History friendly models: retrospective and future perspectives

HISTORY FRIENDLY MODELS:

RETROSPECTIVE AND FUTURE PERSPECTIVES

Gianluca Capone[a,b,+], Franco Malerba[b,c], Richard R. Nelson[d], Luigi Orsenigo[b,e,f] and Sidney G. Winter[g]

[a] Department of Economics and Management – University of Pisa
Pisa, Italy

[b] ICRIOS – Bocconi University
Milan, Italy

[c] Department of Management and Technology – Bocconi University
Milan, Italy

[d] Columbia University
New York, USA

[e] Scuola Universitaria Superiore IUSS Pavia
Pavia, Italy

[f] Science Policy Research Unit – University of Sussex
Brighton, UK

[g] The Wharton School – University of Pennsylvania
Philadelphia, USA

[+] Corresponding author: gianluca.capone@unipi.it

This paper has been published as:

Capone, G., Malerba, F., Nelson, R.R., Orsenigo, L., Winter, S.G., 2019. “History friendly models: Retrospective and future perspectives”. *Eurasian Business Review* 9: 1-23; doi: 10.1007/s40821-019-00121-0.

HISTORY FRIENDLY MODELS:

RETROSPECTIVE AND FUTURE PERSPECTIVES

ABSTRACT

Twenty years ago, we introduced the history friendly modelling approach to formally study industrial dynamics. In this paper we look retrospectively at the results that the history friendly literature has achieved so far and what are the challenges ahead of us. We illustrate the main principles, methods, and building blocks of the approach, and then we illustrate it through two applications. The first one investigates the impact of entry in the mainframes segment of the computer industry. The second application studies the effect of different industrial policies in uncertain technological environments.

Introduction

Twenty years ago, some of us (Malerba et al., 1999) introduced a new modelling approach to formally study industrial dynamics. The approach was labelled "history-friendly modelling" and aimed at explaining specific patterns of evolution empirically observed in certain industries (Malerba et al., 2016). Our choice was motivated by the perception of current economic models about industry dynamics - even evolutionary models - as rather general and abstract, and far from the rich evidence emerging from empirical studies. Historical accounts and policy-oriented analyses of specific industries, in fact, suggested that the factors shaping industry evolution are diverse and quite different from one industry to another. The differences include a variety of elements, such as the nature of innovative opportunities, the formal and informal rules governing the appropriability of returns from innovation, the cumulativeness of technical change, the way firm behave confronting such environments, the link among sectors, the role of institutions.

Still, history friendly models are evolutionary models: they therefore share the theoretical underpinnings of evolutionary theory (Nelson and Winter, 1982). This is particularly relevant for the way agents, and especially business firms, are represented. Model firms are heterogeneous and quite complicated entities, that need to take decisions on many dimensions of interests, such as prices, R&D investments, technological trajectories, domains to explore, marketing expenditures, products to develop, or the vertical scope of the firm. In all of these dimensions, firm choices are driven by decision rules that respond to available data on both the firm and its environment, and are "boundedly" rational (Simon, 1955; Cyert and March, 1963). These general features about how agents behave are also the result of the history of the company and the sector and are dependent by specific representations of the institutional, technological and market environment (e.g. the patent system, the presence and size of alternative markets, the role of actors such as the government, universities, venture capital and so on).

In this paper we look retrospectively at the results that the history friendly literature has achieved so far and what are the challenges ahead of us. In the following section, we present the theoretical principles of the history friendly approach. Then, we illustrate in a comparative way the key ingredients of past models. Sections 4 and 5 report some original applications of existing models, that allow to highlight both the potential and the limits of the approach. Finally, we conclude by discussing future opportunities for research in the area.

History Friendly Models: Principles and Methods

Explaining the mechanisms behind the dynamics of a particular industry requires some elements of theory. Scholars studying the empirical patterns of industry evolution generally rely quite heavily on "appreciative" theories, as a tool to characterize the main mechanisms at work (Nelson et al., 2018).

Appreciative theories are typically verbal (rather than formal) theories, i.e. they are expressed through words (rather than formal symbols and equations); they have some elements of storytelling, and in general are more easily accessible to the general reader. However, appreciative theories are causal theories: they include elements and arguments that analysts believe to play an effective role in driving the empirical phenomenon under study. Although appreciative theories are quite adherent to the empirical reality and offer convincing explanations of the observed patterns, they have a major structural limitation. In fact, it is quite difficult to verify the logical consistency of such theories, due to their complexity and the often lack of precision of the verbal language. History-friendly models allow to overcome this major limit: they aim to capture and represent in formal terms the main mechanisms put forth by appreciative theories, in order to explore their logical consistency through the power of formal methods (Malerba et al., 1999; 2016).

To this purpose, building history friendly models requires three important steps. First, there is the selection of the phenomena deserving attention from a theoretical point of view. Economic history, and specifically the history and evolution of specific sectors, provides guidance in this step, together with informed evaluation by the scholars that want to undertake such an enterprise. In selecting the specific historical episodes and the industries to investigate, we have taken into account both their relevance in the industrial transformation of sectors and more generally for economic growth, and the role specific institutions, technological change and organizational knowledge.

The second step is the choice of how to represent the selected phenomena. In this respect, history-friendly models are fully evolutionary models, since they adopt the same basic representations used in evolutionary economic theory. The central actors are business firms that are heterogeneous and characterized by idiosyncratic capabilities. The focus is not only (and not much) on pricing and current production activities, but mostly on their innovative behaviour, which is seen as a key causal mechanism driving both firm performance and industry evolution. The external environment, in the form of demand conditions, technological change, and competition from rivals, also shapes their choices. A particular importance is also assigned to the time dimension: many features of the firms and the environment change over time, but there is no arbitrary assumption of equilibrium conditions prevailing at any period. However, history plays guidance also in this second step. Empirical studies of many industries show that elements specific to each individual history must be considered to provide explanations that are coherent with the mechanisms highlighted by appreciative theories. These elements may include institutional aspects (e.g. the patent system in the pharmaceutical industry), government decisions (e.g. the intense funding of R&D activities for military purposes in the early stage of the computer industry), the structure of the demand (e.g. the existence of preferences for different products or mixes of product characteristics), the dynamics of technical change (e.g. the introduction of new generations of semiconductor components in the computer industry), or firm strategic choices (e.g. the decision of IBM to develop the personal computer through an independent subsidiary).

Finally, the third step is about manipulation and implementation of the model designed in the second step, to get meaningful results out of it. Here, history-friendly models are definitively part of the broad family of agent-based simulation models (Dawid, 2006). Agent-based models are simulation models dealing with the complexity of economic systems (Tesfatsion, 2002). They are characterized by the emergence of properties at the system level from the complex and non-linear interactions between boundedly-rational agents (Windrum et al., 2007). The use of computer simulations has become more widely accessible due to advances in information technology in recent years and is required to represent the large number of dimensions that are involved in the analysis and to translate heterogeneous agents and their complex interactions into workable models. History is useful also in this step. Typical agent-based models have a large number of variables and parameters, and as many equations specifying interactions among them. Historical evidence about the industry helps to reduce the range of choices that the modeller otherwise has to do arbitrarily in absence of theory indications, as might be the case for the initial number of firms entering an industry or the relevant time span for the running of a simulation. Historical and theoretical considerations provide indications also on the variables to select to assess the results of the model. These variables do not play an active role in the functioning of the model, but are summary indicators of key trends: they must be properly defined and implemented in the programming code since the beginning. The analysis is typically conducted over mean values of these indicators, to reduce the impact of initial conditions and stochastic elements of the model. These values are obtained as averages from a large number of simulation runs, in which the value of the parameters is kept fixed. Still, it can be very useful also an analysis of individual runs of the model, as well as a sensitivity analysis, to evaluate the extent to which observed averages are a correct representation of simulated histories. Once the model is built, there is room for wider applications than just the replication of the observed patterns in the history of the industry. Indeed, a proper exploration of counter-factual simulations is needed to understand the relevance and coherence of causal mechanisms proposed by the underlying appreciative theories, although with the awareness of how tricky this can be, especially in the context of evolutionary economics (Cowan and Foray, 2002).

