Laboratory of Economics and Management
Sant'Anna School of Advanced Studies

Via Carducci, 40 - I-56127 PISA (Italy)
Tel. +39-050-883-341 Fax +39-050-883-344
Email: lem@sssup.it Web Page: http://lem.sssup.it

LEM
Working Paper Series

# Endogenous Growth in Open-Ended Economies with Locally Interacting Agents

*Giorgio Fagiolo*[*]

[*] *St.Anna School of Advanced Studies, Pisa, Italy*

**2000/07** **October 2000**

# Endogenous Growth in Open-Ended Economies with Locally Interacting Agents

Giorgio Fagiolo*

Department of Economics
European University Institute
San Domenico di Fiesole, Firenze, Italy

fagiolo@datacomm.iue.it

September 1999

**Abstract**

Building on some general properties of the empirical patterns of technological diffusion and innovation, the paper presents a simple model in which self-sustaining growth endogenously emerges as the result of imperfect coordination among stylized, boundedly-rational, heterogeneous, firms which locally interact in an open-ended technological space and are able to modify the set of their 'nearest neighbors'. We allow the system to be characterized by: (i) endless opportunities of introducing either 'incremental' or 'radical' innovations; (ii) path-dependency in learning achievements; (iii) dynamic increasing returns grounded upon collectively shared 'knowledge bases'. By means of extensive Montecarlo studies, we identify necessary conditions for patterns of persistently fluctuating exponential growth to be generated in the economy. We also investigate causal relationships between system/behavioral parameters tuning the sources of growth and : (a) the overall performance of the economy; (b) the emergence of the exploration-exploitation trade-off; (c) the ability of the system to self-organize and generate GNP time-series of exponential growth with small growth-rates long-run volatility and statistical properties similar to those exhibited by empirical observed time-series. Finally, the effects of behavioral heterogeneity on aggregate outcomes are analyzed and a simple example is presented in which collective economic growth finds its necessary condition in the presence of a share of 'irrational' individuals in the population.

Keywords: Innovation, Endogenous Growth, Local Interactions, Exploration vs. Exploitation
JEL Classification: O30, O31, C15, C22

* An early version of this work is in Dosi, G. and Fagiolo, G. (1998). The paper has been completed while the author was visiting the Department of Economics of the University of Wisconsin (Madison, U.S.A.). Thanks to W.A. Brock and G. Dosi for helpful and stimulating comments. Financial support by the Bank of Italy (Borsa di Studio 'Bonaldo Stringher') is gratefully acknowledged.

# 1 Introduction

In the past few years, the attention of the economic discipline has been brought back to a deeper analysis of the determinants of self-sustained processes of economic growth fueled by technological advances.

On the one hand, both 'Endogenous Growth' and 'Evolutionary' models have been trying to develop formal theories in order to explain how per-capita incomes grow (also) as the outcome of positive feed-backs in knowledge accumulation[1]. On the other hand, a rapidly expanding, empirically-grounded, literature on the economics of technological change has been exploring the drivers of innovation and diffusion - as well as the mechanism through which they occur and their effects - at the levels of firms, sectors and whole Countries[2].

Notwithstanding all that, many scholars have recently spelled out a negative assessment on the extent to which 'neoclassical' and 'endogenous' growth theories have been able to match 'old' and 'new' growth 'stylized facts'[3]. Relatedly, a large body of literature has argued that there still remains an enormous gap between what we historically know about technical change (and its economic exploitation) and the ways we represent them in growth models[4]. According to this view, formal modelers should attempt to take into account some of the following general properties of the empirical patterns of innovation and diffusion, which seem to be indeed neglected by a good deal of contemporary literature.

First, both 'classic' and 'new' aggregate growth models tend to overlook (and in most cases ignore) the systematic heterogeneity in microeconomic technological competencies recently brought into greater prominence by the empirical literature. As a consequence, any 'representative agent' reduction employed in these formalizations might turn out to be highly misleading whenever the aggregate dynamics depends not only on the average characteristics of any population but also on the microeconomic distributions themselves and on the details of the interaction mechanisms among economic micro entities[5].

Second, there appear to be a striking conflict between the incredibly sophisticated forward-looking rationality one typically imputes to agents in aggregate formal stories and the messy experimentation which empirical students of innovation and business history usually find - full of stubborn mistakes, 'animal spirits' and unexpected discoveries[6].

Third, partly as a consequence, it seems quite hard to interpret macro-dynamics as equilibrium paths isomorphic to some underlying 'representative' behavioral pattern[7].

[1] Concerning the former, see Romer (1986, 1990), Grossman and Helpman (1991b). On evolutionary models of growth cf. Nelson and Winter (1982), Silverberg and Verspagen (1995b) and the references therein.

[2] See, among others, Freeman (1982, 1994), Rosenberg (1982, 1994), David (1975), Dosi (1988), Nelson (1995), Lundvall (1993), Granstrand (1994), Stoneman (1995), and fair parts of Dosi et al. (1988) and Foray and Freeman (1992).

[3] Cf. Durlauf and Quah (1998), Klenow and Rodriguez-Clare (1997) and McGrattan and Schmitz (1998).

[4] See Nelson (1998). Cf. Dosi et al. (1994), for an outline of some historical 'stylized facts' which the theory should ideally account for.

[5] On this and related points, cf. Kirman (1989, 1992) and Allen (1988).

[6] For example, on entry dynamics of new firms cf. the evidence discussed in Dosi and Lovallo (1997).

[7] See Silverberg and Verspagen (1995b).

Finally, economic change appears to be driven at least as much by time-consuming diffusion as from innovation[8].

As a result, self-sustained patterns of economic growth should be interpreted as the emergent property of a complex, evolving system composed of many heterogeneous simple entities locally interacting in a high-dimensional, and possibly open-ended, technological space[9]. Following this intuition, the paper presents a computer-simulated model of endogenous growth that builds on the foregoing five properties, together with other 'stylized facts' stemming from empirical analyses of technological change often disregarded in formal aggregate endeavors. The rest of the article is organized as follows.

In Section 2, we shall outline in more detail the building blocks and theoretical conjectures supporting the model described in Section 3. Next, in Sections 4 and 5, we will present an extensive analysis of the simulation results. Section 6 discusses some econometric properties of the simulated time-series, while, in Section 7, a few extensions of the basic model are introduced. Finally, Section 8 draws some conclusions and flags research developments ahead.

## 2 Decentralized Knowledge Accumulation, Interactions and Collective Outcomes

To begin with, let us try to briefly portray some of the key qualitative properties of the process of technological change, as they are singled out by the relevant empirically-grounded literature.

Technological advances, to a significant extent, are generated, *endogenously*, through resource-expensive search undertaken by a multiplicity of profit- motivated agents. Search itself is generally uncertain and innovative entrepreneurs (or, for that matter, incumbent firms undertaking innovative activities) are driven by the beliefs that "there might be something profitable out there", but are generally unable to form probability distributions on the outcomes of their search efforts.

Innovations are not entirely appropriable: knowledge progressively diffuses to other agents who might well catch-up by investing in imitation - most likely, with a lag proportional to some measure of the distance between the knowledge which they master and the competencies they want to acquire.

Knowledge accumulation generally entails dynamic increasing returns both at the levels of individual agents (typically, business firms) and collection of them (i.e. industries), grounded upon collectively shared 'learning paradigms'. However, radically new technologies involve, to different degrees, ruptures and 'mismatching', so that only part of the old knowledge might be useful to the exploitation of future technologies[10].

[8] This point has indeed been emphasized within otherwise rather orthodox models by Jovanovic and Rob (1989), Jovanovic (1995), and, of course, is near the concerns of evolutionary modelers (cf. Nelson and Winter (1982), Silverberg et al. (1988), Metcalfe (1988, 1996)).

[9] Cf. Lane (1993a,b), Krugman (1996), Leijonhufvud (1996) and the articles in Arthur et al. (1997).

[10] On these points, see in particular Rosenberg (1982) and Freeman (1982) regarding technologies uncertainty; cf. Freeman (1982), Levin et al. (1987), Nelson and Winter (1982), and the remarks in Dosi (1997) and Nelson (1997), on appropriability; see Arrow (1962), David (1975, 1988), Romer (1990), Atkinson and Stiglitz (1969), Nelson and Winter (1982), Dosi (1988) and Malerba and Orsenigo (1993) on different - theoretical and empirical - appreciation

Technological search, as well as information diffusion and knowledge accumulation, can be then fruitfully depicted as an interaction process taking place in some (high-dimensional) technological space[11]. Direct interactions among business firms might indeed arise in the economy at many levels. First, entrepreneurs may spread, either intentionally or not, some of the knowledge they master in their 'technological neighborhoods', possibly leading to localized benefits originated by increasing returns to scale and positive network externalities. Second, business firms - as they gradually explore in the vicinity of the existing practices - might be directly affected in their search by other firms currently employing 'similar' technologies. In either situation, because of the highly path-dependent nature of exploitation and innovation activities, interaction structures will display strong non-stationarities[12], since the set of units (firms and technologies) directly affecting imitative and innovative behaviors of every other business firm in the economy are likely to keep changing at each point in time.

Beside all that, the model presented in Section 3 explicitly takes on board four out of the five 'facts' that Paul Romer (1994) identifies as underlying 'New Growth Theories', namely: (i) multiplicity of agents; (ii) non-rivalry in the use of knowledge; (iii) 'replicability' of physical production activities; (iv) endogeneity of discovery efforts. The fifth one - i.e. the rents associated with successful discoveries - is implicitly there but plays no role. On purpose, the expectations on these rents have been partly de-linked from their actual average values (which is implied by the abandonment of any rational technological expectation hypothesis). Hence, while acknowledging that agents search for innovations because they can *sometimes* earn a rent on them, no monotonic relation between the 'true' expected value of those rents and the propensity to innovate is assumed. As a first approximation, the focus will be on the study of the ways patterns of knowledge accumulation, together with institutionally nested 'animal spirits', affect growth - with rent-related incentives just as permissive conditions, above a minimum threshold[13].

While sharing some overlapping with 'New Growth' models - knowledge diffusion in presence of dynamic increasing returns are identified as the primary sources of growth - the model departs from them in a few important respects. First, knowledge is neither treated as entirely appropriable or a pure externality. Rather, its benefits partly accrue to those who embody it and partly leak out as a sort of spill-over. Second, dynamic increasing returns are, at least to some extent, technology specific. Third, diffusion of information takes time rather then being instantaneous. Finally, the radical uncertainty intrinsic in the innovation process involves the possibility that agents make systematic mistakes in innovative search and adoption.

Moreover, well in the spirit of an evolutionary perspective, the model as-

of dynamic increasing returns; cf. Nelson and Winter (1977), Dosi (1982) and Freeman and Perez (1988) on somewhat complementary notions of 'technological paradigms' and relatively ordered 'trajectories' in learning patterns.

[11] More on technological search as a local interaction process in a technological space is in Chiaromonte and Dosi (1993).

[12] Cf. Kirman (1997, 1998) for a discussion on 'evolving networks'. See also Fagiolo (1998).

[13] In fact, this is quite in tune with the empirical evidence. While it is obviously true that with zero appropriability of innovation no private actor has any incentive to undertake expensive search (e.g. for a long time agricultural research on new varieties of seeds, etc.), on the other side, to our knowledge, there is no convincing evidence, either cross-country or over time, that innovative efforts respond smoothly to the fine tuning of appropriability conditions.

sumes: (i) heterogeneity among agents in their technological and behavioral features, e.g. their problem-solving knowledge and their propensity to search and to quickly imitate[14]; (ii) diversity in the knowledge-bases upon which agents are able to draw; (iii) path-dependency in learning achievements; (iv) bounded rationality in both decisions to allocate resources to search and choices on the directions of search efforts[15]; (v) 'open-ended' dynamics in the technology space (so that learning opportunities are notionally unlimited, but what each agent can achieve at any one time is constrained by what one has learned in the past). However, unlike full-fledged evolutionary formalizations, one does not explicitly account for any selection dynamics through which individual agents (*in primis*, firms) grow, shrink or die according to their revealed technological and market success[16]. Hence, the following could be regarded as a reduced form of an evolutionary growth model, focusing upon the collective outcomes of decentralized patterns of knowledge accumulation, while suppressing - like most traditional growth formalizations - any explicit competitive interaction.

On the grounds of these basic building blocks, the following issues are addressed.

*First*, we will attempt to identify some broad circumstances under which the economy is able to solve the trade-off between 'exploitation' of known technologies and 'exploration' of potentially superior ones and generate self-sustaining growth[17]. Notice that, as the model does not rest on any *a priori* commitment to individual rationality and collective equilibria, this question involves an issue which could be called of Schumpeterian coordination, namely: Can 'boundedly rational', heterogeneous agents - locally interacting in some complex technological space - imperfectly coordinate their efforts of search for novel opportunities and of exploitation of what they already know, such as to yield relatively ordered patterns of self-sustained aggregate growth ?

*Second*, extensive experiments of comparative dynamics will be performed so as to map different conditions of generation and diffusion of knowledge into the resulting growth patterns. For example, what happens to the mean (and higher moments) of the distribution of average growth rates across independent sample paths as technological parameters change ? What if one gradually tunes the parameters governing the magnitude of innovative opportunities, the degree of locality in the imitation/diffusion process and/or the amount of path-

[14] Parts of the overwhelming evidence on this point are surveyed in Nelson (1981), Freeman (1982), Dosi (1988).

[15] Hence, unlike stochastic New Growth models of 'creative destruction' - such as Cheng and Dinopoulos (1991) and Aghion and Howitt (1992) - or 'hybrids' between 'old' and 'new' ones - as Jovanovic and Rob (1990) and Jones and Newman (1994) - the analysis is not confined to those rather special cases whereby decentralized agents on average 'get it right'... . On this point, the empirical evidence indeed matches quite solid theoretical reasons on the impossibility of forming unbiased expectations on future technological advances. After all, innovation is about solving problems that one has been unable to solve so far. But if one could know, even in probability, how to solve them, that would mean that the solution algorithm has already been found. The issue bears on problem-solving complexity and, more generally, on the predictability of discovery. More on this is in Dosi and Egidi (1991) and Dosi et al. (1996), within a vast literature.

