more plausible is the counterfactual reasoning (Lebow, 2007).

Moreover, we argue the need to discuss the plausibility of the change in antecedent X, and the reasons why this change did not occur in the course of events.

“History-friendly” simulations make extensive use of counterfactual investigations. Malerba and colleagues give particular attention to counterfactuals both as a testing procedure and an investigative device. Most papers present an entire section dedicated to counterfactuals as a test for the consistency of the logic of the model and for the sensitivity of the results to the parameterisation.

The use of counterfactual thinking in “History-friendly” models resembles a reconstruction of alternative scenarios through which authors discuss the evolution of the industry under examination, as was done in the area of political history by Ferguson and colleagues (Ferguson, 1997). In Ferguson’s book, there are different contributions on counterfactual historical analyses, discussing deterministic and contingent features in history: various points in history that could have evolved in different ways to how they actually did are investigated, thus providing support on how counterfactual reasoning can be valuable if carried out with historical evidence and with regard to plausibility and implausibility (Bernstein, 2000). Similarly, “History-friendly” models describe the deterministic attributes that give birth to a given industry's evolution and raise discussions on how some patterns and phenomena might have evolved differently as a result of other plausible causal explanations.

Counterfactual analyses in published “History-friendly” models are presented in a clear and simple way for the reader, with specific indications on which is the antecedent and what the effect on the consequent is. Usually, counterfactual reasoning is used to understand the role of policy interventions. In some situations it might be preferable to restrict the number of counterfactual investigations and devote more discussion to the causal connections that link the changes in the antecedent to the changes in the consequent.

The main limitation of these counterfactual investigations is that they provide sufficient circumstances (in the structure of the model) for some events to occur but no necessary conditions. For a discussion on this point we refer to Marks (2007). Moreover, there are very few discussions about the plausibility of the change in antecedent X, and the reasons why this change does not occur.

Some examples of counterfactual reasoning are given in the following papers.

Section 4.2 in Malerba and Orsenigo (2002) raises some interesting questions about market concentration and technological discontinuity in pharmaceuticals: what conditions would lead to a more concentrated industry? Under what circumstances would new small biotechnological firms have gained market leadership? The counterfactual simulation experiments show that an increase in the antecedent ‘absolute costs of research’ would have increased the level of concentration in the industry, whereas economies of scale would have played very little role in stirring up concentration. In fact, the other two antecedents that would have increased concentration are found to be a less fragmented structure of the market and a lower degree of technological opportunities.

With regard to the second question, simulation experiments show no reasonable changes in the antecedent conditions under which new small biotechnological firms would have been able to displace the incumbents: even if the efficiency gap between biotech firms and the incumbents is remarkably (and implausibly) increased in the parameter setting of the model, incumbent firms continue to gain from past innovations and products and resist the threat from new biotechnical firms by signing contracts and agreements with them.

Section 3.3 in Brenner and Murmann (2003) analyses the causes behind the dominance of German firms in the dye industry in the world market. The argument under discussion is a test for the effect of the antecedents — the ‘high initial number of organic chemists available in Germany’ and the ‘high responsiveness of German universities to the demand for chemists’ — on the emergence of German leadership in the market. Results of the counterfactual simulations do not support the argument, showing that this explanation is not independent of the values of other parameters in the model.

Section 6.2 in Malerba et al. (2008b) investigates which conditions would have led to more specialisation or vertical integration among computer firms in the semiconductor industry. The results show that the lack of external markets for components and a minor technological discontinuity in microprocessor technology would have led to more vertical integration of computer firms into semiconductors, while no demand for lock-in effects in semiconductors does not necessarily lead to more vertical integration, thus emphasising the role of the external market and technological discontinuity as the antecedents able to shape coevolution of the semiconductor and computer industries and the boundaries of firms. Finally, it is shown that the absence of lock-in effects in the demand for mainframes would lead computer firms to specialise.

Section 5 in Malerba et al. (2001b) and Sections 7 and 8 in Malerba et al. (2008a) discuss interesting policy issues in a counterfactual fashion.