History Friendly Models: Review and Building Blocks

The first history-friendly model that we developed twenty years ago analyzed the long-term evolution of the computer industry, starting from its birth and the early growth of the mainframes segment to the introduction and growth of the personal computer segment (Malerba et al., 1999). Since then, a few works have applied the methodology to different industries. Malerba and Orsenigo (2002) and Garavaglia et al. (2013) analyzed the evolution of the pharmaceutical industry, focusing in particular on the era of random screening, and explored the effects of technologies with a low degree of cumulativeness and with markets that are strongly fragmented. The co-evolution of two industries (computers and semiconductors) is the focus of Malerba et al. (2008), with the explicit aim of highlighting the role of competences accumulation to explain the dynamics of firms vertical integration and specialisation between the two industries over different technological generations. These three models for computers, pharmaceuticals, and semiconductors and computers are now also available in extended and revised version in Malerba et al. (2016), that also includes the simulation code to reproduce the results. Brenner and Murmann (2003, 2016) analyzed the German synthetic dye industry in the late decades of the nineteenth century until World War I, showing that the catching-up of German firms with respect to France and British competitors and their sustained leadership in the following years was driven mostly by the strength and responsiveness of the national university system, rather than the existing stock of competences and the late introduction of the patent system in the country. Kim and Lee (2003) developed a model to study the evolution of the Random Access Memory chip industry (a relevant segment of semiconductors) in the 1970s and the 1980s, to explain the displacement of small specialised firms by larger diversified firms that entered at a later stage. The model proposes a combination of cumulativeness, process innovations, and large scale investments as main factors driving the observed patterns. More recent models include: (a) Malerba and Yoon (2011), on the recent emergence of a vertical structure in the semiconductor industry, with foundries pursuing chip manufacturing and fabless firms specialising on design, a dynamics associated to the increasing importance of application-specific knowledge; (b) Fontana and Zirulia (2015) on the local area networking (LAN) industry, a multi-market industry, in which the dominant firm in a segment (Cisco Systems in the router market) was able to extend its primacy to a new emerging segment (the switch market) through a strategy of combining the acquisition of the market pioneer and then leader and the exploitation of compatibility and switching costs; (c) Landini et al. (2017) on mobile phones and semiconductors, in which technological changes drives changes in industrial leadership across countries and firms; (d) Li et al. (2018) on the mobile communications industry in China from 2G to 4G, which highlights the interplay of segmented markets and generational technical changes as drivers of the long-term process of capabilities accumulation that allowed Chinese firms to catch-up with more advanced multinational competitors both in the domestic and the international market.

Starting from the baseline models developed for specific industries, further applications have explored public policy issues. Malerba et al. (2001) discuss the challenges of antitrust policy dealing with the monopoly emerging in the computer industry and reinforced by the process of accumulation of technological capabilities and the tendency of customers lock-in, and highlight the role of the timing of the intervention as a key factor in determining its effectiveness. Malerba et al. (2008b) consider further policies rather than antitrust, and show how their effectiveness is again strongly influenced by customers lock-in and the arrival of technological discontinuities. Landini and Malerba (2017) study the role of different public policies, including support to entrepreneurship and protectionist measures, in the context of industrial catch-up. Beyond public policy, other general aspects have been investigated adapting existing models: the role of experimental users in the computer industry (Malerba et al., 2007); more nuanced contractual arrangements between computer industry firms (users) and semiconductor components-producers (Malerba and Orsenigo, 2010), and the timing and intensity of entry (Garavaglia et al., 2006; Capone et al., 2013).

All reported models are built around four main building blocks (see Garavaglia, 2010; Li et al., 2018): demand; technological landscape; firm (innovative) behavior; industry dynamics. In the following we illustrate the main aspects characterizing history friendly models in these four dimensions.

- **Demand**

Demand plays a key role in most models, although sometimes is considered completely exogenous as in Kim and Lee (2002). In general, consumers are represented in a more stylized way that business firms. The typical functional form through which their preferences are expressed is the Cobb-Douglas function. All consumers prefer products with lower prices, or higher cheapness, and better quality, or merit of design (Malerba et al., 2016). However, the effect of quality and price on consumers purchasing decision needs not to be linear: consumers may consider a product only above (below) a certain threshold of quality (price), as in Malerba et al. (1999) and Garavaglia et al. (2013). Other factors affecting consumers are marketing or advertising expenditures by the firm (Malerba et al., 1999; Garavaglia et al., 2013), customers lock-ins (Malerba et al., 1999; Malerba et al., 2008; Brenner and Murmann, 2016) and switching costs (Capone et al., 2013; Fontana and Zirulia, 2015)

Although customers are represented in a highly stylized way compared to other approaches, their preferences can be heterogeneous according to multiple aspects: the minimum/maximum thresholds of quality and price (Malerba et al., 2016; Li et al., 2018), the trade-off between quality and price (Malerba et al., 1999; 2008), or the extent to which they can buy also products distant from their favourite one (Capone et al., 2013; Fontana and Zirulia, 2015; Brenner and Murmann, 2016).

- **Technological change**

There is considerably more heterogeneity in the way innovation landscapes are modelled. In some models firms are assumed to improve their current product along a specified trajectory, due to the cumulativeness of technical change (Malerba et al., 1999; 2008; Li et al., 2018), while in others the quality of a product cannot be changed and innovation is about productivity or the introduction of new products (Kim and Lee, 2002; Garavaglia et al., 2013). Some models explicitly represent the patent system (Garavaglia et al., 2013; Brenner and Murman, 2016). New technologies may arrive over time and displace older ones (Malerba et al., 2008) or just coexist with them (Malerba et al., 1999; Fontana and Zirulia, 2015; Li et al., 2018).

- **Firm innovative behaviour**

Firms innovative activities are strongly intertwined with the technological landscape. In general, these activities are subject to stochastic elements, but are also determined by investment decisions, that is by the amount of resources that firms take from their initial funding (Malerba et al., 1999; Garavaglia et al., 2013) or current profits (Malerba et al., 2008; Li et al., 2008). In some models firms are also allowed to imitate their competitors (Kim and Lee, 2002; Garavaglia et al., 2013; Brenner and Murmann, 2016).

- **Industrial dynamics**

Finally, industrial dynamics refers to the processes governing the entry and exit of firms. Some of the models allow entry at specified periods of time, that is at the beginning of the history or when new technologies arrive (Malerba et al., 1999; 20008). In other cases, entry is endogenous to the model and depends on other variables, such as the level of innovative opportunities (Garavaglia et al., 2013), the stock of potential founders (Brenner and Murmann, 2016), long-term expected profits (Kim and Lee, 2002), or the exit of firms (Li et al., 2018). Exit is typically driven by poor performance either in terms of profits (Malerba et al., 1999; Li et al., 2018) or market shares (Malerba et al., 2008; Garavaglia et al., 2013; Brenner and Murmann, 2016). Fontana and Zirulia (2015) consider also the possibility of mergers and acquisitions.

Applications: Entry in the Computer Industry

In this section we briefly present an original application of the Malerba et al. (1999) model to explore the theme of entry and survival[1]. That paper presented a history-friendly model of the evolution of the computer industry, addressing a peculiar challenge raised by its historical development. A dominant firm (IBM) emerged very early and managed to keep the leadership of the market, although experiencing strong competence-destroying technological change, but it was

[1] For reasons of space we do not report here a full descritpion of the model.

not able to contrast the entry of new, small and innovative firms in the new market segment that emerged as a result of the technical change.