[16] See Silverberg and Verspagen (1995a, 1995c), Dosi and Nelson (1994). Cf. Section 7.2, however, for an example of how one could indirectly address selection issues in the model.

[17] For a thorough discussion on the exploitation-exploration trade-off arising in adaptive systems see March (1991), Schumpeter (1934), Holland (1975), Allen and McGlade (1986) and Kuran (1988). See also Levinthal and March (1981) and Levitt and March (1988) on the trade-off between the refinement of an existing technology and invention of a new one.

dependence in learning achievements ? Are there any monotone relationships between the size (respectively, the growth rate) of the population and the levels (respectively, the growth rate) of the output ?

*Third*, we shall investigate whether processes of innovation, imitation and diffusion with the above characteristics are capable of generating GNP time-series displaying statistical properties similar to the empirically observed ones. For instance, can one robustly associate different regimes of opportunities, path-dependency and information diffusion to significantly different output growth autocorrelation structures or to some measure of the persistence of the business cycle fluctuations ?

*Finally*, the model seems well suited to analyze how behavioral heterogeneity affects aggregate outcomes. Relatedly, it also highlights a few sources of potential conflict between individual and collective rationality. It is indeed an established result that in the presence of externalities and dynamic increasing returns of some kind, one should not in general expect the dynamics generated by self-seeking agents to correspond with the socially optimal one. Abandoning the 'representative agents' compression of the microeconomics of innovation and allowing for e.g. heterogeneity in the firms' willingness to explore, makes the point even more vividly clear: there is no reason to expect that a decentralized economy would handle the dilemma between 'exploration' of novelty and 'exploitation' of incumbent knowledge the same way as an omniscient (and benign) planner would[18]. Moreover, by relaxing the assumption of hyper-rational agents with correct *technological* expectations, one is also able to consider those circumstances where collective growth finds its *necessary* condition in the presence of a number of 'irrational' entrepreneurs.

# 3 The Model

Think of a *knowledge base* (i.e. a technological paradigm) as a metaphorical 'island' on a stochastic n-dimensional lattice (in the following 2-dimensional for simplicity). Each island is characterized by dynamic increasing returns, associated to knowledge-accumulation, which drive the exploitation of any knowledge base. However, notionally unlimited opportunities exist. Hence, at each point in time, firms might introduce in the system 'radical' innovations, while, as time goes to infinity, whatever economic performance measure may go to infinity as well. Moreover, we assume that individual efforts of 'exploration' slowly yield a collective externality, via, first, diffusion of knowledge, and, second, 'incremental' improvements upon specific knowledge bases[19].

[18] But any actual planner, too, would fall well short of that standard, being equally ignorant of long-run learning opportunities.

[19] The distinction between 'incremental' and 'radical' technical progress (i.e. between paradigm changes and within-paradigm improvements) is increasingly accepted also in other modeling perspectives: cf. for example Cheng and Dinopoulos (1992), Jovanovic and Rob (1990), Amable (1995). Similarly, the issue of a time-consuming (and/or resource-consuming) adaptation and diffusion is beginning to make inroads also into equilibrium growth models: cf. Jovanovic and McDonald (1994), Jovanovic (1995) and Jones and Newman (1994). In the model below we especially emphasize 'creative destruction' aspects of technological discontinuities, with relatively lower attention to the possible complementarities among them (on this point, in the formal growth literature, cf. A. Young (1993)). However, note that the complementarity aspect is implicit in the model's assumption that agents are able to 'carry over', so to speak, part of their previous production skills to new knowledge bases.

Search (i.e. exploration of new islands) and imitation require a resource investment, which, as a first approximation, it is assumed to be equal to the opportunity cost of generating no output. Labor is the only formally accounted input - although one can easily think of a much higher dimensionality of the actual search and production input spaces as ultimately projected into labor productivity dynamics.

In this spirit, the economy is represented as a set of production activities, 'spatially' distributed on the 2-dimensional integer lattice $\aleph^2 = \{1, 2, \dots\}^2$, and it is composed of a fixed population of agents $I = \{1, 2, \dots, N\}$, $N \ll \infty$, and a countable infinite number of islands, indexed by $j \in \aleph$. There is only one good, which can be 'extracted' from every island. Time is discrete and the generic time-period is denoted by $t \in \aleph \cup \{0\}$.

The lattice, i.e. the sea, is endowed by the 'Manhattan' metric $d_1$. Each node $(x, y) \in \aleph^2$ can be either an island or not, while each island has a size of one node. Let $p(x, y)$ be the probability that the node $(x, y) \in \aleph^2$ is an island. We will assume throughout that $p(x, y) = \pi$, all $(x, y) \in \aleph^2$, where $\pi \in (0, 1)$[20].

Each island $j \in \aleph$ is completely characterized by its coordinates $(x_j, y_j)$ in the lattice together with an initial (or intrinsic) 'productivity' coefficient $s_j = s(x_j, y_j) \in \Re_+$.

Without loss of generality, we suppose that, at time $t = 0$, the population is randomly distributed on a (small) set of islands $L_0 = \{1, 2, ..., \ell_0\} \in \aleph$ and that productivity coefficients $s_j$ are uniformly distributed with mean $d_1(j) = d_1[(x_j, y_j)] = x_j + y_j$ , $j \in L_0$, so that, on average, the performance of a 'mine' increases with its distance from the origin of the lattice. All agents are then initially mining inside the smallest box containing islands in $L_0$, i.e. $B_0 = \{(x, y) \in \aleph^2 : x \leq x_0^* \textit{ and } y \leq y_0^*\}$, where $x_0^* = \max\{x_j,\ j \in L_0\}$ and $y_0^* = \max\{y_j,\ j \in L_0\}$[21]. In Fig.1 a very simple example of a conceivable initial configuration of the economy is depicted in order to make clearer the above assumptions.

Finally, assume that each agent $i \in I$ has an exogenously determined willingness to explore defined by the scalar $\epsilon_i \in [0, 1]$.

## 3.1 Dynamics and Endogenous Novelty

Let us turn now to describe how the economy evolves. At time $t = 1, 2, \dots$ each agent can be in one of three different states, namely be a 'miner', an 'explorer' or an 'imitator'. Let $a_{i,t}$ the state of agent $i \in I$ at time $t$, where $a_{i,t} \in \{\text{'}mi\text{'}, \text{'}ex\text{'}, \text{'}im\text{'}\}$, and denote by $j \in \aleph$ the island currently occupied by the 'miner' $i \in I$, i.e. the agent $i \in I$ such that $a_{i,t} = \text{'}mi\text{'}$.

Agents are allowed (with a certain probability) to leave the island they are working on, gradually explore the lattice around and, possibly, discover previously unexploited (and possibly more productive) islands. In order to illustrate how this is capture in a formal way, we need some additional notation.

Denote by $m_t(x_j, y_j)$ the number of miners working on island $j \in \aleph$ at time t. Then, define an island $j \in \aleph$ to be 'known' at time $t$ if $m_\tau(x_j, y_j) > 0$ for at least a $\tau : 0 \leq \tau \leq t$, i.e. if it currently has some people on it or if it was so at

[20] In Section 7.2 this assumption will be however relaxed.

[21] This does not mean, however, that islands $j = 1, 2, \dots$ (both in $L_0$ and in $\aleph - L_0$) are sorted (in some way) by their distance from the origin.

least some finite time in the past. Accordingly, let the set of currently 'known' islands be given by:

$$L_t = \{j \in \aleph : \exists\, 0 \leq \tau \leq t : m_\tau(x_j, y_j) > 0\} \tag{1}$$

Among all known islands, let us call 'colonized' those currently exploited, i.e. all $j \in L_t : m_t(x_j, y_j) > 0$. Conversely, all islands in $\aleph - L_t$ will be called 'unknown', since no agent has previously exploited them. Furthermore, denote the cardinality of $L_t$ by $\ell_t$ and the current location in the lattice of agent $i \in I$ by the pair $(x_{i,t}, y_{i,t})$. Finally, similarly to $L_0$, consider the smallest box containing all islands in $L_t$, i.e. let

$$B_t = \{(x, y) \in \aleph^2 : x \leq x_t^* \textit{ and } y \leq y_t^*\} \tag{2}$$

where $x_t^* = max\{x_j, j \in L_t\}$ and $y_t^* = max\{y_j, j \in L_t\}$. Since the node $b_t^* \equiv (x_t^*, y_t^*)$ will only coincide by chance with a 'known' island, we can think of $b_t^*$ just as a 'proxy' of the most efficient island (i.e. the best practice) currently exploited by the agents[22].

The model allows for an endogenous dynamics on the set $L_t$ and, consequently on the box $B_t$, in the sense that $L_t$ changes through time because of the actions of agents in $I$. A crucial distinction has to be made here between what we will call the 'currently realized' economy and the economy *tout court*. As the box $B_t$ contains all exploited technologies up to time $t$, it therefore represents a proxy of what is actually at disposal of the economy, i.e. the current set of 'fundamentals' or the 'realized economy'.

However, outside $B_t$ there is a whole - eventually better - world waiting to be discovered. The model depicts precisely the process of the gradual endogenous discovery of the economy by the agents themselves. Hence, given the endogenous nature of innovation/imitation activities, it is crucial to account for the process by which agents in different states make 'crucial decisions', i.e. irreversible choices that change forever the economic environment[23]. Let us consider the production process, first.

## 3.2 Production

A 'miner' $i \in I$ currently located on island $j \in L_t$ with co-ordinates $(x_j, y_j)$, will necessarily get, at no cost, a (gross) output $q_{i,t}$ according to the simple production function:

$$q_{i,t} = s(x_j, y_j)[m_t(x_j, y_j)]^{\alpha - 1} \tag{3}$$

where $s(x_j, y_j)$ is the initial 'productivity' coefficient defined above, $m_t(x_j, y_j)$ is the number of 'miners' currently working on island j and $\alpha > 0$. Hence, the current total (gross) output of island $j \in L_t$ and the economy total (gross) output (GNP) will be:

[22] Unlike most neoclassical models, generally based on technical change embodied in different vintages of equipment - cf. Solow (1960) and Kaldor and Mirrlees (1962) - at any given moment in time there is no a single best-practice technique, but many competing technologies located near the frontier of the box $B_t$, see also Silverberg et al. (1988).

[23] See also Shackle (1955) and Davidson (1996) for some hints in a similar spirit.

$$Q_t(x_j, y_j) = s(x_j, y_j)[m_t(x_j, y_j)]^\alpha \tag{4}$$

and

$$Q_t = \sum_{j \in L_t} Q_t(x_j, y_j) \tag{5}$$

In principle, three 'returns to scale' (RTS) regimes can be defined, namely: (i) $0 < \alpha \leq 1$, RTS are individually negative and technologically decreasing; (ii) $1 < \alpha \leq 2$, RTS are individually decreasing but technologically increasing; and, finally, (iii) if $\alpha > 2$, RTS are both individually and technologically increasing. In the following, however, we will mainly focus on technologically increasing RTS, i.e. on cases (ii) and (iii) (cf. also Section 5.1).

### 3.3 Exploration and Innovation

At time $t$, each miner has the opportunity to become 'explorer'. For the sake of simplicity, we will assume here that this happens with probability $\epsilon_i = \epsilon$, for all $i \in I$ which are in the state of 'miner'[24]. As soon as a 'miner' currently working on island $j \in L_t$ decides to become 'explorer' (i.e. $a_{i,t+1}$ = '$ex$'), it leaves its island, 'sailing' around until it finds another one - possibly not known. Notice that, up to now, we have not endowed agents with any 'forecasting' skill. However, when a 'miner' leaves its island at time $\tau$, we let it carry the memory of the last output which the agent was able to get in the state of 'miner' (i.e. its past knowledge and skills), that is $q_{i,\tau}$.

During the search, explorer $i$ does not extract any output and, from time $t+1$ on, moves through the lattice following the 'naïve' stochastic rule:

$$Prob\{(x_{i,t+1}, y_{i,t+1}) = (x, y) | (x_{i,t}, y_{i,t})\} = \tag{6}$$
$$= \begin{cases} \frac{1}{4} & \text{if} \quad |x - x_{i,t}| + |y - y_{i,t}| = 1 \\ 0 & \quad \text{otherwise} \end{cases} \quad \text{all } (x, y) \in \aleph^2$$

that is, at each time period, the 'explorer' moves from its current node $(x_{i,t}, y_{i,t})$ by randomly selecting one out of the four adjacent nodes. This simple behavioral rule implies that, for any $(\xi, \psi) \in \aleph^2$:

$$Prob\{(x_{i,t}, y_{i,t}) = (x_0 \pm \xi, y_0 \pm \psi) | (x_{i,0}, y_{i,0}) = (x_0, y_0)\} = \tag{7}$$
$$= \begin{cases} \binom{t}{\frac{t+d}{2}}\binom{t}{\frac{t+c}{2}} 4^{-t} & \text{if } t \leq d \text{ and } [(t,d) \text{ odd or } (t,d) \text{ even}] \\ 0 & \text{if } t > d \text{ or (t odd and d even) or (d odd and t even)} \end{cases}$$

[24] In Section 7.3 this assumption will be relaxed, allowing for heterogeneous willingness to explore.

where $c = |\xi - \psi| \in \aleph$ and $d = \xi + \psi \in \aleph$. Notice that we are assuming that agents are not aware of the fact that islands are (on average) more and more productive the further away one goes from the origin of the lattice. Indeed, the expected distance from the starting island after $k$ periods of exploration is 0, as by (7): $\mathbf{E}[(x_{i,t+k}, y_{i,t+k})|(x_{i,t}, y_{i,t})] = (x_{i,t}, y_{i,t})$.