As we mentioned in Section 3 of this paper, Malerba et al. (2001b) focus their analysis on the role of the antitrust policy in confronting the rise of the near monopolist (IBM) in the mainframe market. Antitrust laws represent one of the main public interventions for controlling monopoly power that characterises the industry under examination. In particular, the emphasis of the analysis is placed on the relevance of the timing of the Antitrust authority’s intervention: this is the antecedent in this analysis. The results of the simulation runs show that in the case of immediate intervention the market very soon becomes concentrated.
again, given that the increasing returns are still quite high, thus requiring continuous control by the authority. Intervention which is not immediate leads to a situation in which the emergence of a monopolist takes much more time and in the end the level of the concentration ratio is significantly reduced. The reason for this could be that the break-up has the effect of making firms more similar in terms of size and resources. If the Antitrust authority intervenes late, on the other hand, the largest firm will have already reached a significant size, so that the two resultant firms are already large compared to other competitors. Thus, one of these two new large firms will gain the leadership, leading the market toward renewed high concentration. Finally, if intervention is very late, the increasing returns will have started to fade away and consequently the market becomes a duopoly. The conclusion of the counterfactual investigation is that a change in the timing of intervention by the Antitrust authority would affect the evolution of the industrial structure of the computer industry.

In Malerba et al. (2008b), the counterfactual investigations focus on the efficacy problems that policies face in dynamic environments. Unintended policy results and inter-industry effects may emerge from different policy interventions. For example, it is shown that Antitrust policies that aim at reducing demand lock-ins and supporting open standards, on the one hand, would increase competition among producers but, on the other hand, given the disappearance of a monopolist firm (and consequent high investment in R&D), they would reduce the quality of the computers sold. These policies would also affect the boundaries of firms: the amount of vertical integration in the mainframe industry is reduced in comparison with standard simulations. Another example relates to the role of public procurement. If public procurement is modelled as selective and temporary by buying products from the best producer for a given number of periods, the effects on industry evolution would be a higher concentration in the transistor and integrated circuit industries, an increased quality of semiconductors and an increased tendency toward specialisation by mainframe firms. On the contrary, if public procurement is modelled as permanent and targeted at many component producers, the transistor industry would be more competitive while the integrated circuit industry would become even more concentrated with higher quality products. This would, in the end, push the mainframe leader to specialise more frequently. What emerges is the conventional Schumpeterian trade-off between temporary monopoly and higher innovation (i.e., product quality in the model).

An interesting remark we raise here relates to the possibility that counterfactuals stimulate the investigations of some relations from an empirical point of view, thus establishing a link between theory and empirical analysis (Cowan and Foray, 2002; Dawid and Fagiolo, 2008). For example, the policy issues we mentioned need to be empirically investigated concerning the effects of political interventions in different stages of industry evolution. Empirical research, where possible, may weigh against the argument of these simulation results. Also, empirical investigations on the unintended effects of policy interventions, as well as inter-industry effects, could be crucial for public authorities.

5. The debate on empirical validation in simulations: the “History-friendly” models

Tesfatsion (2006) describes the difficulty of empirical validation as one of the main drawbacks of agent-based models. It is important to understand how “History-friendly” simulations are placed in the debate about empirical validation of simulation models (see among others Brenner and Murmann, 2003; Brenner and Murmann, 2004; Boero and Squazzoni, 2005; Pyka and Fagiolo, 2005; Richiardi et al., 2007; Windrum et al., 2007).

There are, in the economics literature, different interpretations of the concept of validation and several attempts at empirically validating the model that generate a basic confusion (Bailey, 1988; Richiardi et al., 2007). We mentioned in Section 4.1 that the purpose of “History-friendly” models is to resemble the functioning of a given industry in a stylised and formal way in order to simulate the patterns of evolution of the main features of that industry. The recognition that the world may have evolved in so many more different ways than it did leads the researcher to investigate the motivations and principles that explain logically why things actually happened the way they did, with the help of a logically coherent model that does not necessarily need to closely empirically match the quantitative specific values of historical records.