The model depicts an industry where at the beginning a single component technology is available (the “transistor”) and later on a new component technology (the “microprocessor”) emerges. The two technologies differ along two dimensions, performance and cheapness, and the microprocessor has an objectively higher technology potential on both of them. These two dimensions provide value for customers: all users prefer faster and less expensive computers, but they differ in the relative importance they give to them. Big firms prefer mainframes because they assign more importance to performance; small users are much more interested in cheapness and therefore prefer personal computers. The two groups of customers also differ in the minimum requirements that a computer must satisfy in both dimensions in order to consider it for purchasing. The transistor technology does not meet the cheapness minimum requirements of the small users. Therefore, this market segment emerges only once the microprocessor technology becomes available.

Within each component technology, there is high cumulativeness of technical change: firms improve both the attributes of their product by investing in R&D laboratories and by learning from experience. Each firm has a specific and invariant technological trajectory that determines the relative weight assigned to improvements in performance vis-à-vis cheapness.

The dynamics of the model is as follows. At the beginning, a number of firms starts R&D spending to explore technological trajectories based on the transistor technology; once they meet the minimum requirements by mainframes customers, they start selling products and use their profits to push technological improvements. Then, the microprocessor technology enters the stage and new firms start their R&D spending using it to target either the mainframes segment or the personal computer segment, although in the former they face competition from incumbents. Finally, also incumbents can adopt the new technology and enter the personal computers segment.

The number of firms that attempted entry at the early stage of the industry was very low, and most of them came from already mature industries such as punched-card machines and electronics. This is reflected in the model, where the number of firms (N) that try to enter the industry by exploiting the initial “transistor” technology is fixed at a low level (6).

In this exercise, entry will still be considered as an exogenous process. However, we introduce the relevant distinction between potential entrants and actual entrants. Potential entrants are those firms that start their R&D activities in an industry: only some of them become actual entrants, and the period between potential entry and actual entry is endogenous to the model. This distinction points to a more general point. Standard theories of entry typically posit the existence of a pool or queue of potential entrants, the origin of which is unspecified and the size of which is assumed to be infinite or again is unspecified (with the notable exception of the flourishing literature on the role of pre-entry experience and spinoffs pioneered by Klepper, 2002). Potential entrants then decide to enter evaluating the prospects for profitability. However, the size of this queue might be crucial for determining the actual amount of entry. If the queue is small, entry will be limited irrespective of any other condition. When the pool is very large, “excessive entry” might take place, as soon as heterogeneity among firms, imperfect information, or uncertainty are considered. Thus, an effort towards beginning to disentangle the role played by the size of the queue and “real” entry might be a worthwhile undertaking.

So, our initial question is whether the presence of a limited number of potential entrants at the early stage of the computer industry was a key determinant of its following structure and dynamics. In principle we would expect no big change, as the effects of the cumulativeness of technical change, the existence of a homogenous demand and the presence of lock-ins for the customers in the mainframe market should not depend on the number of competing firms.

The experiment is run as follows: we study the evolution of the industry for all the values of N from 6 (baseline value) up to 100. Moreover, for each value of N we generate a set of 100 simulations in order to reduce the impact of random elements. Figure 1 shows the effects of these changes on the competitive structure of the mainframes segment.

INSERT FIGURE 1 HERE

In any given period, as N increases the Herfindahl index decreases; however, as the initial number of firms is made larger, the fall in concentration in the mainframe segment associated with the entry of the new microprocessor firms becomes smaller. Last, the Herfindahl index has a steep increase towards the end of the simulation even when N takes its maximum value (100). These results are not surprising: the existence of more potential entrants reduces concentration, but at a decreasing rate. And as the three key determinants of industry structure in this sectoral context (cumulativeness, demand homogeneity, lock-ins) are at work, a dominant firm emerges in all cases. Therefore potential entrants have only a short-run impact on market structure.

While the initial number of firms does not affect much concentration, it does have an influence on who becomes the dominant firm. Figure 2 shows that the share of the first generation firms (those entering with the “transistor” technology) in the mainframe segment decreases abruptly when the number of entrants is quite high, leaving room for the newcomers based on the “microprocessor” technology.

INSERT FIGURE 2 HERE

To dig deeper in this result, we then apply statistical methods to the simulation results by considering each simulation as an observation and building a binary variable that takes the value “1” if the share of the first generation firms in the mainframes segment in the last period is higher than 80% and the value “0” if the share is lower than 20%. Therefore we drop a small number of observations where the share is between 20% and 80% and all simulations in which either no first generation firm or no second generation firm manage to enter the market. The final number of observations is 90.777 out of 95.000: only less than 5% observations are not considered.

INSERT TABLE 1 HERE

Model 1 in Table 1 presents the baseline model having the market success of first generation firms as dependent variable and using only a constant term as a regressor. In Model 2 we include also N (the number of first generation potential entrants). Results are striking: the new model provides only a limited improvement on the baseline, as it can explain only 29% of the cases that were not correctly predicted by Model 1; moreover, the marginal effect associated to N is very low. In sum, while the initial number of firms does not affect much concentration, it would seem to favour the performance of second generation firms. Yet, regression results do not appear to support this conclusion. The issue deserves clearly some further investigation.

Therefore, in addition to the number of entrants, we consider now four more variables that might help to explain the success of incumbents vis-à-vis newcomers. They are:

1. the level of concentration in the industry, as measured by the Herfindahl Index (HERF);
2. the distance of the leader firm from the technological frontier (FRONTIER);
3. the distance of the leader firm from the best technological trajectory (TRAJECTORY);
4. the distance of the leader firm from the technological trajectory of its next follower (FOLLOWER).

The intuition is simple. Lower concentration implies that the rise to dominance of any one firm is slower and in any case the leader faces tougher competition. Similarly, the existence of a close follower weakens the position of the dominant firm vis-à-vis new competitors. Hence, later entrants introducing the new technology may find it easier to challenge the incumbent(s). This will also be the case if the leader is still far away from the technological frontier at the time of the discontinuity: the gap faced by the new entrants is smaller. Last, if the dominant firm happens to follow a relatively inefficient trajectory, there are higher chance that a newcomer might pick up a better direction.

For each of these variables, we consider the distinction between “potential” entrants and “actual” entrants. When the “microprocessor” technology emerges, a number of potential entrants (POTENTIAL or P) start their R&D activity along the new trajectories opened by this new technology. This variable indicates the time of entry and it is exogenously determined by a parameter of the model. Actual entry (ACTUAL or A) occurs instead when a new potential entrant starts selling its product in the mainframe segment: this time variable is endogenously determined within the model and therefore it does change across simulations.

To begin with, we observe that the results indicate that actual entry (A, Model 4) performs better than potential entry (P, Model 3) in predicting the outcomes of the industry. However, potential entry does matter: the full model (Model 5), in which the relevant variables measured at both periods P and E are taken into account, still adds something to the picture.

Second, results clearly indicate that newcomers have more possibilities to displace the incumbents if at the time they start their R&D operations and, more importantly, at the time they start selling their product on the mainframes segment:

1. there is lower concentration in the industry (positive sign of HERF);
2. the incumbent is far from the technological frontier (positive sign of FRONTIER);
3. the incumbent follows a technological trajectory which is not perfectly addressed to the needs of demand (negative sign of TRAJECTORY);
4. the incumbent has a close competitor within the firms that belong to the first generation (positive sign of FOLLOWER).

Thus, the statistical analysis confirm the intuitions. It is worth noting that the independent variables pertain to different domains: the competitive structure of industry (HERF), the competitive dynamics among incumbents (FOLLOWER), the position in the technological environment (FRONTIER) and the position in the demand environment (TRAJECTORY).

This analysis helps also to explain the puzzling results obtained in the previous section. We observed a positive relation between N (the number of first generation potential entrants) and the emergence of a leader among the second generation firms, but this finding was not supported by regression results. The explanation is that in fact at least three of the variables of the previous analysis are affected by N: concentration (HERF) has a negative relationship with the number of firms in an industry; there is a higher probability that a strong competitor emerges (FOLLOWER) if the pool on which it can be selected is larger; finally, tougher competition among many firms slows down and limits the growth of a leading firm (FRONTIER). Even more interestingly, once we take into account these variables the effect of N switches from negative to positive. First, quite simply, since the size of the pool of early entrants is larger, the chances that a first generation firm becomes the leader increase correspondingly. Second, new entrants face more competition from incumbents at the time of their entry: the current leader might be weaker, but there is more diffuse competition. One final remark regards the trajectories followed by successful new entrants. Are they systematically different from those "chosen" by previous leaders? We test whether the mean of the trajectories of leader firms significantly differ according to their generation (Table 2).