The new location of the explorer $(x_{i,t+1}, y_{i,t+1})$ might obviously be: (i) 'sea'; (ii) a 'known' island $j \in L_t$; (iii) a 'new' island $j \in \aleph \backslash L_t$. In the first case, i.e. $(x_{i,t+1}, y_{i,t+1}) \neq (x_j, y_j)$ for all $j \in \aleph$, we still have $a_{i,t+1}$ = '$ex$' and the exploration goes on. In the second case, there will be a $j \in L_t$ such that $(x_{i,t+1}, y_{i,t+1}) = (x_j, y_j)$ and hence the explorer $i \in I$ becomes miner on $j \in L_t$, i.e. $a_{i,t+1}$ = '$mi$'.

The third case is the most important. Suppose, for simplicity, that at time $t$ each explorer is allowed to find new islands only outside the box $B_t$[25]. As stated above, the node occupied by the 'explorer' $i \in I$ at time $t+1$ could be a 'new' island with probability $\pi$. In case of discovery, the new island $j^*$ with co-ordinates $(x_{j^*}, y_{j^*}) = (x_{i,t+1}, y_{i,t+1})$ is added to the set of 'known' islands, i.e. $L_{t+1} = L_t \cup \{j^*\}$ and $\ell_{t+1} = \ell_t + 1$. Moreover, both the set $B_{(.)}$ and the 'best practice' proxy $(x^*_{(.)}, y^*_{(.)})$ are accordingly updated.

## 3.4 Path-Dependence and 'Ordinary' vs. 'Extraordinary' Discoveries

In the model we allow discoveries to be either 'ordinary' or, to different extents, 'extraordinary'. In order to capture the distinction from the innovation literature between innovations within existing knowledge bases and the introduction of radically new 'technological paradigms' (Dosi (1982)), the 'initial' productivity coefficient of a 'new' island $j^*$ discovered by the 'explorer' $i \in I$ carrying the output memory $q_{i,\tau}$, will be given by:

$$s_{j^*} = s(x_{j^*}, y_{j^*}) = (1 + W) \cdot \{d_1[(x_{j^*}, y_{j^*})] + \varphi \cdot q_{i,\tau} + \varpi\} \tag{8}$$

where $d_1[(x_{j^*}, y_{j^*})] = x_{j^*} + y_{j^*}$ is, as usual, the distance of $j^*$ from $(0,0)$; $W$ is a random variable distributed as a Poisson with mean $\lambda > 0$; $\varpi$ is a uniformly distributed random variable, independent of $W$, with mean zero and variance $\sigma_\varpi$ and, finally, $\varphi \in [0,1]$. The interpretation of Eq. (8) is straightforward. The initial productivity of a 'new' island depends on four factors, namely: (i) its distance from the origin (as for initial islands); (ii) a cumulative learning effect directly linked to the past 'skills' of the discoverer, i.e. $\varphi \cdot q_{i,\tau}$; (iii) a random variable $W$ which allows low probability 'jumps', that is, changes in technological paradigms[26]; (iv) a stochastic i.i.d. zero-mean noise $\varpi$.

Two considerations are in order. First, the mechanism through which innovations are introduced in the economy is both path-dependent (see Arthur (1994)) and influenced by random (small) events (cf. Arthur (1989) and David (1992a,b)). On one hand, a large $\varphi$ implies that more skilled 'explorers' (i.e.

[25] This is not a necessary assumption, however. As we will see below, the economy is naturally driven, although only on average, toward more efficient islands by the process of diffusion of information, so that the event of finding a new island inside $L_t$ is in fact irrelevant in our description.

[26] As happens in Nelson and Winter (1982) or Silverberg and Verspagen (1994), innovation is a local process.

more efficient past 'miners') are likely to discover more productive islands and to produce more in the future, thanks to a sort of 'learning-to-learn' mechanism (cf. Stiglitz (1987)). Moreover, the stochastic nature of innovation, together with increasing returns associated with learning by doing (as in Arrow (1962) and Parente (1994)), allow even 'ordinary' discoveries to drive the process of growth. Second, notice that, as by independence:

$$\mathbf{E}s(x_{j^*}, y_{j^*}) = (1+\lambda)[(x_{j^*} + y_{j^*}) + \varphi \cdot q_{i,\tau}] \tag{9}$$

then, on average, a larger $\lambda$ lets 'extraordinary' discoveries to be more likely in the economy. The parameter $\lambda$, together with $\pi$, are measures of the degree of notional 'opportunities'. Indeed, a large $\lambda$ lets, in expectation, the productivity of a newly discovered island to be sensibly larger than those associated to the currently 'known' islands; likewise, a larger $\pi$ implies a larger average number of per-period discoveries.

### 3.5 Interactions, Diffusion of Knowledge and Imitation

Due to the uncertainty of the exploration process and to within-island dynamic increasing returns, the system provides an incentive for both 'miners' and 'explorers' to imitate the most productive islands existing in the 'currently realized' economy. In the model we formalize a process of local diffusion of knowledge which tries to capture some basic features of empirically observed patterns of imitation and diffusion[27].

Let $m_t$ be the number of 'miners' currently present in the economy. At time $t$, the agents mining on each 'colonized' island $j \in L_t$ deliver a signal which is instantaneously spread all around. Signals are characterized by an *intrinsic intensity* proportional to the share of miners present on $j \in L_t$ - i.e. $m_t(x_j, y_j)/m_t$ - and a *content* given by the actual productivity of the island - i.e. $Q_t(x_j, y_j)/m_t(x_j, y_j)$. Moreover, they decay exponentially with the distance from the source, so that the *actual intensity* with which a signal delivered from $(x_j, y_j)$ reaches an agent currently located at $(x, y)$ is given by:

$$w_t(x_j, y_j; x, y) = \frac{m_t(x_j, y_j)}{m_t} exp\{-\rho[|x - x_j| + |y - y_j|]\}, \quad \rho \geq 0 \tag{10}$$

A signal delivered at $(x_j, y_j)$ will be *received* by agent i located at $(x, y)$ with a probability proportional to $w_t(x_j, y_j; x, y)$. Agent $i$ will then collect the 'contents' of all *received* signals (those coming from, say, islands whose indices are $\{j_1, \ldots, j_{\widetilde{\ell}}\} \subseteq L_t$, where $\widetilde{\ell} \leq \ell_t$ is a random variable) and contrast them with *its own performance.* The latter is simply agent $i$'s current productivity if it is a 'miner' (say on island $j$), or the 'memory' on the productivity of its island of origin (say, $j$), if it is an 'explorer'. Hence, it will choose among the $\widetilde{\ell} + 1$ available options by drawing an element from the set $\{j, j_1, \ldots, j_{\widetilde{\ell}}\} \subseteq L_t$, with probabilities proportional to the associate productivities. If the choice is $j$, then

[27] See for instance Nelson and Winter (1982), David (1975); Dosi (1988, 1991), Freeman (1994). Cf. also, along the same lines, Jovanovic and Rob (1989) and Jovanovic and McDonald (1994).

it will decide not to imitate any island but rather to remain in the current state. Otherwise, it will become an 'imitator' - i.e. $a_{i,t+1} =$ '$im$' - and it will move toward the imitated island, say $(x\prime, y\prime)$, reaching it after $k = d_1[(x\prime, y\prime); (x, y)] = |x - x\prime| + |y - y\prime|$ time periods - i.e. making one step per period and following the shortest path[28]. Finally, once the imitated island is reached, it will turn again its state into 'miner', i.e. $a_{i,t+k+1} =$ '$mi$'.

The processes of production and knowledge diffusion govern interactions in the model[29]. First, agents locally interact in an indirect and deterministic way in the mining process. The parameter $\alpha$, by tuning dynamic increasing returns to scale at the island level, also measures the strength of the incentives to converge on the same technologies - i.e. conformist effects. Second, firms interact in a direct and local way by means of the knowledge diffusion process: on average, firms' behaviors are directly affected by the stochastic signals coming from those agents which, on average, are currently closest in the technological space (i.e. the lattice). Here, the parameter $\rho \geq 0$ tunes the 'degree of locality' of the interactions: the larger $\rho$, the more the process of diffusion of knowledge is local, since signals will tend to reach, in probability, only the 'nearest neighbors'. Two extreme cases are: (i) $\rho$=0, i.e. interactions are global, as information diffusion does not depend on the distance between source and receiver; and (ii) $\rho = \infty$, i.e. no signals are spread and interactions are shut down.

Some considerations are worth noting. The model, as in many other 'local interaction' setups, assumes that agents' decisions (e.g. whether to imitate or not, which technology to adopt, etc.) are affected by some function of the past choices of their nearest neighbors (appropriately weighted). The kind of dependence postulated is a stochastic one, in order to capture all those frameworks where either information is imperfectly conducive or the agents themselves, faced with a complex environment, possess no adequate skills to correctly master the information they hold. In our model, unlike the majority of formalizations in the literature, the interaction structure - albeit rudimentary - does change in time. In fact, as agents gradually explore the technological space and move across the lattice, they are likely to modify both the cardinality and the composition of the set of their 'neighbors', allowing for an evolving interaction network.

## 3.6 System Variables, Timing and Implementation

At each time period $t = 0, 1, 2, \ldots$ , the economy is completely characterized by the following *micro variables*. At the technologies (i.e. islands) level: (a) the set of 'known' islands $L_t$; (b) their co-ordinates $\{(x_j, y_j), j \in L_t\}$; (c) the initial productivity coefficients $S_t = \{s_j, j \in L_t\}$. Concerning agents, we define the mappings:

$$A_t : I \rightarrow \{\text{`}mi\text{'}, \text{`}ex\text{'}, \text{`}im\text{'}\}$$
$$C_t : I \rightarrow \aleph^2$$
$$O_t : I \rightarrow \Re_+$$

[28] For the sake of simplicity, notice that an imitator cannot be reached by any other signal while committed to a particular destination.

[29] For a thorough discussion on local interaction models in economics, see Kirman (1998) and Fagiolo (1998).

recording agents' current states, coordinates and individual outputs.

The *macro variables* of interest are: (i) the triple $(m_t, e_t, i_t) \in \aleph^3$, $m_t + e_t + i_t = N$, i.e. the current number of 'miners', 'explorers' and 'imitators' in the economy; (ii) the pair $(\ell_t, \ell_t^C) \in \aleph^2$ (where $\ell_t$ is the number of currently *known* islands and $\ell_t^C \leq \ell_t$ is the number of the *colonized* ones), together with their coordinates and their initial productivity; (iii) the log of GNP, namely $q_t = \log Q_t$.

The timing of decisions and events occurring in a generic iteration (i.e. in the time interval $(t-1, t]$ ) runs as follows. First, given macro and micro system variables at time $t-1$, agents take their decisions: miners update output and choose whether to start searching; explorers select the next portion of the lattice to explore (and, possibly, they find a new island); imitators continue to approach the technologies they have chosen to adopt. Third, interactions take place through information diffusion. Finally, all time-$t$ system variables are accordingly updated and the next iteration starts.

The model is an example of a so-called 'artificial economy'[30]. Unless the focus is not on particular stationary cases - as in Section 4.1 below - one is bound to analyze its main properties by resorting to computer simulations. Analytical solutions are not indeed achievable for the full-fledged form, because of the underlying complication of the stochastic processes updating micro - and accordingly macro - system variables. In Appendix 1, the pseudo-code of the implemented computer program is reported, together with the list of all relevant variables[31].

In the next sections we will present an extensive analysis of simulation results. The focus will be on the aggregate properties of the simulated time-series of the log of GNP, i.e. $q(\omega) = \{\log Q_t,\ t = 0, 1, ..., T; \omega\}$, where $\omega$ is a point in the parameter space $\Omega$, that is:

$$\omega \in \Omega \equiv \{(\rho, \varphi, \lambda, \pi, \alpha, \epsilon, N, T) \in \Re_+^3 \times [0,1]^3 \times \aleph^2\} \tag{11}$$

To begin with, we will analyze how the model behaves in some 'benchmark' parametrizations, in order to assess the role played by knowledge-specific increasing returns, imitation and exploration in the dynamics of the economy. In particular, we will start by addressing the question of whether the model is able to display self-organized patterns of persistent growth and - if so - under which behavioral and system parametrizations (especially concerning the degree of 'open-endedness' of the economy, as well as innovation and diffusion rates).

[30] Cf. Lane (1993a,b). The choice to label as 'artificial' all economic models implemented and run on a computer could be however criticized, at least if the term is employed with its common meaning (something 'not genuine'). If, on the contrary, one interprets the label 'artificial' in another of its acceptation (i.e. 'made in imitation of something existing in nature'), then it simply boils down to the very definition of an (economic) 'model'. The same argument also applies to the so-called 'artificially adaptive agents' models, cf. for example Axtell and Epstein (1994), Holland and Miller (1991), Arthur (1993).

[31] It is not possible, of course, to provide here an overview of methodological (and philosophical) issues involved in computer simulated economic models, as well as a discussion on the relative merits of simulated and analytically solvable models. On these topics, see Nelson and Winter (1977), Lane (1993a,b), Silverberg and Verspagen (1995b) and Leijonhufvud (1996).

# 4 The Emergence of Self-Sustaining Growth : Closed vs. Open-Ended Economies

A key feature of the model resides in its ability to allow for an endogenous evolution of the set of fundamentals of the economy. But, in the first place, what would happen if one bounds, to some extent, the dynamics governing the progressive enlargement of the technological frontier ? Put it differently, is the economy able to generate patterns of self-sustaining aggregate growth if one considers stationary environments where agents behave on the grounds of an unchangeable set of fundamentals ?