What we said does not imply that the structure, values and parameters of the model can be chosen freely. We wish to conduct this research in an accurate and empirically oriented way. Among the different approaches of empirical validation in agent-based models (Fagiolo et al., 2007a,b; Windrum et al., 2007), we suggest a process for developing the models that should explicitly structure the behaviour of the agents and the choice of the parameters in order to be as accurate and coherent as possible with available empirical evidence. This process of “input validation” (Bianchi et al., 2007) is necessary for at least two reasons. First, there may be different settings of the model that lead to a given outcome. Then, the strategy of being closely related to the empirical observations in the choice of parameters’ values and structure helps the researcher to define the starting characteristics of the model. Second, selecting parameters and values of the model that resemble empirical observations may help in giving accuracy to the structure of the model and credibility to its functioning and results. The availability of accurate data, empirical studies and anecdotal evidence about the phenomena under investigation becomes, thus, important both initially in helping to define the initial setting of the model, and in the end, in empirical validation of the “simulated traced history”.

With regard to empirical validation, Brenner and Murmann (2003) discuss an interesting way of implementing simulations which can induce with great confidence that “the hypothesized causal relationship indeed represents a true description of the real world”. Following their procedure of developing the model, they obtain interesting results, in a “History-friendly” style, about the evolution
of the synthetic dye industry in Germany at the beginning of the twentieth century. However, some values, parameters, representations and above all, agents' behaviour are inherently not measurable or quantitatively representable. This obviously implies that in every simulation model some degree of arbitrariness is unavoidable. Models are by definition subjective in their development of assumptions and structure of behaviour. In this paper, we suggest that "History-friendly" researchers focus on simulations that enable the evolution of an industry to be logically explored, studied and investigated in a coherent and appreciative way rather than trying to represent a 'true' description of the real world. Given the intrinsically limited empirical expressiveness that characterises some processes of the model, what matters to the researcher is to implement interpretations that exceed this limitation (David et al., 2005). In fact, even though there is the possibility of using precise empirical data to set the parameters, the algorithms, methods and behaviour of the simulations cannot 'represent' the real functioning of the reality but are an interpretation of the functioning of reality, as perceived by the modeller.

The validation approach followed by Malerba and colleagues is less detailed than in Brenner and Murmann (2003). However, these authors are clear about their goals. The match between the results of the model and the evidence is understood in a qualitative way and they aim to resemble the evolution of a given industry in a stylised way. Their models "do not attempt detailed calibration of parameters. Because most parameters fall into groups within a particular mechanism in the model, common-sense guidance is available for choosing plausible orders of magnitude" (Malerba et al., 2008a).

We stress the importance of accentuating the difference between validity and accuracy in the debate on empirical validation of agent-based economic models (Robinson, 2004; Schmid, 2005). The process of validation is strictly related to the purpose. This concept refers to a coherent structure of beliefs that serve the purpose of the simulation to show that simulation may help in investigating and explaining some logical relationships and causal links. The accuracy of the model is, instead, related to the correspondence of the structure and results of the model to the available empirical observations.

In this sense, "History-friendly" models should follow an explicit process of validation by declaring the aim and the way of achieving it. With regard to this debate, we believe "History-friendly" modelling procedures match at least two of the various validation techniques discussed in the social science literature (Carley, 1996): grounding and calibration. The central scope of grounding "is to establish that the simplifications made in designing the model do not seriously detract from its credibility and the likelihood that it will provide important insights" (Carley, 1996), such that the simulation model is able to capture the basic elements of a social process as expressed in a verbal theory. Moreover, in the development of the models, a process of setting and resetting the parameters, algorithms and routines of the simulation in order to coherently represent the logic behind some processes is repeatedly carried out, so that the internal workings of the model are robust.

A final remark is needed. Criticism about simulation models is often made that, with a sufficiently large number of parameters that can be manipulated, the researcher can make the model do anything he/she wants. We argue that this criticism is not fully appropriate for "History-friendly" models (see also Carley, 1996; Malerba et al., 2008a). The processes of the models have a structure based on algorithms, rules, and interrelated causal mechanisms that interact in complex and non-linear ways. Changing the model's parameters and procedures in order to generate the desired pattern would thus affect the rest of the model and, consequently, its outcome in many other respects. It is not likely, then, that manipulations of some parameters can generate the intended overall model patterns. Moreover, the time sequence in the model gives another restriction in the arbitrariness of parameter definition: in fact, the choice of the value of a parameter must take into account its consistency with the meaning that one period stands for (one year, one month, one day or else) in the chosen simulation time sequence (Malerba et al., 2008a).

6. Challenges for “History-friendly” models and conclusions

This paper aims at creating some order in the new branch of models called "History-friendly" simulations. After a brief recall of key principles of the industrial dynamics approach, we have analysed the basic features and methodology of the "History-friendly" approach.