INSERT TABLE 2 HERE

First generation leaders are likely to have a trajectory that leads as quickly as possible to the market, in order to lock-in customers, gain profits to increase R&D expenses and therefore sustain further technological advances. This is the trajectory that mirrors the consumer preferences of users of mainframes before the discontinuity and it is identified by the position of the minimum thresholds of performance and cheapness. Is this technological "strategy" still successful for the second generation firms? It should not, because the new microprocessor-based firms cannot exploit the early-mover advantage of first generation firms. In fact, the test gives a negative answer. The technological trajectory that characterizes the new second generation leaders is quite different and it exploits the strong preference of mainframes customer towards performance rather than cheapness.

Applications: Exploration of unknown opportunities

The second application focuses on an extreme scenario where firms explore an unknown space of opportunities. The model used to carry on the following exercises is based on the history-friendly representation of the pharmaceutical industry developed by Malerba and Orsenigo (2002), Garavaglia et al. (2013), and Malerba et al. (2016). The model depicts an industry characterized by market fragmentation, an innovative process in presence of low cumulativeness, and a competitive context where imitation plays an important role. However, the model representation might also apply to many cases where firms operate under deep uncertainty about the properties of the environment, be it an industry, a region, a country; very much in the spirit of Rodrik (2004).

Here is a brief sketch of the model. A number of firms compete to discover, develop and market new products to satisfy a large variety of consumers' needs. They face a space of opportunities that, at the beginning, is largely unexplored. R&D activity is split in two stages: research, i.e. the attempt to discover promising product designs; and development, when these potential products are transformed into actual market products. In the first stage, firms randomly explore the space of products: only a very limited number of product designs has the potential to become an actual product, while the others have "zero quality". There is little scientific knowledge guiding search, and success is a random event. When a firm finds a design that might become a marketable product, the firm patents it. The patent provides protection from imitation for a certain amount of time and over a range of similar designs.

Once a promising design has been found, a firm engages in its development, which requires time and resources. The firm does not know how difficult, time consuming and costly the process will be, or what the actual quality of the new product will be. Development projects sometimes just fail. When development is completed, the firm uses these resources to launch the product in the marketplace. Sales are influenced by the quality of the product, but also by the marketing efforts and by the price that is charged. Pricing is based on a mark-up rule that takes into account the competitive pressure in the submarket.

Firms are represented as strategically heterogeneous, displaying differing propensities towards innovation on the one hand, and towards imitation and marketing on the other. The first successful product offered to satisfy a particular need faces no competition, and the firm may experience a burst of growth. But after some time, other firms may discover and develop competitive products. Moreover, after patent expiration, imitation may occur. As a consequence of the competition from competing or imitative products, the market share and revenue of the original innovator will be eroded away by competitors and imitators.

A key concept in the model is that of a submarket. The term is meant to represent a group of customers that are related in the sense that they share similar or related needs, that can be satisfied by similar or related products. In the model, submarkets are anonymous as well as abstract: they do not have names that correspond to real world analogues. The discovery of a product in a particular submarket does not increase the probability that a firm can discover another product in the same submarket or in a different one, given the firm's innovative effort. However, as it discovers and develops new products, a firm will progressively diversify into new submarkets. Thus, diversification into new submarkets is treated as a random process: firms are always searching everywhere in the space of products. A firm's growth will then depend on the number of products discovered and commercialized, the size and the growth of the submarkets they are present in, the number of competitors and the relative quality and price of the products.

If it has the resources, a firm can choose to invest simultaneously in several parallel projects of research for developing different product it has discovered. If some project needs to be postponed because of lack of resources, the choice of which project to develop depends on the economic value of the submarket and the residual length of patent protection. Initially, firms are endowed with external seed resources to conduct their early research and development activities. Once the discovered products are commercialized, profits coming from the sales of products are reinvested in R&D and marketing. As time passes in the simulated history, there is a broad tendency to diminishing returns to R&D that is attributable to the fact that the space of untapped opportunities is getting depleted by the ongoing discovery process – but the complex details of this evolution are determined endogenously.

In the context of this model, a good metric of social welfare is represented by the number of new submarkets discovered through the innovation process. In fact, until the first product within a submarket is discovered, all consumers within the submarket cannot satisfy their needs. A second metric is represented by firms' profits (the producer surplus): high profits are also positive from a welfare point of view because in the model we assume that a fraction of it is invested in R&D activities. However, high profits are also an indicator of concentration, which can be harmful for social welfare. Therefore, for each exercise we conduct, we also consider the dynamics of concentration as measured by the Herfindahl index.

Benchmark case. The relevance of policy actions depends much on how difficult is for economic agents to operate within the context set by the market forces and the technological environment. If the environment is munificent in terms of innovation opportunities, policy might well become irrelevant. Therefore, as a benchmark for our analysis we use a context in which innovation is quite difficult because of both costs and uncertainty. Figure 3 shows the results of the benchmark case: in Panel A, after an initial period of search, submarkets start to be discovered by firms. Their number increases quite quickly at the beginning, but then the growth rate slows down. The rapid growth at the beginning is due to the entry of many new innovative firms in a relatively short period. Later, the discovery process becomes slower also because innovative firms face the competition by imitators, that reduces their profits. This can also be seen by looking at Panel B: the average firm profit increases steadily at the beginning, and then declines once the entry of imitators reduces margins. Towards the end, profits start increasing again, as an effect of a selection process across firms. Finally, the Herfindahl Index (depicted in Panel C) starts growing steadily quite soon and reaches a relatively high level.

Watering-Can Policy. In this context, the simplest incentive policy that can be designed is just to give money to all firms, according to the watering-can principle. All firms receive an incentive in the form of an amount of money, that must be spent in innovation activities. In the early phase, this money is calibrated to double the level of private investment of the firms. Later, once firms start gaining profits, the incentive is financed by a 20% tax on profits, but this amount is again doubled by the government. This way, in the second phase there is also a shift of resources from imitative to innovative firms. Results are showed in Figure 4 and are quite striking. The number of discovered submarkets (Panel A) shows a steeper growth in the early phase, and by the end of the simulation it reaches a value 60% higher than in the benchmark case. This is also possible because in the early phase, firms get much higher profits, that allow them to finance further innovation activities and to discover more submarkets (Panel B). However, the increase in profits is only temporary: by the end of the simulation, the level of profits is lower than in the benchmark case. This is due to the fact that over all the simulation, the concentration is always much lower than in the benchmark case (Panel C). So, the results of this policy is that we get more innovation, but we also end up in a quite competitive environment.

Timing effects of Watering-Can Policy. A natural follow-up question is whether the timing of the incentives matters. Therefore, we run two more simulations in which this incentive scheme applies either only in the early phase (early life cycle, or ELC) or only in the late phase (late life cycle, or LLC). Results are showed in Figure 5: it is quite evident that this type of incentives is particularly effective in the early phase because they allow more firms to discover more submarkets at the beginning and to use the resulting profits to finance further innovation activities, without harming, but actually sustaining competition. When incentives are given too late, when the growing and exploratory phase of the industry life cycle is already exhausted, we do not observe significant differences with respect to the benchmark case with no incentives.