In order to answer these questions, we will start by analyzing the benchmark case of a closed economy (i.e. one in which $L_t \equiv L, \forall t$). This will be done by focussing on two distinct setups. In the first one, it will be assumed no possibility of exploration (and hence innovation) whatsoever, i.e. $\epsilon_i = \epsilon \equiv 0, \forall i \in I$. Therefore, agents will only be supposed to exchange information about an initial, fixed set of technologies and eventually exploit them, but they will not be able to endogenously introduce innovations in the system. In the second setup, exploration (as well as information diffusion) will be permitted, but only inside the initial 'realized economy'. In other words, miners can become explorers with some probability $\epsilon_i = \epsilon > 0$, but they are only able to 'sail' within the box $B_0$.

Then, we will turn our attention back to the more general case of 'open-ended' economies in which agents can to introduce innovations in the system.

## 4.1 A Closed Economy without Exploration

When $\epsilon_i = \epsilon \equiv 0$, all $i \in I$, the model allows an analytical treatment. Without loss of generality, assume that the economy is composed of two islands, i.e. $|L_0| = \ell_0 = 2$. In this case we can neglect any spatial consideration by assuming that the productivity coefficients $(s_1, s_2) \in \Re^2_+$ also represent the technological distance between islands[32]. More precisely, let $(s_1, s_2) = (1, s), s \in \aleph$, and suppose that if a miner working on island $j \in \{1, 2\}$ at the beginning of time $t - s$ decides to imitate island $j\prime \neq j$, then he will reach $j\prime$ at the end of time $t - 1$ and start producing at time $t$[33].

In this simple setting, island 2 plays the role of the 'best practice' for $s \geq 2$, while the case $s = 1$ depicts the benchmark case of homogeneous technologies[34]. Here, the dynamics of the economy is entirely driven by the process of information diffusion (see Section 3.5) until one out of the two technologies, say $j$, manages to capture all $N$ agents. In that case, the economy would lock-in at the steady state where total output is $Q^* = s_j N^\alpha$. As intuition suggests, however, path-dependency entailed by increasing returns will tend to drive all agents, through waves of imitation, toward the island with the *actual* (not *initial*) best

[32] When exploration is shut down and only information diffusion drives the dynamics of the economy, the two-dimensional lattice structure - which provides an additional underlying source of technological distances - is indeed redundant. This is not true, however, in the full-fledged model.

[33] In other words, the parameter $s$ represents the number of time periods an imitator needs to get to the adopted technology, while the actual productivity distance between islands is $\Delta = s - 1$.

[34] Notice that when $s = 1$, the information diffusion rate $\rho$ plays no role, as the information is spread and received instantaneously. Moreover, no imitators are present in the economy.

productivity, therefore implying non-ergodicity in the stochastic process governing output evolution and, consequently, potential inefficiency.

More formally, it can be proven the following:

**Proposition**

Consider the stochastic process $\underline{M}_t = \{(M_{1t}, M_{2t}), t \geq 1\}$, where $M_{jt}$ is the random variable: "number of agents 'mining' on island $j$ at time $t$", $j \in \{1, 2\}$. Given any initial distribution $P_0$ on $M_{10}$ with support $\{0, 1, \ldots, N-1, N\}$, $M_{20} = N - M_{10}$, then, for any $(\alpha, \rho) \in \Re^2_+$:

1. $\underline{M}_t$ is a Markov process with discrete state space $\{(m_{1t}, m_{2t}) \in \{0, 1, \ldots, N\}^2 : m_{1t} + m_{2t} \leq N\}$;
2. If $s = 1$ then $\underline{M}_t$ is a stationary, not irreducible, Markov chain displaying two absorbing states $\underline{m}_+ = (0, N)$ and $\underline{m}_- = (N, 0)$;
3. If $s \geq 2$ then $\underline{M}_t$ is a non-stationary Markov process with two absorbing states $\underline{m}_+ = (0, N)$ and $\underline{m}_- = (N, 0)$.

   **Proof.** See Appendix 2. ■

The above Proposition simply states that, for any $s \in \aleph$, the economy will be ultimately absorbed either in the efficient outcome $Q^*_+ = sN^\alpha$ or in the inefficient one $Q^*_- = N^\alpha \leq sN^\alpha$. In order to characterize the probabilities of absorption in either state, let us consider some deterministic initial condition lying in the transient set, i.e. $m_{10} = m_0 \in \{1, \ldots, N-1\}$, $m_{20} = N - m_{10}$. It is then of interest trying to assess how the probability of an inefficient lock-in depends on the technological gap between the island 1 and 2 (i.e. the parameter $s$), on the strength of increasing returns ($\alpha$), on the speed of information diffusion ($-\rho$) and, of course, on the number of miners initially adopting the inefficient technology ($m_0$). We have:

**Result**

Let the absorption probability in island 1 (i.e. the inefficient limit state if $s \geq 2$) to be defined as:

$$p^s_-(m_0; \alpha, \rho) =$$
$$= Prob\{\exists t^* > 1 : \forall t \geq t^*, \underline{M}_t = (N, 0) | \underline{M}_0 = (m_0, N - m_0); \alpha, \rho; s\}$$

for any $m_{10} = m_0 \in \{1, \ldots, N-1\}$. Then, everything else being constant[35]:

1. $p^s_-(m_0; \alpha, \rho)$ are non-increasing in $s$ and non-decreasing in $m_0$;
2. If $s = 1$, then $p^s_-(m_0; \alpha, \rho)$ are non-increasing in $\alpha$ if $m_0 \leq N/2$ and non-decreasing in $\alpha$ when $m_0 \geq N/2$; if $s \geq 2$, then $p^s_-(m_0; \alpha, \rho)$ are always non-decreasing in $\alpha$.

[35] Unfortunately, absorption probabilities cannot be computed explicitly (see Appendix 2). However, when $s = 1$ one can tabulate them by means of numerical analyses. Montecarlo techniques must instead be used when $s \geq 2$ to estimate both absorption probabilities and average absorption times.

3. $p^s_-(m_0;\alpha,\rho)$ are non-increasing in $\rho$ when $s \geq 2$ and the magnitude of this effect increases with $s$.

In order to support the claim above, Figure 2 shows some examples of the estimation of $p^s_-(m_0;\alpha,\rho)$ for $s \in \{1,2,3\}$, as $(m_0;\alpha,\rho)$ vary in the relevant parameter space, cf. Panels (a) to (c). As one could have expected, when the initial number of 'inefficient' adopters is below a certain threshold (itself increasing with the strength of returns to scale and the technological gap), then the system will inevitably converge to the efficient outcome no matter how large are the incentives to stick with the initial choice. However, when $m_0$ goes through that threshold, the probability of ending up in the inefficient state becomes strictly positive and grows as the incentives to knowledge accumulation increase. In the limit, when only a few miners are initially aware of the superior technology and returns to scale are increasing ($\alpha > 1$), the probability that the system is absorbed by island 2 converges to zero[36]. Finally, when $s \geq 2$, the more information diffusion is local (i.e. the greater $\rho$), the less the average number of miners which leave their islands in each time period and, consequently, the less likely the event that waves of imitation would trigger a migration from the efficient technique toward the inefficient one (see Panel (d)). Therefore, for a given $(m_0,\alpha)$, the probability of being absorbed in island 1 will decrease with $\rho$ (increasingly fast as $s$ grows).

The system parameters accordingly affect the average time of absorption, see Figure 3. When technologies are homogeneous, lock-in time is decreasing with $m_0$ no matter how large are returns to scale - Panel (a) - while, when $s=2$, two regimes appear: if $\alpha$ is small, the length of the transition dynamics leading to absorption is only retarded as $m_0$ grows; conversely, when returns to scale keep increasing, the relationship becomes unimodal, with the maximum shifting to the left as $\alpha$ goes up - see Panel (b). Moreover, when $s=2$, a smaller speed of information diffusion (larger $\rho$) entails larger absorption time, as expected. Analog patterns arise in the case $s=3$, cf. Panels (c) to (f).

Further light on the ways in which the relevant parameters influence the dynamics of the system can be shed by analyzing the behavior of the time evolution of the conditional expectation of the number of miners currently working on islands 1 and 2, before lock-in occurs. To that end, in Appendix 2 we study the system of difference equations governing the evolution of such paths, i.e. $\{\mathbf{E}_t(\underline{M}_t|\underline{M}_{t-1},\underline{M}_{t-2},\ldots;m_0),\ t\geq 1\}$.

However, non-ergodicity of the stochastic process $\underline{M}_t$ implies that the actual dynamics of the system is entirely determined by early random events characterizing the process - cf. David (1992b) - that is by early, unpredictable, waves of imitation. As a consequence, GNP time-series and growth rates look like those depicted in Figure 4. In this simple setting, growth is a transitory phenomenon: once the lock-in is achieved, no fluctuations will arise thereafter.

## 4.2 A Closed Economy with Exploration

Suppose now that exploration is allowed, i.e. let $\epsilon_i = \epsilon > 0$, $\forall i \in I$, but only within an unchanged set of 'knowledge bases'. That is to say, we assume that

[36] Notice also that, when technologies are homogeneous, absorption probabilities are symmetric around $m_0 = N/2$ for a given $\alpha$; moreover, as $\alpha$ increases, $p^1_-(m_0;\alpha)$ tends to a step-function whose value at $m_0 = N/2$ is undetermined. See Appendix 2.

once an initial set of islands $L_0$ is drawn, the probability of finding new islands is zero, i.e. $\pi = 0$, and that explorers are allowed to search only inside the fixed 'realized economy' defined by the box $B_0$.

In this setup, agents are still not able to 'innovate' (i.e. to discover islands other than the already 'known' ones) and must necessarily exploit the existing technologies. However, unlike the previous case, they can always decide to leave the island they are working on, even though all agents are mining on it.

All that introduces a potential source of 'irrationality' and 'idiosyncrasy' in individual behaviors. Although the decision to become explorer is not linked - in this version of the model - to any system variable, it is tempting to define this behavior as a 'nonconformist' one, as in a few models of 'social interaction' and 'herd behavior'[37]. Indeed, when exploration is allowed, the lock-in of the system will not generally occur, since there is always a positive probability that 'non conformist' decisions will induce phase-transitions in the system.

In a two-islands setup, the economy is characterized by the Markov process $\underline{M}_t$ (as before), together with the stochastic process describing the current number of explorers. However, unlike the previous case, transition probabilities are not only influenced by the propensities to imitate technologies with a higher revealed efficiencies, but also involve a certain probability of 'exploring'. Islands represent here 'basins of attraction' among which the system continually oscillates[38].

The stochastic process of exploration/imitation yields persistent output fluctuations but only transitory growth. Indeed, as depicted in Fig.5, the simulated time-series of GNP display a stationary autoregressive pattern - as econometric analyses (not reported) usually show. Over finite time periods, increasing returns and knowledge diffusion induce agents (on average) to move toward currently more efficient islands - cf. Figures 6(a) and 6(b) for the two cases $s_1 = s_2$ and $s_1 < s_2$. However, exploration allows with positive probability 'de-locking' bursts, also toward notionally less efficient islands. In a sense, persistent fluctuations are in this case generated by a problem of *imperfect Schumpeterian coordination* in presence of dynamic increasing returns to learning[39].

### 4.3 Exploring in an Open-Ended Economy: Some Qualitative Results

In both stationary environment analyzed so far, self-sustaining growth could emerge only if one superimposes an exogenous Solow-like drift on the best-practice production function. Otherwise, as long as agents behave on the grounds of fixed fundamentals, economic growth is a temporary phenomenon.

Consider now the more general case where $\epsilon_i = \epsilon > 0$, $\forall i \in I$ and the economy is open-ended. In this full-fledged setup, firms are able to endogenously induce a drift in the technological frontier, in ways that are both path-dependent and imitation-driven.

It turns out that the economy exhibits, for a wide range of parameters, patterns of self-sustaining growth. Typically, the simulated time-series of GNP

[37] See Kirman (1993), Brock and Durlauf (1995), Hirschleifer (1993).

[38] These properties are quite similar to those displayed by models based on Fokker-Planck equations. Cf. also Kirman (1993) and Orléan (1992).

[39] Notice here the loose analogy with the coordination-related dynamics treated by Cooper and John (1988) and Durlauf (1994).

are exponentially shaped, so that its natural logarithm displays a linear trend, as in Figure 7[40]. However, exponential growth is not the sole rough regularity in output patterns that open-ended economies are able to generate. Indeed, in other regions of the parameter space, the model can mimic either 'mild growth' economies, whereby the GNP time-series fluctuate around an *S-shaped* trend and long-run growth rates tend to become a mean-zero stationary process - cf. Figure 8(a); or 'no growth' ones - see Figure 8(b) - as in the stationary setup with exploration discussed in Section 4.2.

Under parametrizations yielding self-sustaining growth[41], the model exhibits some additional interesting regularities. First, the time series of the number of 'miners', 'explorers' and 'imitators' can typically be described as stationary processes, see Figure 9.

Second, although the number of currently 'known' islands displays a linear time trend, both the ratio between 'colonized' and 'known' islands, and the number of 'colonized' ones - Figure 10(a) and (b) respectively - fall quickly and then follow a stationary process. Hence, starting from a fairly uniform distribution of $N$ agents on the initial set of 'known' islands $L_0$, diffusion of knowledge is likely to drive agents to concentrate on (i.e. 'colonize') a relatively small cluster of 'known' islands which, by dynamic increasing returns, might be, often but not always, the most efficient ones. In general, the expected size of the clusters of colonized islands is likely to increase as returns to scale and path-dependency get smaller; the probability of finding an island increases; agents are more willing to explore; and/or information diffuses more locally.

Third, since the number of 'explorers' follows a stationary pattern, the average per-period number of 'discoveries' keeps constant. Moreover, as the symmetry of the 'exploration' rule should suggest - cf. Eqs. (6) and (7) - the distance from the origin of any new technology increases linearly with the number of discovered islands - see Figure 11(a). However, the path-dependent nature of innovation implies that the initial productivity of a new island (i.e. the coefficient $s_{j^*}$) is generally greater than the average current productivity over all 'known' islands - cf. Figure 12 - while the one-time push irregularly caused by the introduction of 'new paradigms' keeps the order of magnitude of initial productivity of new technologies constantly above their distance from the origin - see Figure 11(b).