We emphasise the concept of complexity that characterises the socio-economic world and its ongoing process of change. We discuss the flexibility of simulation models for assessing with complexity, as well as the property of emergence of agent-based simulations. Finally, we recognise that the dynamics in industries present high specificity: each industry follows particular trajectories and demonstrates specific traits. These considerations represent the background in which the need for a "History-friendly" modelling style arises.

"History-friendly" models provide a theoretical tool that can engage in dialogue with the logical examination and causal explanations presented by empirical studies, by putting together a deep understanding of the historical events of an industry's evolution and a formalisation of the relationships that govern the behavioural routines for capturing the industry dynamics. These models construct a consistent and logically coherent framework to explore and investigate from an appreciative perspective which features are responsible for the evolution and functioning of the complex system under examination.

In some respects, we think that "History-friendly" models have contributed richly to the economics literature but that their potential has not yet been fully exploited. The claim of this paper is that, on the one hand, "History-friendly" methodology so far represents a useful toolkit for the analysis and understanding of industrial dynamics. Since the very first discussions on "History-friendly" simulations about 10 years ago, a lot of work has been done and improvements have been made. "History-friendly" simulations have also contributed to the important debate on the empirical embeddedness of simulations and some
progress has been made in recent years, but here we stress the importance that more work needs to be done in this direction. On the other hand, this family of models still has to grow and improve. Some of the initial goals and ambitions have only partially been fulfilled (if fulfilled at all).

For example, if we can recognise the fruitful dialogue between researchers involved in “History-friendly” simulations and historians and empirical researchers, we cannot say the same for the contribution that “History-friendly” models have made in raising new empirical insights or questions.

Also, with respect to the declared purpose of ‘generalisation’, we believe that “History-friendly” simulations may represent “case-based models” aiming at moving toward ‘typifications’, i.e. aiming at investigating “some properties that apply to a wide range of empirical phenomena that share some common features” (Boero and Squazzoni, 2005).

Despite the potential role of these models, the generalisation of their results is lacking in existing simulations. This remains a challenge for future development. This lack may also be related to the fact that there is significant heterogeneity among existing models.

Related to this, Richiardi et al. (2007) claim the need for a common protocol in agent-based models. The heterogeneity of the approaches in agent-based computational modelling is currently still too broad in order to reach a common protocol soon. However, we believe that at least inside the given sub-families of models, as in the “History-friendly” case, such an attempt could be pursued. The development of common libraries that could generate a platform for modular models would contribute to diffusion of the approach.

With regard to developing analyses in “History-friendly” simulations, a discussion of parameter selection and model calibration is often taken for granted, reporting simply that extensive sensitivity analyses and calibration refinements have been carried out. So far, there is little possibility of replicating the simulation results of published papers.

Finally, it would be stimulating to work out ‘future counterfactuals’, in which the researcher investigates potential future conditions that could lead to different outcomes. So far, we have driven our research agenda ‘looking in the rearview mirror’: the challenge lies in looking further ahead. This prospect is highly ambitious but it may contribute to stimulating a debate about the normative role of simulation models in economics.

To conclude, we believe this paper helps to facilitate the implementation of “History-friendly models”, provides a more robust methodology for model development, encourages the use of this promising toolkit to study industrial dynamics, and raises challenges that may indicate the direction for future research.

Acknowledgements

The author acknowledges the financial support of the Italian Ministry for Education, Universities and Research (FIRB, Project RISC - RBNE039XXA: “Research and entrepreneurship in the knowledge-based economy: the effects on the competitiveness of Italy in the European Union”).

I would like to thank the following for their useful suggestions: Luca Berga, Franco Malerba, Luigi Orsenigo, participants of the Workshop “Evolutionary Models on Industrial Dynamics: Issues, Debates and Prospects” (29 June 2004, Ecis, Eindhoven), participants of the FIRB - RISC Conference “Research and entrepreneurship in the knowledge-based economy” (7–8 September 2009, KITeS, Milan) and two anonymous referees. The usual disclaimers apply.

References


Andersen, E.S., Valente, M., 1999. The two software cultures and the evolution of evolutionary economic simulation. DRUID Papers, DRUID and Aalborg University, Denmark, Version 1 December 1999.