Merit-Based Incentives. Rather than simply distribute money to all firms, the policymaker might choose to select only some of the firms as recipients of the incentives. We model this policy alternative by changing the probability distribution of innovation activities letting the chances to innovate depend on past innovative performance as measured by the number of discovered innovative products. Therefore, best performing firms in terms of innovation outcomes (discovered and commercialized product designs) will get a higher probability to discover new products, analogously to what would happen if they would receive a monetary subsidy . Results of this exercise are presented in Figure 6, and are compared to the Benchmark and Watering-Can cases. The new policy generates only a limited increase in the number of discovered submarkets (around 10%), and the additional discoveries occur quite late in the life cycle of the industry. This pattern is also associated to a relevant increase in both firms profits and concentration. Overall, the Merit-Based policy performs worse than the simple Watering-Can policy: the main reason is that a Merit-Based policy cannot be properly applied when the industry is young, because it is more difficult to recognize which are the best firms in this period. However, as it was clear from the Timing experiment, the early phase is also the one in which incentives are more important.

Ex-Ante Selection. An even more sophisticated policy might consider the possibility to identify and select the most promising areas for research and to concentrate in those areas the extra-resources provided by the policymaker. We model this exercise by shifting the distribution of search efforts from a Uniform density function to a Pareto-like density function, where the submarkets with the highest probability to be explored are those with the highest number of customers. Because of this change, the probability of selection of a product design satisfying the needs of a particular submarket is not equal for all product designs. So, this policy indirectly operates like a subsidy given to firms for conducting specific research activities. This way, we partially relax the extreme uncertainty assumption and we consider that a high number of customers can be more easily recognised and therefore become the indirect target of a policy, or might even act directly to lobby the policymaker. Results are shown in Figure 7. This policy produces higher profits than the benchmark in the growing stage of the life cycle (Panel B), but they are not used to explore more submarkets as in the Watering-Can case. In fact, Panel A shows that the number of discovered submarkets is 15% lower than the benchmark case, and is less than half the number of discovered submarkets in the Watering-Can exercise. Moreover, there is also no benefit in terms of concentration, which remains stable around the same levels of the benchmark case (Panel C).

The reason behind these results is that the incentive scheme induces firms to concentrate only on a limited set of submarkets. Many areas are subject to a limited innovative effort, whereas the most popular submarkets are subject to strong selection forces that reduce the number of surviving firms and therefore increase concentration.

Timing effects of Ex-Ante Selection Policy. It is interesting to explore a follow-up exercise changing the timing of the incentive scheme, analogously to what we did in the Watering-Can case. Again, we run two more simulations in which the Ex-Ante Selection (EAS) incentive scheme applies either only in the early phase (EAS_ELC) or only in the late phase (EAS_LLC). The basic expectation is that an application of the scheme only in the early phase might improve the outcomes: early selection might help firms to focus on the right areas, and then later on they can use profits to explore more submarkets. The late life cycle scheme, instead, would be less effective in that it does not affect much the behaviour of the firms in the growing stage of the industry life cycle. Results are shown in Figure 8. Limiting the incentive to the early life cycle has a positive effect as compared to the full time scheme, since the number of discovered submarkets increases in the later stages of the simulation. However, this is not enough to match the outcomes of the benchmark case. Firms do not get enough profits from the most populated submarkets, due to the strong competition that occurs in these selected submarkets. As expected the late life cycle policy has a limited effect on the number of discovered submarkets, but it also has a positive effect on concentration: the Herfindahl index increases to levels higher than the Benchmark case due to the indirect protection that is ensured to monopolists or oligopolists in the less explored submarkets.

Imperfect Ex-Ante Selection. An important variant of the Ex-Ante Selection policy would require that the policymaker has not perfect information about the submarkets characteristics, and therefore it might actually select the wrong ones – maybe because groups taking advantage by the development of some submarkets are better able to exercise pressure or gain visibility, or simply because we are back in the extreme uncertainty scenario, but the policymarker is not aware of it. Results presented in Figure 9 show that the outcomes become worse in this situation: since the amount of profits available to firms decreases, the number of discovered submarkets is lower than in the perfect information case. Concentration, however, remains stable at lower levels, because lower profits reduce the possibility for the best firms to discover new products and drive out of the market laggard firms.

Few Submarkets. The analysis conducted so far refers to an industry characterized by a high number (200) of independent submarkets. An interesting question to ask is whether results would differ by changing this aspect of the environmental context. In this exercise, we reduce the number of submarkets to 1/10 of the original value (20), keeping constant the total number of potential products to be discovered and the global size of the industry. In this modified context, we compare the results of the Benchmark case with the Watering-Can policy and the Ex-Ante Selection policy, as defined in the previous sections. Since the number of submarkets is much lower, we also consider a different metric to evaluate results: the number of discovered (innovative) products. Results are presented in Figure 10. Although in the Ex-Ante Selection the dynamics of discovery is slower, in all three cases all submarkets are discovered in the early phase of the industry (Panel A). In terms of products discovered, the best performance is guaranteed by the Ex-Ante Selection: in this case the higher profits earned by firms (Panel B) determine higher investment opportunities that can be translated in the discovery of new products, since there are no neglected submarkets. Analogously to what found in the previous cases, the Ex-Ante Selection policy determines a higher level of concentration (Panel C).

Conclusions

The two cases reported in the previous sections provide interesting insights about the potential of history-friendly models to go beyond the simple task of replicating historical occurrences and providing counterfactual analyses. The application on entry in the computer industry shows that even in a context where cumulativeness and lock-ins made monopoly the only possible outcome, entry can have long-lasting consequences for the selection of the winner firm and technology. Although this might be less interesting for industrial economists, it is a question that could be central in the analysis of strategy scholars interested in the sources of superior performance of dominant firms and also of researchers investigating catching-up by latecomer firms and technologies. Similarly, the application of innovation policies in the model of the pharmaceutical industry generates insightful reflections on how such policies should take into account the role of uncertainty and the nature of the innovation problem at stake.

On a more general level, history-friendly models today still offer many research opportunities, as large historical datasets become available, complexity is more and more recognized as the real challenge to understand economic phenomena, and new insights about how heterogeneous firms and people within them take decisions. A first and obvious opportunity is to enlarge the set of industries analyzed through this method. These may include other industries in which formal R&D plays an important role, such as the aircraft industry or the medical devices industry. However, the work might be extended to more traditional sectors, such as textiles or steel, to study specific historical episodes where major changes occurred in the technologies used. Even more relevant, due to the current trend in economic activities, would be the extension of history friendly models to services, that would require major changes (see Pereira and Dequech, 2015, for a first attempt). Further advancements might come from comparative modelling of different industries on a specific theme (a recent example about spinoffs is the work by Capone et al., 2019) or rather of the same industry across different countries, as in the historical account by Mowery and Nelson (1999). History-friendly models could also provide insights into complex questions related to specific themes. The role of spinoffs in the evolution of industries (Klepper, 2016) might be one of such themes. There is now overwhelming evidence about the important advantages that experienced founders inherit from their parents, but there is still debate on what kind of experience could be more beneficial, and what impact can have on all actors of the industry. The dynamics of industrial catching up by developing countries (Malerba and Nelson, 2012; Lee 2013) might be another one. On this topic, actually, there has been some recent attempt at modelling what happened in specific industries like global mobile phones and semiconductors (Landini et al., 2017; Landini and Malerba, 2017), and mobile communications in China (Li et al., 2018). A third theme is the role country and industry specific institutions. Empirically, the role of institutions has been examined in detail for several industries by the literature on sectoral systems (Malerba, 2002; Malerba and Adams, 2014), but also for the evolution of industrial structures in countries and regions (Boschma and Capone, 2015; Cortinovis et al., 2017). The institutional economics literature can offer relevant case studies and appreciative theories to inform history-friendly models aimed at explaining the role of institutions in industry evolution (Grabner, 2016). Finally, history-friendly models could also be enriched by a deeper consideration of the strategic behavior by different types of firms, that so far has been represented through simple and inertial decision rules. The approach might benefit from integration, at least to some extent, of recent approaches considering the complexity of relations within a firm (e.g. the NK model introduced by Levinthal, 1997), or power and knowledge distribution (Dosi and Marengo, 2016).