Fourth, relatively ordered spatial patterns of colonized islands are likely to emerge, due to the local nature of both the exploration and imitation processes. In Figure 13, the path of expansion of the 'best practice' proxy $b_t^*$ is plotted, together with four 'snapshots' showing the locations of currently 'colonized' islands in the technological space for different time periods $t = 0, 500, 1000, 1500$. While in the early time periods of the simulation small (stochastic) events select the region of the lattice where exploration is going to take place, the path-dependent nature of the overall process tends to keep the economy inside that region. In such a situation, 'rare' events (i.e. the exceptional discoveries), feeding path-dependently upon diffusion and incremental innovations thereafter, might be able to trigger a self-reinforcing process whose ultimate outcome could

[40] More on the statistical properties of simulated GNP series is in Sections 5.2 and 6.

[41] All results reported in this sub-section refer to the parametrization: $\pi = 0.1, \rho = 0.1, \epsilon = 0.1, \lambda = 1, \varphi = 0.5, N = 100, \alpha = 1.5, T = 1000$, but they robustly arise in all setups yielding exponential growth. See Section 5.1. for an extensive Montecarlo analysis of the parameter space.

be a pattern of exponential growth. Indeed, some 'lucky' explorers - which have decided not to imitate one out of the cluster of colonized islands - will eventually find intrinsically superior islands outside the 'realized economy'. Although they might not be able to adequately exploit the opportunities of the 'new' island by themselves, the 'extraordinary' character of their discovery might nevertheless induce other agents to move there in the future and, consequently, increase its actual productivity.

Finally, accordingly to empirically detected patterns of innovation, diffusion and adoption (see e.g. Dosi (1982)), the model generates *S-shaped* diffusion curves in the number of agents currently mastering a given technology. Moreover, because many techniques are allowed to coexist over the same time intervals if they exhibits sufficiently similar realized productivities, one usually observes overlapping diffusion patterns as those depicted in Figure 14. As the set of current available technologies keeps enlarging due to the unceasing process of exploration and innovation, firms migrate from less toward more productive islands, entailing processes of innovation diffusion which occur at different rates. As Figure 15 shows, the latter typically depend on the characteristics of the technologies involved in the process, the incentives provided by the economic environment and the features of the adopters themselves. In very general terms, the speed at which innovations are adopted (and substituted) is increasing in both their absolute *intrinsic* productivity gap and the extent to which interactions are global - cf. Panels (a), (b) and (d). Also, if information is diffused not too locally, radical innovations tend to retain their leadership much longer than incremental ones (see Panels (a) and (b)). Yet, the rate at which innovations are substituted is decreasing with the average willingness to explore of the agents in the system - cf. Panel (c).

# 5 The Sources of Self-Sustaining Growth : A Montecarlo Analysis

A basic insight stemming from the foregoing qualitative analysis is that, in order for the economy to be able to attain self-sustaining growth, the following conditions (or a suitable mix of them) ought to apply, namely: (i) production is characterized by increasing returns ($\alpha > 1$); (ii) the level of both opportunities (as measured by $\lambda$ and $\pi$) and willingness to explore ($\epsilon$) is sufficiently large; (iii) knowledge diffusion is not too 'local' ($\rho$ small); (iv) there is some path-dependency in innovation ($\varphi$>0). However, the presence of direct interactions among spatially distributed agents, as well as non-linearities in microeconomic behaviors, suggests that threshold effects and non-monotonic relationships are likely to emerge as one attempts to map sub-regions of the parameter space into robust aggregate regularities characterizing growth patterns.

In this section, we will present extensive Montecarlo (MC) analyses supporting the above conjectures and further investigating their implications.

More specifically, we will consider $M$ independent realizations of $\underline{q}(\omega)$, i.e. a simulated log(GNP) time-series for a given parametrization $\omega \in \Omega$ and define $\mathbf{Q}(M, \omega)$ as the $T \times M$ matrix whose $m$-th column is given by: $\underline{q}_m(\omega) = [q_{m,0}, \ldots, q_{m,T}]'$. Let us also set $\underline{h}_m(\omega) = [h_{m,1}(\omega), \ldots, h_{m,T}(\omega)]'$, where $h_{m,t}(\omega) = [(Q_{m,t} - Q_{m,t-1})/Q_{m,t-1}]$, and call $\mathbf{H}(M, \omega)$ the $T \times M$ matrix whose $M$ columns are made by MC growth rate time series $\underline{h}_m(\omega)$. Given any statistics $S : \Re^T \to \Re^\kappa$, mapping $\underline{q}_m(\omega)$ into a $\kappa$-valued real vector, our interest is in assessing how the sample moments of the MC distribution of $S$ - e.g. means and standard deviation of $S(\omega) = \{S(\underline{q}_m(\omega)), m = 1, \ldots, M\}$ - depends on $\omega$[42].

In Section 5.1, we shall firstly present a study of how the overall performance of the economy, as measured by its average growth rate (AGR) $S(\underline{q}_m(\omega)) = g_m(\omega) = [(q_{m,T}/q_{m,0})^{1/(T+1)} - 1]$, is affected as $\omega$ varies in $\Omega$[43]. Next, in Section 5.2 we will try to investigate the emergence of stochastic non-stationarity in log(GNP) time-series, in order to attempt to distinguish those sub-regions of the parameter space which are able to generate (with sufficiently high likelihood) patterns of self-sustaining growth. Finally, in Section 5.3, we will turn our attention to output growth rates matrix $\mathbf{H}(M, \omega)$, so as to discuss some finite-time properties of the average within-simulation growth-rate volatility and its long-run behavior in different growth regimes.

## 5.1 Economic Performance and the Exploitation - Exploration Trade Off

A first clear-cut result that MC simulations point out is that - everything else being constant - the overall performance of the economy, as measured by its AGR, appears to be monotonically increasing with respect to: (a) the extent to which the system is fueled with innovation 'opportunities' (i.e. both $\lambda$ and $\pi$); (b) the magnitude of path-dependency affecting the innovation process (i.e. $\varphi$); (c) the degree of globality of the information diffusion in the interaction process (i.e. $-\rho$). Support to this general conclusion is presented in Figure 16, where MC sample means of AGRs (i.e. $\overline{g}_M(\omega) = M^{-1} \sum_{m=1}^{M} g_m(\omega)$) are plotted against $(\log_{10} \rho, \varphi)$ in four distinct opportunity setups (and for given $\alpha$ and $\epsilon$).

The shape of the surfaces $(\log_{10} \rho, \varphi | \cdot)$ also suggests that, while path dependency linearly affects AGRs, locality of interactions is likely to induce some threshold effects. Indeed, as Figure 17 clearly shows, when one gradually tunes the rate of information diffusion from its minimum (i.e. $\rho = \infty$) toward its maximum (i.e. $\rho = 0$), an abrupt change in AGRs usually arises around $\rho^*(\omega) \cong 1.0$: if $\rho < \rho^*(\omega)$, AGRs are barely influenced if the rate of information diffusion is slightly altered; however, when $\rho > \rho^*(\omega)$, small changes in the degree of locality of interactions bring about very large effects on the overall performance of the economy.

As the economy is gradually injected by increasingly powerful sources of growth (larger opportunities, higher path dependency and more conducive information diffusion), one should also expect an increasing MC sample volatility of AGRs. In fact, as reported in Figure 18, a strong positive correlation usually emerges between MC sample means of AGRs and MC sample standard deviations $\sigma_{g_M}(\omega) = [M^{-1} \sum_{m=1}^{M} g_m^2(\omega) - \overline{g}_M^2(\omega)]^{\frac{1}{2}}$, as the latter appear to be monotonically increasing with either parameter, everything else held constant[44].

[42] See Appendix 3 for a brief discussion on the reliability of the first and second Montecarlo moments as estimators of the 'true' moments of the data generating process.

[43] Notice that all implications presented in this Section are not affected by the particular choice of the AGRs. Indeed, employing alternative measures of average growth rates, as $g'_m = [(q_{m,T} - q_{m,0})/T]$ or $g''_m = [(Q_{m,T}/Q_{m,0})^{1/(T+1)} - 1]$, will only change the scale of attainable growth rates. Cf. also Appendix 3.

[44] For a similar property displayed by actual time-series in a cross-section of countries, cf.

MC sample standard deviations never 'explode', however, as one increases the strength of the sources of growth. Hence, despite the self-reinforcing nature of the mechanisms triggering economic growth in the system (i.e. exploration, innovation and production), the model yields sufficiently ordered growth paths, which turn out neither to overlap nor converge as long as one considers sets of GNP time-series generated by points in the parameter space far enough, cf. Figure 19 for an illustration of such property[45].

If the degree of opportunities, path-dependency and globality of interactions seem to engender a non ambiguous effect on the overall performance of the economy, different regimes arise when the same experiment is carried through with respect to the extent to which agents are willing to explore (i.e. $\epsilon$) and the strength of returns to scale in the production process (i.e. $\alpha$). This is not surprising, indeed, as the latter parameters are responsible for tuning the forces underlying the trade-off in the choice between searching and producing, the solution to which is a key ingredient for achieving better economic performances.

In Figure 20 we have plotted some relevant portions of the surface $(\epsilon|\lambda,\ \pi,\ \rho,\ \varphi;\ \cdot) \longmapsto \overline{g}_M(\omega)$ to first investigate how AGRs depend on $\epsilon$. As one could have expected, larger AGRs can be attained on average if the system somehow manages to optimally solve the trade-off between exploitation and exploration (March (1991), Allen and McGlade (1986)). It turns out, however, that one can single out four distinct regimes in the causal relation between $\epsilon$ and MC mean of AGRs, namely:

R1. If no interactions take place ($\rho = \infty$) and opportunities are low, then $\overline{g}_M(\omega)$ is monotonically decreasing with $\epsilon$ - no matter how large is path-dependency - and displays a maximum at $\epsilon^* = 0$; since $\overline{g}_M(\epsilon^*) \equiv 0$ (closed economy and no information diffusion), AGRs are always negative for $\epsilon > 0$;

R2. If information is spread locally (i.e. $0 << \rho << \infty$), then, irrespective of path-dependency and opportunities, $\overline{g}_M(\omega)$ are decreasing for small $\epsilon$, increasing when $\epsilon$ is large, and exhibit two maxima at $\epsilon^* = 0$ and $\epsilon^* = 1$;

R3. If information is spread globally ($\rho = 0$) and path-dependency is high, then $\overline{g}_M(\omega)$ is monotonically increasing with $\epsilon$, independently of opportunities, so that a maximum arises at $\epsilon^* = 1$;

R4. Finally, in all other 'intermediate' cases, i.e. either global interactions and low path-dependency or no information diffusion and high opportunities, $\overline{g}_M(\omega)$ exhibits an interior maximum $0 < \epsilon^* < 1$.

---

Fatas (1995).

[45] In Figure 19 we plotted the time-series describing the 5% and the 95% percentiles of the MC distributions $\underline{q}_t(\omega) = \{q_{m,t}(\omega), m = 1, \dots, M\}$, as $t = 1, \dots, T$, in four different parameter setups ($M = 10000$). Notice that even in the global information / high opportunities case, the band including the 90% of MC observations does not enlarge as $T$ grows. Moreover, 90% bands do not overlap even for very small econometric sample sizes. Furthermore, additional simulation exercises seem to support the following general conjecture. Let $d$ be a metrics measuring the distance between any two MC samples $\underline{q}(M, \omega') = \{q_m(\omega'), m = 1, \dots, M\}$ and $\underline{q}(M, \omega'') = \{q_m(\omega''), m = 1, \dots, M\}$ $\omega', \omega'' \in \Omega$. Then, for any MC sample size $M$, there will exist an upper bound $\vartheta_M > 0$ - decreasing in $M$ - such that, taken any $0 < \vartheta < \vartheta_M$ and $\omega \in \Omega$, it should be always possible to find a $d(\vartheta_M, \omega) > 0$, so that in the $d(\vartheta_M, \omega)$-neighborhood of $\omega$ one can find a $\omega' \in \Omega : d[\underline{q}(M, \omega), \underline{q}(M, \omega')] > \vartheta$.

The interpretation of the above evidence is straightforward. When information diffusion is not active, no newly discovered technology can exploit increasing returns to scale and exploration is totally harmful (because, in the long run, it leads to negative AGR). Conversely, economies in which information is globally diffused and innovators strongly benefit from learning by doing (high $\varphi$) typically maximize their AGR when all agents commit themselves to exploration and production on new islands only lasts one period.

As information diffusion becomes local, the overall performance of the economy increases either if few explorers are around or if there are many. In the first case, a large population of miners can continually exploit both increasing returns to scale and incremental, path-dependent innovations through small-scale migrations driven by local imitation. In the second one, still thanks to local information diffusion, small clusters of colonized islands can immediate benefit from the large-scale introduction of innovations.

The most interesting regime, however, arises when MC mean of AGRs are maximized by an interior value of $\epsilon$. The intuition here corresponds to that suggested in March (1991, p.71). As he points out, systems that engage in exploration to the exclusion of exploitation "exhibit too many undeveloped new ideas and too little distinctive competences", while, at the opposite extreme, they "are likely to find themselves trapped in sub-optimal stable equilibria". In the model, this condition applies in two setups, namely: (a) agents face very large opportunities but they are unable to completely exploit returns to scale because information is not spread around; (b) interactions are global but knowledge does not accumulate as the economy evolves. In both situations, higher economic performances cannot be attained by entirely committing themselves either to technological search or to production. As a result, the losses stemming from the exploitation-exploration trade-off are minimized by an appropriate balance between the two forces, which, however, agents are unable to correctly evaluate ex-ante.

Let us turn now to discuss how $\alpha$ affects the overall performance of the economy. In very general terms, larger values of $\alpha$ should imply, above some thresholds, higher returns from exploitation. Consequently, imitation should drive agents, on average, to concentrate on the most efficient islands (no matter if they are newly discovered or already 'colonized'). However, the extent to which this process is able to entail larger AGRs strongly hinges on the features of information diffusion and the amount of opportunities in the economy.

Some paradigmatic examples of the surface $(\alpha|\lambda, \pi, \rho, \varphi; \cdot) \longmapsto \overline{g}_M(\omega)$ are plotted in Figure 21. First of all, notice that, as long as information is globally diffused in the economy, AGRs are monotonically increasing in $\alpha$ for any $(\lambda, \pi)$ and $\varphi$. In this case, irrespective of the RTS regime, no trade-off between exploration and exploitation arises: agents easily manage to move toward islands with larger actual productivities so that higher $\alpha$'s imply better performances. Conversely, when interactions are shut down ($\rho = \infty$), miners tend to remain locked into sub-efficient states, especially when opportunities are low. In this case, both explorers and miners are likely to be captured by inefficient technologies which have been exploited in an extensive way during early periods of time. As a consequence, MC means of AGRs tend to be decreasing with $\alpha$ when opportunities are low and only mildly increasing for large $\alpha$'s when they are high. The latter pattern arises more vividly when interactions are local ($0 < \rho < \infty$). Here a threshold effect emerges, as $\overline{g}_M(\omega)$ starts growing with $\alpha$ only after RTS become both technologically and individually increasing (i.e. $\alpha > 2$). To understand this nonlinear effect, recall that the two-stage procedure governing interactions requires that agent $i \in I$, whose last output on $j^*$ was $q_{i,\tau} = s_{j^*} \cdot [m_\tau(x_{j^*}, y_{j^*})]^{\alpha-1}$, would choose to imitate the signal already received from island $j \neq j^*$ with a probability proportional to $j$'s actual productivity $s_j[m_t(x_j, y_j)]^{\alpha-1}$, while he would decide to remain on $j^*$ with probability proportional to $q_{i,\tau}$. Thus, in general, imitation begins to trigger a strong self-reinforcing process of knowledge accumulation when $\alpha > 2$. Therefore, if information is locally diffused agents tend to be trapped into inefficient islands for small $\alpha$'s, whereas they will be able to fully exploit the economy's opportunities when RTS are large enough to promote not wasteful migrations.

### 5.2 Stationarity vs. Stochastic Non-Stationarity of GNP Time-Series

Better overall economic performances (e.g. higher average GNP growth rates) seems to be generated in the system - above certain thresholds - by a synergetic mechanism involving in non-additive ways innovation, path-dependency, increasing returns and diffusion of knowledge, and *not* by any of these forces taken in isolation. But, in the first place, is there a well-defined mapping between larger magnitudes of AGRs and the emergence of patterns self-sustaining growth ?

As a first approximation, one might attempt to discriminate regions of the parameter space $\Omega$ yielding patterns of self-sustaining growth from those in which growth is only a transitory phenomenon, depending on whether log(GNP) time-series $\underline{q}_m(w)$ display an $I(1)$ pattern or turns out to be stochastically stationary once standard ADF tests are performed[46].

Some examples of such an exercise are presented in Figures 22. Given low or high opportunities setups as before, we have firstly computed the percentage of acceptance of ADF(1) test statistics $t_1(q_{m,t}(w))$ - at 5% of significance - over $M = 10000$ independent MC simulations as a function of the degree of globality of information diffusion ($\rho$) and path-dependency in knowledge accumulation ($\varphi$) - for some given willingness to explore ($\epsilon$) and returns to scale ($\alpha$)[47].

Similarly to the AGRs case, the mean of MC distributions of $t_1(q_{m,t}(w))$ are increasing exponentially with $\rho$ and linearly with $\varphi$ (see Panels (a) and

[46] In the model, exponential growth is usually associated with 'difference stationary' log(GNP) time-series. In fact, according to standard ADF tests - and irrespective of the employed Dickey-Fuller regression specification - one is always unable to reject (at 5% of significance) the null of a unit root, which, on the contrary, is systematically not accepted for both first differences $\Delta q_{m,t}$ and growth rates $h_{m,t}$. In Table 1 we report as an illustration the results of ADF tests performed on the GNP time-series displayed in Figure 7. See however Section 6 for a discussion on the drawbacks (in particular, lack of power against nearby stationary alternatives) of ADF tests and for additional statistical properties of simulated GNP time-series.

[47] The choice of ADF(1) tests - constant and trend included - has been suggested by a preliminary experiment in which an extensive battery of conventional Dickey-Fuller regressions: $\Delta q_{m,t} = \gamma_0 + \delta t + \gamma_1 q_{m,t-1} + \phi_1 \Delta q_{m,t-1} + \ldots + \phi_k \Delta q_{m,t-k} + u_t$ has been estimated for a small subset of points in the parameter space. All diagnostics (Schwartz, Akaike, etc.) have indicated that $k = 0$ is inadequate while $k > 2$ is unnecessary (and therefore wasteful of degrees of freedom). Also, notice that this choice for $k$ is in line with the current literature testing for trend vs. difference stationarity in empirical GNP time-series, cf. Diebold and Senhadji (1996). Note also that the results presented here are sufficiently robust across different parametrizations of willingness to explore and returns to scale.

(b)). This implies that the null hypothesis (i.e. presence of a unit-root in the log(GNP) time-series) is accepted with an increasing MC frequency as one tunes up the sources of growth. In fact, a sort of threshold emerges in the $(\varphi, \log_{10} \rho)$-plane: beyond some given combinations of path-dependency and globality of interactions the model delivers almost always difference-stationary log(GNP) time-series. In addition, the portion of the $(\varphi, \log_{10} \rho)$-plane containing MC frequencies of the ADF(1) test acceptance greater than 90% is larger the greater the magnitude of opportunities, as expected.

Results in line with the analysis in Section 5.1 also arise when one investigates how the MC acceptance frequency of the ADF tests varies with the magnitude of the willingness to explore ($\epsilon$). According to the evidence presented above for AGRs, four regimes can be singled out as far as the emergence of patterns of self-sustaining growth is concerned. Apart from the trivial results of regime $R1$ (always yielding negative AGRs and then $I(0)$ patterns), Figure 23 illustrates what happens to ADF(1) 5%-acceptance frequencies in regimes $R2$, $R3$ and $R4$. An interesting insight is that, because of the positive correlation usually arising between the likelihood of $I(1)$ GNP patterns and large magnitudes of AGRs, one is likely to get some within-regime variation in the behavior of acceptance frequencies. To see this, Panels (b) and (c) depict two examples of regime 3 behavior (global interactions, high path-dependency). Since the performance of the economy typically increases with the magnitude of opportunities, ADF tests acceptance frequency is constantly around 100%, for any $\epsilon > 0$, when $\lambda$ and $\pi$ are high, while it attains a 100% steady state only for larger $\epsilon$'s when opportunities are small. The same argument also holds for regime 2 (see Panel (a)): patterns of self-sustaining growth typically appear when $\epsilon$ is large as soon as path-dependency gets larger. Finally, Regime 4 still implies an 'interior solution' for the exploitation-exploration trade-off (Panel (d)). In all those cases whereby the economy is characterized either by global interactions and low path-dependency or by no information diffusion and high opportunities, the system is able to generate self-sustaining patterns of growth *only if* a suitable balance between R&D and production is achieved.

## 5.3 Volatility of Growth Rates Time - Series and Self - Organization

Let us turn now the attention to MC samples of growth rates time-series (GRTS):

$$\underline{h}_m(\omega) = \{h_{m,t}(\omega) = [(Q_{m,t}(\omega) - Q_{m,t-1}(\omega))/Q_{m,t-1}(\omega)], \quad t = 1, \dots, T\},$$

$m = 1, \dots, M$ and explore their aggregate properties.

An interesting issue to address - motivated by the foregoing evidence on AGRs variability *across* independent MC simulations, see Section 5.1 - concerns the behavior of the *within-sample* volatility of GRTS, e.g. their standard deviation $S(\underline{h}_m(\omega)) = \sigma(\underline{h}_m(\omega))$. In particular, we will firstly discuss whether self-sustaining growth always imply a larger volatility in GRTS, by studying how $\sigma(\underline{h}_m(\omega))$ depends on opportunities, path-dependency and information diffusion for a given sample size $T$. Second, we will study the average *within-simulation* behavior of $\sigma(\underline{h}_m(\omega))$, to understand what happens to the volatility of GRTS across successive phases of development.

To begin with, let us consider Figure 24, where sample means of MC distributions of the GRTS standard deviations are plotted against the main sources of growth. Unlike the behavior exhibited by the variability of AGRs across MC observations, the correlation between magnitudes of GRTS volatility and average growth rates attained by the system is not always positive.

As one could have expected, when $\lambda$ is low, more volatile GRTS are generally implied by more global interactions (and, to a smaller extent, by higher path-dependency). In such a case, the economy is likely to go through punctuated output upsurges caused by the arrival of 'new' paradigms whose higher productivity is (possibly) further boosted by knowledge accumulation. Thanks to the effectiveness of information diffusion, agents almost instantaneously decide to adopt the new knowledge bases. This will in turn trigger prolonged periods of negative (but low in absolute value) growth rates, because of the time-consuming nature of the adjustment process followed after imitation.

Conversely, when radical innovations are very likely, the relation between relevant parameters and average $\sigma(\underline{h}_m(\omega))$ undergoes a dramatic change. In fact, setups typically yielding self-sustaining growth (small $\rho$'s, large $\varphi$'s) are characterized by lower magnitudes of average volatility, whereas economies usually generating stationary GNP time-series or very mild growth display an higher GRTS variation.

As opportunities become larger, the amount of path-dependency increasingly influences GRTS variability - cf. Figure 25. On the one hand, larger $\varphi$'s do not affect at all GRTS standard deviations when interactions are global, as knowledge accumulation is efficiently driven by information diffusion. On the other hand, as one keeps decreasing the rate at which information is spread around, path-dependency gradually becomes the main force allowing for a self-enforcing process of accumulation of existing competencies. The large amount of additional knowledge coming from extraordinary discoveries is then expected to be carried on to the next innovations (at a rate measured by $\varphi$). As a consequence, the economy is likely to be characterized by periods of near-zero growth rates (exploitation periods) intertwined by recurring huge jumps caused by the introduction of radical innovations, i.e. by a very large GRTS volatility.

To sum up, in all possible setups eventually yielding self-sustaining growth, the economy displays an unexpectedly low average GRTS variability, in particular when opportunities are very high. Furthermore, the GRTS sample volatility coming from GNP time series characterized by exponential growth is at least of the same magnitude of - but more often lower than - that generated by stationary output realizations (i.e. no or mild growth).

Even more unexpectedly, self-sustaining growth economies appear to attain persistently higher output growth rates through a self-organizing process characterized by GRTS volatility decreasing in time. To illustrate this property, we have considered four prototypal environments yielding[48]: (a) stationary GNP time-series; (b) levels of GNP evolving around a S-shaped trend; (c) self-sustaining growth emerging from a low opportunities setup; and (d) self-sustaining growth emerging from a high opportunities one. In each environment,

[48] The following qualitative growth patterns arise robustly across $M = 10000$ independent MC simulations, see Table 2.

we have computed MC means of the distributions of recursive sum of squares and recursive standard deviation of GRTS, as well as of standard deviation of GRTS over disjoint samples, cf. Figure 26. As one takes into account the time evolution of GRTS volatility of enlarging econometric samples $\{T_0, T_0+1, \ldots, \overline{T}\}$, for $\overline{T} = T_0+20, T_0+21, \ldots, T$ and $T_0 = 50$, a striking pattern arises. Recursive sum of squares of GRTS deviations usually increases as $\overline{T}^{\beta}$, with $\beta > 0$. However, $1 \leq \beta < 2$ in the stationary GNP cases, so that recursive sum of squares tend to explode, while recursive standard deviation grows as $\overline{T}^{\beta-1}$. Conversely, as soon as some evidence of persistent economic growth emerges in the system, $\beta$ becomes less than unity - see Panels (a) and (b) - and recursive standard deviations turn out to be monotonically decreasing and converge toward some positive constant. In general, a negative correlation appears between $\beta$ and the overall performance of the economy: the more one fuels the system with opportunities and path-dependency, the higher the rate at which GRTS volatility, as measured by average recursive standard deviation, decreases in time. The same kind of regularity arises when one plots MC means of GRTS standard deviation computed over disjoint subsamples $[T_0, T_0+k-1]$, for some fixed $k$, against $T_0 = 1, k+1, 2k+1, \ldots$ (see Panel (c), Figure 26).

Hence, the model is able to account for the appearance, over finite time periods, of distinct patterns (or 'phases') of development. Exponential growth emerges as the outcome of a self-organization process leading to 'ordered' GNP time-series characterized by fairly moderate variability both across independent histories and, more importantly, within time realizations. By means of an imperfect adjustment process to unpredictable structural changes endogenously introduced in the system, our boundedly rational agents are able to reach nearly efficient aggregate states. This 'Schumpeterian' coordination is mainly achieved through networks of direct and (partly) local interactions which establish whenever new knowledge bases appear in the economy. The fact that interaction structures can change in time allows the system to eventually unlock itself from conceivable inefficient states in which the economy could be trapped if agents should always master a fixed set of technologies and interact with the same 'relevant others'. As a result, under structural conditions above certain thresholds, the economy manages to self-organize by conveying initial phases of turbulence into an aggregate dynamics in which phases of almost steady positive growth rates are punctuated by temporary slowdowns characterized by small volatility.

# 6 Statistical Properties of Simulated GNP Time-Series

In the foregoing Sections we have attempted to shed light on the mechanisms allowing for the emergence of persistent growth in the model. However, an interesting question still to be answered concerns the actual ability of the model to generate time-series displaying statistical properties similar to those exhibited by the empirically observed ones.