REFERENCES

Boschma, R., Capone, G., 2015. Institutions and diversification: Related versus unrelated diversification in a Varieties of Capitalism framework. *Research Policy* 44: 1902–1914.

Brenner, T., Murmann, J.P., 2003. The Use of Simulations in Developing Robust Knowledge about Causal Processes: Methodological Considerations and an Application to Industrial Evolution. *Papers on Economics and Evolution* #0303. Jena, Max Planck Institute.

Brenner, T., Murmann, J.P., 2016. Using simulation experiments to test historical explanations: The development of the German dye industry 1857-1913. *Journal of Evolutionary Economics* 26, 907–932.

Capone, G., Malerba, F., Orsenigo, L., 2013. Are switching costs always effective in creating first-mover advantage? The moderating role of demand and technological regimes. *Long Range Planning* 46, 348–368.

Capone, G., F. Malerba, and L. Orsenigo. 2019. "Spinoffs in context: Entry and performance across industries." *Industrial and Corporate Change*, forthcoming.

Cortinovis, N., J. Xiao, R. Boschma and F.G. van Oort. 2017. "Quality of Government and Social Capital as Drivers of Regional Diversification in Europe." *Journal of Economic Geography* 17: 1179-1208.

Cowan, R., Foray, D., 2002. Evolutionary economics and the counterfactual threat: on the nature and role of counterfactual history as an empirical tool in economics. *Journal of Evolutionary Economics* 12, 539–562.

Cyert, R.M., March, J.G. 1963. *A Behavioral Theory of the Firm.* Prentice Hall, Englewood Cliffs, NJ.

Dawid, H., 2006. Agent-based models of innovation and technological change. In: Tesfatsion, L., Judd, K. (Eds.). *Handbook of Computational Economics. Agent-Based Computational Economics, vol. 2.* Elsevier, Amsterdam, pp. 1235–1272.

Dosi, G., Marengo, L., 2016. The dynamics of organizational structures and performances under diverging distributions of knowledge and different power structures. *Journal of Institutional Economics* 11, 535-559.

Fontana, R., Zirulia, L., 2015. "…then came Cisco, and the rest is history": A history friendly model of the Local Area Networking industry. *Journal of Evolutionary Economics* 25, 875-899.

Garavaglia, C., 2010. Modelling industrial dynamics with "History-friendly" simulations. *Structural Change and Economic Dynamics* 21, 258-75.

Garavaglia, C., Malerba, F., Orsenigo, L., 2006. Entry, market structure and innovation in a history-friendly model of the evolution of the pharmaceutical industry. In: Mazzucato, M., Dosi, G. (Eds.), *Knowledge Accummulation and Industry Evolution. The Case of Pharma-Biotech.* Cambridge University Press, Cambridge, pp. 234–265

Garavaglia, C., Malerba, F., Orsenigo, L., Pezzoni, M., 2013. A simulation model of the evolution of the Pharmaceutical industry: A history friendly model. *Journal of Artificial Societies and Social Simulation* 16(4), 5.

Grabner, C., 2016. Agent-based computational models – a formal heuristic for institutionalist pattern modelling? *Journal of Institutional Economics* 12, 241-261.

Kim, C. W., Lee, K., 2003. Innovation, technological regimes and organizational selection in industry evolution: A 'history friendly model' of the DRAM industry. *Industrial and Corporate Change* 12, 1195-1221; doi:10.1093/icc/12.6.1195.

Klepper, S., 2002. The capabilities of new firms and the evolution of the US automobile industry. *Industrial and Corporate Change* 11, 645-666.

Klepper, S., 2016. *Experimental Capitalism: The Nanoeconomics of American High-Tech Industries*. Princeton University Press, Princeton, NJ.

Landini, F., Lee, K., Malerba, F., 2017. A history-friendly model of the successive changes in industrial leadership and the catch-up by latecomers. *Research Policy* 46, 431-446; doi: 10.1016/j.respol.2016.09.005.

Landini, F., Malerba, F., 2017. Public policy and catching up by developing countries in global industries: A simulation model. *Cambridge Journal of Economics* 41, 927-960.

Lee, K., 2013. *Schumpeterian Analysis of Economic Catch-Up: Knowledge, Path-Creation, and the Middle-Income Trap.* Cambridge University Press, New York.

Levinthal, D., 1997. Adaptation on rugged landscapes. *Management Science* 43, 934-950.

Li, D., Capone, G., Malerba, F., 2018. The "Long March" to catch-up: A history-friendly model of China's mobile communications industry. *Research Policy*, forthcoming; doi: 10.1016/j.respol.2018.10.019.

Malerba, F., 2002. Sectoral systems of innovation and production. *Research Policy* 31, 247-261

Malerba, F., Adams P., 2014. Sectoral systems of innovation, in: Dodgson, M., Gann, D., Phillips, N. (Eds.), *The Oxford Handbook of Innovation Management*. Oxford University Press, Oxford, pp.183–203.

Malerba, F., Nelson, R.R. (Eds.), 2012. *Economic Development as a Learning Process: Variation across Sectoral Systems*. Edward Elgar, Cheltenham, UK.

Malerba, F., Nelson, R.R., Orsenigo, L., Winter, S.G., 1999. 'History-friendly' models of industry evolution: The computer industry. *Industrial and Corporate Change* 8, 3-40.

Malerba, F., Nelson, R.R., Orsenigo, L., Winter, S.G., 2001. Competition and industrial policies in a History-friendly model of the evolution of the computer industry. *International Journal of Industrial Organization* 19, 635–664.

Malerba, F., Nelson, R.R., Orsenigo, L., Winter, S.G., 2007. Demand, innovation, and the dynamics of market structure: The role of experimental users and diverse preferences. *Journal of Evolutionary Economics* 17, 371-399.

Malerba, F., Nelson, R.R., Orsenigo, L., Winter, S.G., 2008a. Vertical integration and disintegration of computer firms: A history-friendly model of the coevolution of the computer and semiconductor industries. *Industrial and Corporate Change* 17, 197-231.

Malerba, F., Nelson, R.R., Orsenigo, L., Winter, S.G., 2008b. Public policies and changing boundaries of firms in a History-friendly model of the co-evolution of the computer and semiconductors industries. *Journal of Economic Behaviour and Organization* 67, 355–380.

Malerba, F., Nelson, R.R., Orsenigo, L., Winter, S.G., 2016. *Innovation and the Evolution of Industries: History Friendly Models*. Cambridge University Press, Cambridge, UK.

Malerba, F., Orsenigo, L., 2002. Innovation and market structure in the dynamics of the pharmaceutical industry and biotechnology: Towards a history-friendly model. *Industrial and Corporate Change* 11, 667–703.

Malerba, F., Orsenigo, L., 2010. User-producer relation, innovation and the evolution of market structures under alternative contractual regimes. *Structural Change and Economic Dynamics* 21, 26-40.

Malerba, F., Yoon, M., 2011. *Modelling the Evolution of Firm Boundaries: A History-Friendly Model of the Emergence of the Fabless Ecosystem.* Working Paper, Milan and Seoul.

Mowery, D.C., Nelson, R.R. (Eds), 1999. *Sources of Industrial Leadership: Studies of Seven Industries.* Cambridge University Press, New York.

Nelson, R.R., Dosi, G., Helfat, C.E., Pyka, A., Winter, S.G., Saviotti, P.P., Lee, K., Malerba, F., Dopfer, K., 2018. *Modern Evolutionary Economics. An Overview.* Cambridge University Press, New York.

Nelson, R.R., Winter, S.G., 1982. *An Evolutionary Theory of Economic Change.* Harvard University Press: Cambridge, MA.

Pereira, M., Dequech, D., 2015. A History-Friendly Model of the Internet Access Market: The Case of Brazil. In: A. Pyka, J. Foster (Eds). The Evolution of Economic and Innovation Systems. Springer, New York, pp. 579-610.

Rodrik, D., 2004. *Industrial Policy for the Twenty-First Century.* Working paper, John F. Kennedy School of Government, Harvard University.