In this section, we will address this 'exercise in plausibility' by trying to single out some statistically measurable features detected in the actual 'business cycle' (e.g. the magnitudes of the auto-correlations of output growth, the persistence of oscillations, etc.) which robustly arise also in simulated data. In principle, two related lines of analyses might be pursued. First, as in Section 5, one could try to roughly map regions of the parameter space into different statistical regimes of 'business cycle' statistics. Second - and somewhat similarly to what has been done in the literature under the heading 'calibration of Real Business Cycle' models (see e.g. Cogley and Nason (1995)) - one might attempt to distinguish regions of the parameter space (if any) which are capable to generate simulated GNP time series whose features exhibit statistically insignificant differences with those computed on the empirically observed ones.

An initial *caveat* about the extent to which these kinds of analyses can be informative is however required. Despite a huge empirical literature aiming at uncovering 'stylized facts' in the business cycle, a very little consensus has emerged through the last years about the actual properties of aggregate output. Many debates still remain open, mainly because the statistical procedures routinely employed have low power in presence of the generalized lack of large spans of data[49]. Years after the seminal work of Nelson and Plosser (1982), for instance, the question of whether empirical log(GNP) time-series are trend- or difference-stationary is still unresolved[50]. Furthermore, as long as linear univariate time-series analysis is concerned, there seems to be no conclusive indication about the best parametric (ARIMA) specification fitting, e.g., U.S. de-trended real log(GNP) series[51]. In general, the only clear suggestion coming from the empirical literature is that of pursuing case-by-case analyses, possibly employing a multivariate approach[52].

Because of all that, we have chosen not to push further univariate parametric studies - either focusing on the issue trend- versus difference-stationarity or uncovering preferred univariate ARIMA models for MC samples $\{\underline{q}_m(w), m = 1, \ldots, M\}$[53]. On the contrary, we will employ here non-parametric analyses to

[49] Cf. among others Blanchard and Fischer (1989) and Romer (1996).

[50] It is a well-known result that standard ADF tests for 'stochastic non-stationarity' suffer from very low power against nearby 'trend-stationary' alternatives. Many authors have indeed proven that they are inherently incapable to discriminate between the null and the alternative hypotheses on the basis of a finite sample of observations (see e.g. Bloug (1992), Christiano and Eichenbaum (1989) and Rudebusch (1993)). Conversely, many other contributions have recently appeared suggesting that unit-root tests can be nonetheless informative, at least over long spans (DeJong and Whiteman (1991, 1994)). In this connection, Cochrane (1988) has pointed out that the use of longer GNP samples (as in our case) may produce sharper unit-root inference. Yet, evidence stemming from this strand of literature seems to conclude that U.S. aggregate output is not likely to be difference stationary (Diebold and Senhadji (1996)). Notice also that the distinction between trend- and difference-stationarity is potentially important only in economic forecasting, but might not be so critical in many other contexts.

[51] See, among others, the contrasting findings stemming from Watson (1986), Campbell and Mankiw (1987), Clark (1987), Stock and Watson (1988), Gagnon (1988), Blanchard and Fischer (1989).

[52] Cf. Blanchard and Quah (1989) and Cochrane (1994). Unfortunately, the endogeneity of all variables generated in the model prevents us from any meaningful multivariate analysis.

[53] As we have already pointed out in Section 5.2, the simulated log(GNP) time-series $\underline{q}_m(\omega)$ turn out to be difference stationary almost always above well-defined threshold in the parameter space. Despite all the drawbacks of ADF tests, this result seems to match those obtained earlier for U.S. GNP by Nelson and Plosser (1982) and Stock and Watson (1986). For opposite findings, cf. Diebold and Senhadji (1996). As to preferred univariate models for $\Delta\underline{q}_m(\omega)$, preliminary analyses generally display results similar to those in Campbell and Mankiw (1987). Given the well-know drawbacks of selection procedures based on both Akaike and Schwartz criteria, we have simply selected the ARMA$(p, q)$ specification which most frequently get the maximum likelihood given $p + q$ (over $M = 50$ Montecarlo replica-

investigate some important features concerning the autocorrelation structure of output growth rates and the persistence properties of the 'business cycle'.

One of the few unquestioned 'stylized facts' of U.S. quarterly GNP growth - also observed in other Countries, with some notable exceptions: see Campbell and Mankiw (1989) - is that it is positively autocorrelated over short horizons, while the autocorrelation function (ACF) over higher lags is not significantly different from zero. In particular, the first GNP growth ACF coefficient appears to be large and positive for U.S., Canada, Italy and, to a less extent, Japan.

Figure 27 shows the MC means (over $M = 1000$ replications) of the ACF of GNP growth in some paradigmatic parameter regions yielding $I(1)$ patterns of log(GNP) time-series. Clearly, the model is perfectly able to robustly replicate the above statistical property, in particular when interactions are global and opportunities are large enough. In all these cases, the positive shocks to output growth rates coming from innovations are almost instantaneously spread in the economy, leading to high and positive first- and second-order autocorrelation coefficients, followed by insignificant values over longer horizons. Conversely, when interactions are local and opportunities are low, GNP growth does not display ACF coefficients significantly different from zero (as they almost always fall inside the 5% Bartlett confidence bands). Notice also that the associated estimates for the spectral densities[54], albeit much smoother than the empirical ones, usually display a peak around low frequencies and then tend to decrease as the length of the period becomes small, cf. Figure 28.

To get a quantitative measure of whether the model is really able to replicate empirically observed autocorrelation functions (ACF) for output growth, we have also computed generalized Q-statistics defined as:

$$Q_{acf} = (\underline{r} - \overline{\underline{r}}(\omega))'[\mathbf{V}(\omega)]^{-1}(\underline{r} - \overline{\underline{r}}(\omega))$$

where $\underline{r}$ is the $k$-vector containing the empirically observed ACF; $\overline{\underline{r}}(\omega)$ is the k-vector containing the MC means (computed over $M = 1000$ independent simulations) of the estimated ACF generated by the model under the parametrization $\omega \in \Omega$, i.e.:

$$\overline{\underline{r}}(\omega) = \frac{1}{M}\sum_{m=1}^{M}\underline{r}_m(\omega)$$

and $\mathbf{V}(\omega)$ is the $k \times k$ MC covariance matrix:

$$\mathbf{V}(\omega) = \frac{1}{M}\sum_{m=1}^{M}(\underline{r}_m(\omega) - \overline{\underline{r}}(\omega))(\underline{r}_m(\omega) - \overline{\underline{r}}(\omega))'$$

Generalized Q-statistics are approximately $\chi^2(k)$ distributed. Table 3 shows

tions). Telegraphically, $\Delta\underline{q}_m(\omega)$ appear to be ARMA$(1,0)$ or ARMA$(0,2)$ when interactions are local and path-dependency is low, while seem to be better described by ARMA$(3,0)$ or ARMA$(2,2)$ when interactions are global and path-dependency is high.

[54] Spectral densities have been estimated by smoothing the periodogram using a Bartlett window of $k = 100$ and plotted in the frequency domain $[0, 3.14..]$, scaled to match the unit interval.

some results obtained for $k = 50$[55]. The model appears to generate very similar autocorrelation structures for U.S. and Italy in both global and local interaction setups. Moreover, because ACFs for some European countries present very different patterns, with negative low-lags coefficients (e.g. France, U.K., Germany), one is not able to reject the null of insignificant differences between empirical and simulated ACF also in the local interaction case.

Another prominent debate in the study of business cycle, strongly related to the analysis of the structure of ACF output growth, concerns the issue of whether GNP fluctuations are characterized by a permanent component and, if so, how big such a component might be. Indeed, if output levels were stationary around an exogenous trend, any shocks to GNP would eventually die out and long-term forecast will be unaffected. On the contrary, if GNP levels could be described by a random walk (or an even more persistent process), then the series would continue to diverge from its previously forecast once an innovation occurs.

Following Campbell and Mankiw (1987, 1989) and Cochrane (1988), we have considered two non-parametric measures of persistence of GNP fluctuations based on sample estimates of auto-correlations of output growth (cf. Appendix 4 for details).

As Table 4 clearly shows, both measures decreases as the window size $k$ grows, but generally stabilizes around values exceeding unity in all the experimented parametrizations. This result, quite in tune with the findings of Campbell and Mankiw (1989), implies that our simulated GNP time-series do not appear to revert toward any smooth exogenous trend and exhibit very persistent fluctuations: a 1% shock to output should indeed change the long-run univariate forecast of GNP levels by far more than 1%. Also, persistence turns out to be higher the more interactions are global, the larger the likelihood of 'radical' innovations and the *smaller* the density of islands in the economy.

Notice, however, that both estimated measures, albeit non-parametric, do display significant drawbacks. First, as noticed above, they strongly depend on the availability of long spans of data, so that Campbell and Mankiw's empirical results could not be so informative. Moreover, standard errors of the estimated $V^k$ are usually very large (cf. Table 4 and Appendix 4). This implies that is often very hard to distinguish if an estimated measure greater than unity really comes from a non-stationary high-persistent process. However, as standard errors are increasing with the window size $k$, one is more likely to get a low $V^k$ from an highly persistence process than a large $V^k$ from a less persistent one.

# 7 Some Extensions of the Basic Model

In the present Section, we will show how the model can be naturally extended to deal with some issues of interest in formal growth literature. First, the assumption of a constant population size will be relaxed so as to study whether 'scale effects' arise in the model. Second, we will introduce some alternative specifications of the exploration pay-off structure in order to reconsider the sources of the exploration-exploitation trade-off. Finally, we will illustrate the

[55] Our results are not dramatically affected by this choice. International data are homogeneous to those in Campbell and Mankiw (1989) and refer to quarterly output growth from 1960:1 to 1990:4 (Source: International Monetary Fund).

potential conflicts between individual behaviors and collective economic performance, arising in the model because of the introduction of heterogeneity, either in agents 'willingness to explore' or in the very structure of agents' behavioral rules.

## 7.1 Size of the Economy and Scale Effects

A well-known drawback of many models of endogenous growth based on some forms of increasing returns - involving dependence of a *flow* variable upon a *stock* variable, e.g. arrivals of technological 'blueprints' as a function of their levels - is that sheer scale effects influence output growth rates[56]. For instance, many one-factor models, such as Romer (1986), predict that growth rates are increasing, other things being equal, in the size of the population. Furthermore, when one considers extensions of these basic models - such as multi-factors models (Aghion and Howitt (1992), Grossman and Helpman (1991a) and Romer (1990)) and with international trade (Grossman and Helpman (1991c)) - the standard result is that growth rates are increasing in the factor used intensively in the 'innovative' activity (e.g. the stock of skilled labor).

The present model, notwithstanding increasing returns to learning, *does not* display that unreasonable property[57]. Figure 29 depicts the behavior of MC means of AGRs as a function of the population size ($N$) and of the econometric sample size ($T$), in a parameterization setup usually yielding self-sustaining growth. If any, weak evidence on *falling* AGRs the larger the size of the economy for a given time-length emerges. Moreover, AGRs do not display any monotone pattern when $N$ and $T$ both increase. The intuition behind this result is that, while *ceteris paribus* larger economies face potentially higher returns to knowledge exploitation, it is also true that they must cope, in probability, with higher 'adjustment lags' to new knowledge bases (as proxied in our model by the time it takes to move a certain fraction of the $N$ agents to the notionally superior islands). Hence, larger economies which are potentially able to fully exploit increasing returns to any one knowledge base need also a relative longer time to achieve persistently higher growth rates.

In order to offer a possible solution to the 'scale effect' problem common to many 'endogenous growth' models, Jones (1995a) has developed an alternative specification implying, in the steady state, a positive correlation between output growth rate and *population growth rate*. This implication, which does not seem

[56] We refer here to R&D-based models of endogenous growth, such as e.g. Romer (1990) and Grossman and Helpman (1991a,b). In these models, size-effects stem from three related assumptions, namely (i) technology is non rival, so that increases in the scale of the economy entail larger profits for all innovators; (ii) there are strong inter-temporal spillovers, i.e. each innovator can improve existing technology at any time; and (iii) new technologies are substitute for the old ones, so that returns to innovation are decreasing in the rate of innovation. All this implies that steady state analysis has to be conducted in the zero population-growth case. Conversely, in many models in which growth is endogenously generated by the accumulation of human and physical *rival* capital, any increase in the scale of the economy has no impact on growth rates (cf. Lucas (1988), Jones and Manuelli (1990) and Rebelo (1991)).

[57] Even though empirical results seem to strongly reject that e.g. growth rate of per capita GDP bears a significantly positive relation with the size of the working-age population *at a country level* - cf. Jones (1995b), McGrattan and Schmitz (1998) and Barro and Sala-y-Martin (1995) - some authors have argued that this is not so when higher levels of aggregation (e.g. world population) or lower ones (i.e. regional economies) are considered, cf. Kremer (1993).

to be rejected by empirical data[58], is also shared by our model when one assumes that, instead of being constant, the population exogenously grows at a rate $\eta > 0$, i.e.:

$$N_{t+1} = f(N_t) = \lfloor (1 + \eta) N_t \rfloor, \qquad t \geq 0,$$

and that the 'new born' agents $\Delta N_{t+1} = f(N_t) - N_t$ are randomly distributed across the $L_t$ currently known islands[59]. Moreover, as Figure 30 clearly shows, AGRs are linearly increasing with $\eta$ in all relevant parametrizations, as the model in Jones (1995a) implies as far as comparative statics exercises are concerned.

## 7.2 The Structure of the Exploration - Exploitation Payoff: An Alternative Specification

Two key ingredients driving the emergence of self-sustaining growth in the economy are, first, the particular behavioral rule governing exploration (see Eqs. 6 and 7) and, second, the assumption that technological search takes place over a somewhat 'flat' stochastic payoff landscape (i.e. $\pi$ constant across the lattice). An immediate implication of the latter hypotheses is that, once a miner has decided to start searching, the probability of finding an island in a finite number of time periods $t^*$ converges to 1 as $t^* \to \infty$[60]. Consequently the 'expected' returns from the activity of exploration should be - despite their extreme volatility - typically larger than those (certain) stemming from the exploitation of the existing knowledge bases.