Simon, H., 1955. A behavioral model of rational choice. *Quarterly Journal of Economics* 69, 99–118.

Tesfatsion, L., 2002. Agent-based computational economics: Growing economies from the bottom up. *Artificial Life* 8, 55-82.

Windrum, P., Fagiolo, G., Moneta, A., 2007. Empirical validation of agent-based models: alternatives and prospects. *Journal of Artificial Societies and Simulation* 10(2), 8.

FIGURES

Figure 1. Herfindahl Index in the mainframes segment over time as a function of N.

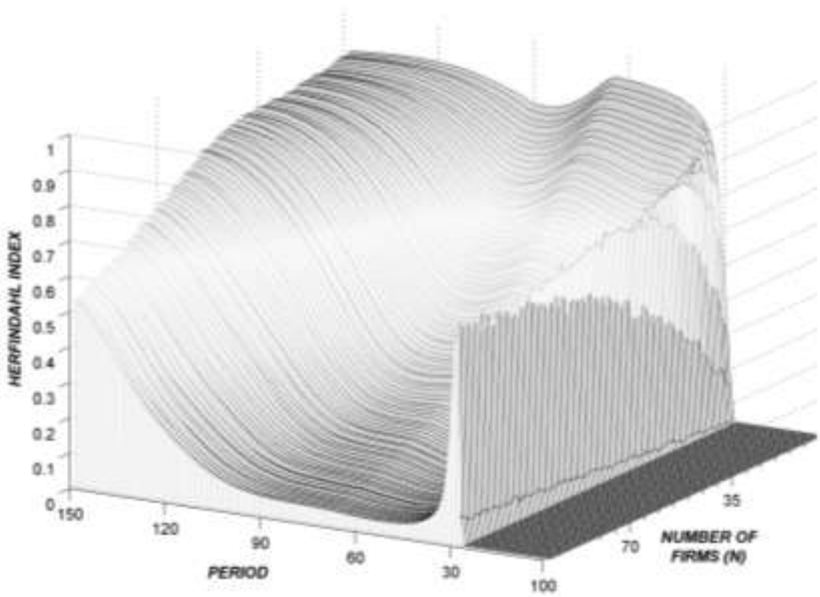


Figure 2. Share of first generation firms in the mainframes segment over time as a function of N.

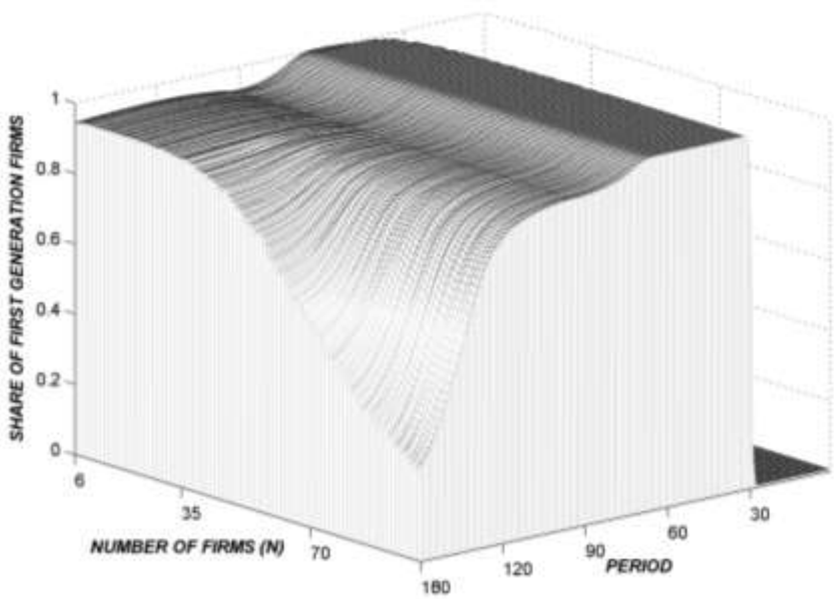


Figure 3. Benchmark case

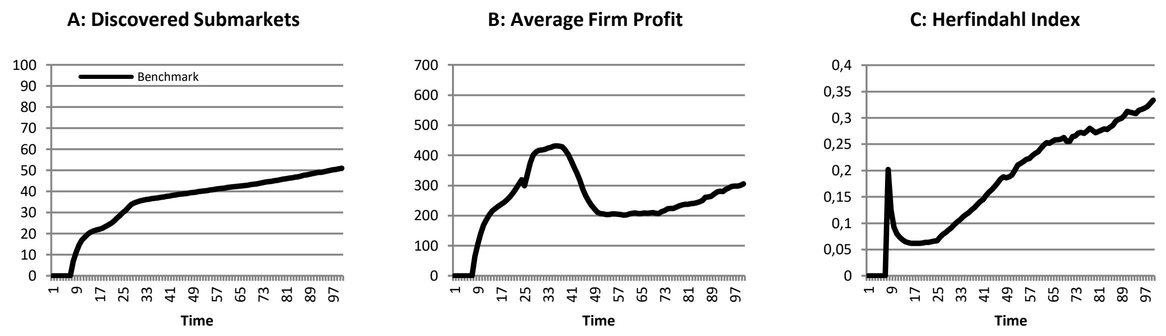


Figure 4. Watering-Can Policy

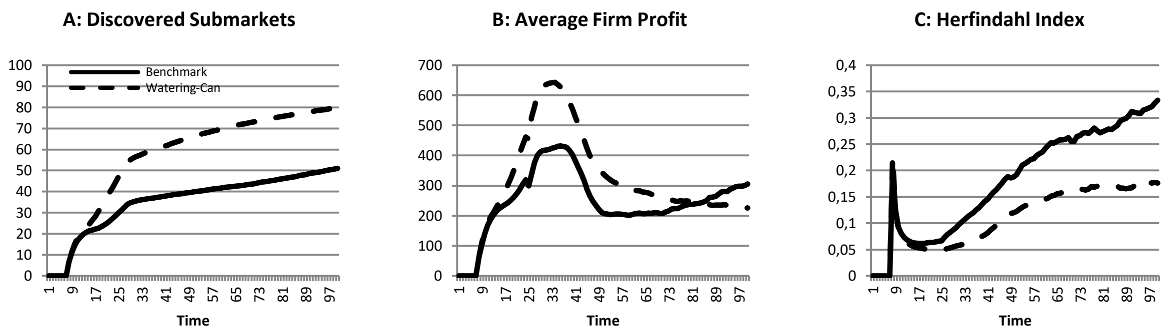


Figure 5. Timing of Watering-Can Policy

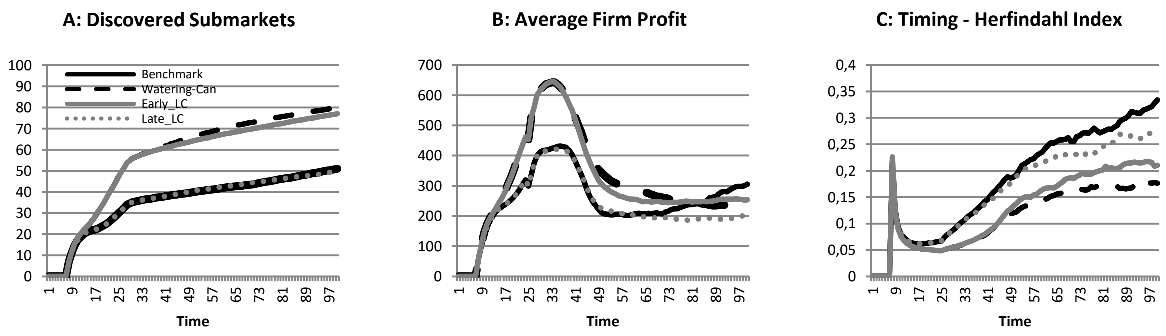


Figure 6. Merit-Based Policy

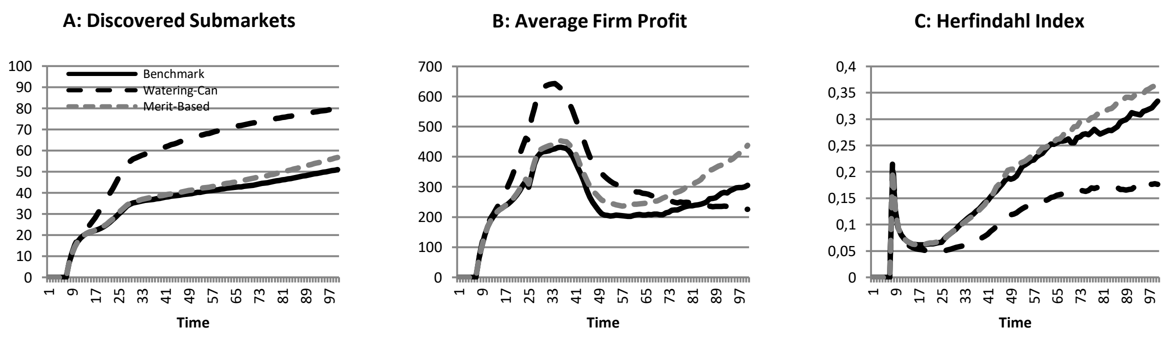


Figure 7. Ex-Ante Selection Policy

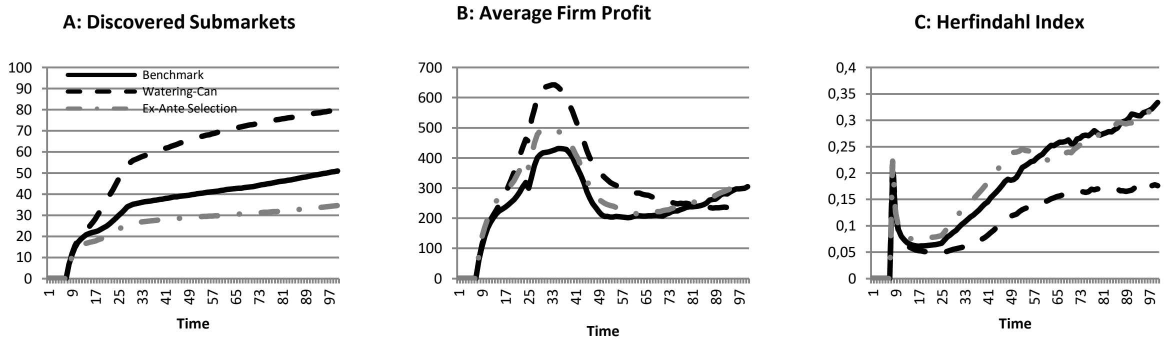


Figure 8. Timing of Ex-Ante Selection Policy

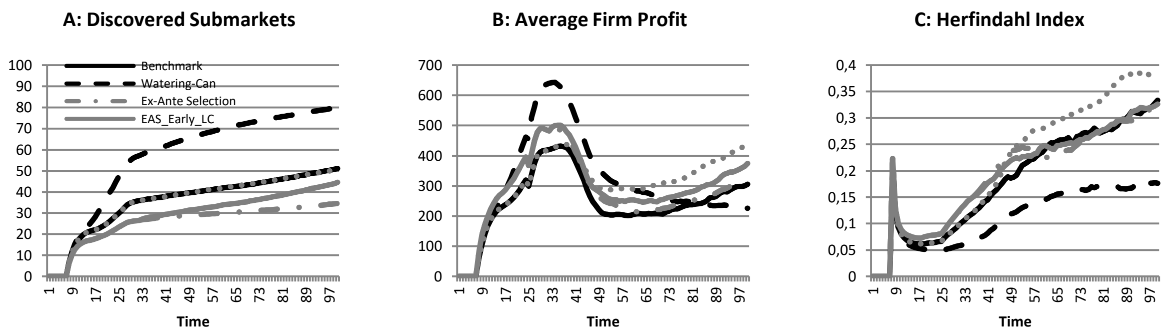


Figure 9. Ex-Ante Imperfect Selection Policy

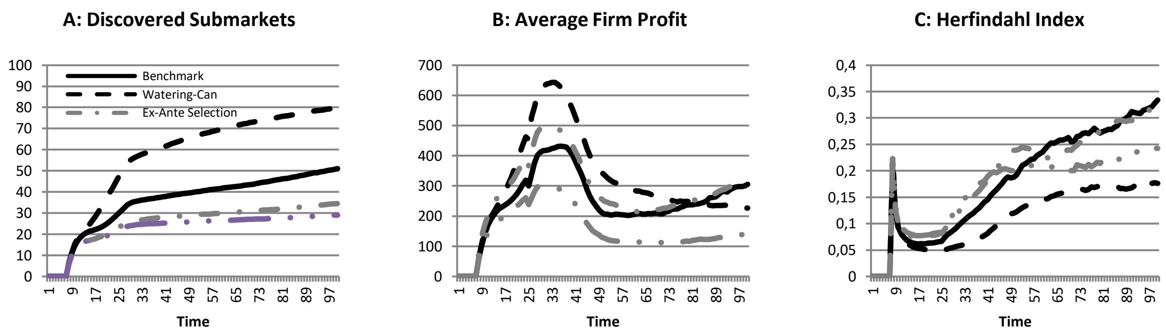
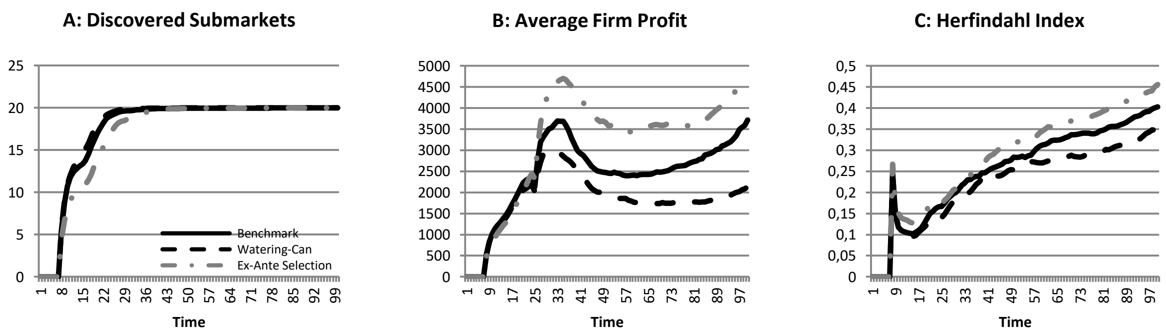


Figure 10. Few Submarkets



TABLES

Table 1. Probability that the mainframes market winner is a first generation firm.

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	0.2129***	0.7008***	-1.7231***	-1.33***	-1.3699***
N		-0.0089***	-0.0039***	0.0021***	0.0013***
HERF(P)			1.4781***		0.3454***
FRONTIER(P)			3.4417***		0.6235***
TRAJECTORYP			-0.7423***		-0.2165***
FOLLOWER(P)			0.1363***		0.037^{**v}
HERF(A)				1.6013***	1.2589***
FRONTIER(A)				1.3093***	0.9972***
TRAJECTORYA				-0.6963***	-0.4419***
FOLLOWER(A)				0.26***	0.1032***
% Correctly Predicted Cases	74.7	82.1	93.5	96.2	96.5
Errors	22924	16285	5898	3467	3183
% Correctly Predicted – Adj	-	29.0	74.3	84.9	86.1
McFadden R^2	-	0.365347	0.717132	0.829435	0.844550
BIC	102607.3	65135.65	29089.56	17567.80	16062.69
Observations	90777	90777	90777	90777	90777

Table 2. Difference in the technological trajectories of the mainframes market winner: first *versus* second generation.

	First Generation	**Second Generation**
Mean	0.822071	0.8754
S.D.	0.0133	0.0474
H_0 (FG = SG)	Rejected; p(0.0001); t(-168.01).	