To support this conclusion, let us let us start by considering the payoff landscape faced by the miner $i \in I$, working at time $t = t_0$ on island $j_0 \equiv (x_0, y_0)$ if it decides to become an explorer. Given an initial (certain) output of $q_{i,t_0}$, the payoffs of agent $i$ are described by the realizations of the stochastic process $\{\hat{q}_{i,t}\}_{t=t_0+1}^{\infty}$ , $\hat{q}_{i,t} \geq 0$, where $\hat{q}_{i,t} = \hat{q}_{i,t}(x_0, y_0, q_{i,t_0}; \omega)$ and $\omega \in \Omega$. In general, the probability structure of such a process is a very complicated object. Consequently, it might be convenient to analyze a setting in which information diffusion is absent (i.e. let $\rho = \infty$). Notice that in this case, since agents cannot imitate other technologies, the starting island $j_0$ will keep its population (on average) constant. Hence, $q_{i,t_0}$ can be taken as a measure of returns to exploitation and contrasted with some measure of returns from exploration[61]. Concerning the latter, let us define the *expected return* of agent $i$ at time $t$, as:

[58] At least if the model is taken as relevant for the World economy, see McGrattan and Schmitz (1998).

[59] This implication holds *a fortiori* when the new $\Delta N_{t+1}$ units are distributed among the currently colonized islands only, either randomly or with probabilities proportional to the number of miners currently working on them.

[60] Indeed, disregarding the effect of information diffusion, the probability that an explorer will find an island after (exactly) $s$ periods of exploration is $\pi(1-\pi)^{s-1}$, which sums to 1 over $s = 1, 2, \ldots$ .

[61] It is worth noting that in the following we are not addressing any individual-choice issues. Agents are not supposed to hold the rational capabilities necessary to compare expected streams of returns from different activities before taking their decisions. A simple setting in which the aggregate consequences stemming from *quasi rational* vs. *irrational* decisions are compared is presented in the Section 7.3.

$$\mathbf{E}_{t_0}[\hat{q}_{i,t}] = \int_0^\infty q f(\hat{q}_{i,t}) dq =$$

$$= \int_0^\infty q f(\hat{q}_{i,t}|\hat{q}_{i,k}, k = t_0+1, \ldots, t-1) \cdot f(\hat{q}_{i,k}, k = t_0+1, \ldots, t-1)\, dq =$$

$$= (1-\pi)^{t-t_0-1} \int_0^\infty q Pr\{A(s_t = \hat{q}_{i,t})\}\, dq \; = \; \pi(1-\pi)^{t-t_0-1}\mathbf{E}_{t_0}[s_t] =$$

$$= \pi(1-\pi)^{t-t_0-1}(1+\lambda)(d_0 + \varphi \cdot q_{i,t_0}) \; = \; \tilde{\pi}(1+\lambda)(d_0 + \varphi \cdot q_{i,t_0}), \tag{12}$$

where $A(s_t = \hat{q}_{i,t})$ stands for the event 'finding a new island at time $t$ with a productivity coefficient $s_t = \hat{q}_{i,t}$', $\tilde{\pi} = \pi(1-\pi)^{t-t_0-1} = \pi(1-\pi)^e$ is the probability of finding an island at time $t$ having left island $j_0$ at time $t_0$, and $e + 1 = t - t_0$ is the *actual* duration of exploration. Moreover, as the events 'finding a new island a time $t$' are mutually exclusive, the *total expected return* from exploration $\mathbf{E}_{t_0}[\hat{q}_i]$ reads[62]:

$$\mathbf{E}_{t_0}[\hat{q}_i] = \mathbf{E}_{t_0}[\sum_{t=t_0+1}^{\infty} \hat{q}_{i,t}] = (1+\lambda)(d_0 + \varphi \cdot q_{i,t_0}) \tag{13}$$

Notice that $\mathbf{E}_{t_0}[\hat{q}_{i,t}]$ is not monotonically increasing in the opportunities parameter $\pi$. Indeed, expected returns at time $t$ are increasing in the probability of finding an island only when the *actual* duration of exploration $t - t_0$ is less than its *expected* length $\tau = 1/\pi$. On the other hand, total expected returns are independent on $\pi$ as any explorer will eventually find a new island.

A straightforward comparison between *total* and *time-t expected* payoffs from exploration and returns to exploitations yields:

$$\mathbf{E}_{t_0}[\hat{q}_i] \geq q_{i,t_0} \Longleftrightarrow \begin{cases} q_{i,t_0} \leq \frac{1+\lambda}{1-(1+\lambda)\varphi} d_0 \; = \; c_1(\lambda, \varphi) \cdot d_0 \quad \text{if} \quad (1+\lambda)\varphi \leq 1 \\ always \quad \text{if} \quad (1+\lambda)\varphi > 1 \end{cases} \tag{14a}$$

$$\mathbf{E}_{t_0}[\hat{q}_{i,t}] \geq q_{i,t_0} \Longleftrightarrow \begin{cases} q_{i,t_0} \leq \frac{\tilde{\pi}(1+\lambda)}{1-\tilde{\pi}(1+\lambda)\varphi} \; = \; c_2(\lambda, \varphi, \pi, e) \cdot d_0 \quad \text{if} \quad (1+\lambda)\varphi\tilde{\pi} \leq 1 \\ always \quad \text{if} \quad (1+\lambda)\varphi\tilde{\pi} > 1 \end{cases} \tag{14b}$$

Hence, irrespective of the density of islands in the economy, *total expected returns* from exploration always overcome returns from exploitation whenever path-dependency and/or the likelihood of radical innovations are large enough - cf. Eq. (14a). As $\lambda \geq 1$, this is true whenever $\varphi \geq \frac{1}{2}$. Conversely, if $(1+\lambda)\varphi \leq 1$,

[62] Eqs. (12) and (13) hold because: (i) all random variables are stochastically independent; and, (ii) $\forall k = 1, 2, \ldots, \forall t : \Pr\{\hat{q}_{i,t+k} > 0 \mid \hat{q}_{i,t} > 0\} = 0$.

inequality $\mathbf{E}_{t_0}[\hat{q}_i] \geq q_{i,t_0}$ still holds if initial output is smaller than a multiple $c_1(\lambda, \varphi) \geq 0$ of $d_0$. As expected, when $\varphi$ and $\lambda$ are small enough, explorers get better payoffs the greater the distance from the origin of the starting island and the smaller initial output. Indeed, when path-dependency is weak, initial output does not strongly affect learning by doing ($\varphi$ small). Also, the likelihood of getting a higher productivity coefficient for a newly discovered island is larger the greater the distance from the origin of the island the explorer has started from. Finally, as $\partial c_1/\partial\lambda > 0$ and $\partial c_1/\partial\varphi > 0$, returns from exploration will be larger the more likely are radical innovations and the stronger path-dependency.

When *time-t expected returns* from exploration are compared to initial output - see Eq. (14b) - the duration of exploration $e+1 = t - t_0$ becomes a crucial variable. Indeed the probability $\tilde{\pi}$ of finding a new island at time $t$ (having left island $j_0$ at time $t_0$) is decreasing in $e$. However, $\tilde{\pi}$ is increasing in $\pi$ when the actual duration of exploration $(e+1)$ is smaller than the expected one $(1/\pi)$. Then, in addition to large values of $\varphi$ and $\lambda$, the inequality $\tilde{\pi}\varphi(1+\lambda) > 1$ is more likely to be satisfied the smaller $e$ (for a given $\pi$) and the more $\pi$ is sufficiently close to $1/(1+e)$ (for a given $e$). Intuitively, the time-$t$ net returns from exploration are always positive when the likelihood of finding an island at time $t$ (i.e. $\pi$) is large enough and when the duration of exploration is close to its expected value.

A straightforward consequence is that, when information diffusion is superimposed to the processes of exploitation, exploration and innovation, net returns from exploration are likely to be, on average, always positive - both in the long and in the short run - for a very large region of the parameter space. This is primarily due to the 'flatness' assumptions made about the probability structure of the payoff landscape faced by the explorers. In particular, for all $t$, the random output $\hat{q}_{i,t}$, given $\hat{q}_{i,k} = 0$, $k = t_0+1, \ldots, t-1$, will be equal to $(1+W)(d_0 + \varphi q_{i,t_0} + \varpi)$ with probability $\pi$ and 0 with probability $1-\pi$. Therefore, each explorer will on average face a constant (though path-dependent and individual-specific) payoff, no matter how large is $t$ and how far it will be from the island it started from.

An interesting question then arises, namely: Will the model still be able to generate patterns of self-sustaining growth when explorers have to face harder environments in which, e.g. opportunities are decreasing the more R&D is pushed ahead ?

To address this issue, let us assume that the probability of finding an island in the node $(x, y)$ is decreasing with the distance $d_1$ between $(x, y)$ and the origin, i.e.:

$$\pi(x,y) = \pi_0 \exp\{-\theta(x+y)\}, \quad \pi_0 \geq 0,\ \theta \geq 0 \tag{15}$$

This setup will generally reflect economies in which opportunities are decreasing in a globally shared measure of technologically distance computed with respect to some evaluation of the initial technological endowment. Though the actual payoff landscape faced by each explorer still is path-dependent and individual-specific, we can think of it as it if were the result of random drawings of $\hat{q}_{i,t}$ superimposed on an exogenously declining surface $\pi(x, y)$. Notice that, when $\theta = 0$, $\pi(x,y) = \pi_0$ as before, while, as $\theta$ grows, the probability of finding a new island will drop faster and faster for explorers searching far from the origin of the lattice.

Simulation exercises on this version of the model are presented in Figure 31. For a parameter setup yielding exponential growth in the case $\theta = 0$, we have initially investigated the effects of increasing the slope of the payoff landscape on the overall performance of the economy. As Panel (a) shows, MC mean of AGRs are declining as $\theta$ grows, but larger AGRs can be sustained by larger $\lambda$'s (and accordingly by larger $\varphi$'s and smaller $\rho$'s) even when explorers face very 'steep' environments. Moreover, the variation of GNP time series across independent simulations tends to shrink as $\theta$ increases, because the economy gradually converges toward zero-output realizations.

Nevertheless, for small $\theta$'s, the economy is still able to generate self-sustaining growth for meaningful parameter setups (see Panel (b)). As the slope of the payoff landscape gradually becomes steeper, other interesting growth patterns appear. When $\theta$ is not too large, the economy experiences either temporary $I(1)$ patterns discontinued by periods of stagnation and recession or initial phases of growth followed by cycles around a steady output thereafter. Finally, when it becomes inherently hard to introduce innovations as research is pushed far from the initial fundamentals, and interactions are local, GNP time-series progressively display negative trends characterized by increasing volatility (see Panel (c)).

Similar results also arise when one assumes that the probability of finding an island is individually decreasing with the duration of exploration. In particular, assume that if explorer $i$ leaves island at time $t_0$, then the probability of finding an island at time $t$ will be: $\pi_i(t, t_0) = \pi_0 \ \theta^{\ t-t_0}$ , $\pi_0 \geq 0$, $0 \leq \theta \leq 1$. In this specification, each explorer will face an individually-specific, locally decreasing opportunity landscape, the probability of finding a new island becoming smaller the longer the search is carried on. Notice that, even though we have chosen not to introduce any explicit selection dynamics (i.e. a mechanism governing the evolution of individual firms according to their revealed technological and market success), this alternative specification of the exploration-exploitation payoff structure would allow us to address selection issues in an indirect way.

### 7.3 Behavioral Heterogeneity, Individual Rationality and Collective Outcomes

Heterogeneity may arise in the model (at least) at three levels. First, technologies might (and do) differ both by their *ex-ante* productivity coefficients $s_j$ and, path-dependently, by their *ex-post* exploitation histories (i.e. the realizations of the stochastic processes $m_j(t)$). Second, agents might be characterized by different propensities to explore (i.e. distinct $\epsilon_i$, $i \in I$), given the decision rules described in Section 3. Third, agents might even be endowed by heterogeneous decision rules, e.g. some agents might hold 'more' rational strategies governing the choices of imitation, technological search, etc. .

In this section, we will first relax the assumption of a homogeneous propensity to commit to technological search and consider an economy in which an initial distribution $E = (\epsilon_1, \epsilon_2, \ldots, \epsilon_N)$, $\epsilon_i \in [0, 1]$ and $\epsilon_j \neq \epsilon_{j'}$ for some $j \neq j'$ is given. To keep things simple, let us suppose that $E$ is such that $\epsilon_i = 0$, $i = 1, 2, \ldots, \lfloor \mu N \rfloor$ and $\epsilon_i = \epsilon_0$, $i = \lfloor \mu N \rfloor + 1, \ldots, N$, where $\mu \in [0, 1]$ and $\epsilon_0 \in (0, 1]$. In other words, a share $\mu$ of the population is completely unwilling to leave (i.e. the 'sedentary' agents), while the remaining share $(1-\mu)$ behaves as before.

As intuition suggests, the aggregate consequences of increasing $\mu$'s (in terms of economy's AGRs) are once again strictly related to the eventual solution of the exploitation-exploration trade-off, arising in this case even more vividly. When the share of potential explorers declines, the AGRs schedule goes indeed through three distinct regimes, see Figure 32(a). If the sources of growth are very weak, the only way to avoid decreasing output levels is to commit completely to exploitation of existing resources. On the contrary, when they are very strong, the economy's performance is optimized at $\mu^*=0$ and AGRs decrease with $\mu$. In all intermediate setups (e.g. global interactions and no path-dependency), AGRs typically exhibit an interior maximum (cf. Section 5.1 and 5.2), with poor performances when the economy commits either small or too many resources in the exploration of unknown knowledge bases. Accordingly, when the size of the set of 'sedentary' agents grows, output time-series are likely to exhibit increasingly longer periods of steady GNP levels, cf. Figure 32(b). Hence, in the limit, the system is going to mimic the closed-ended, no-exploration case presented in Section 4.1, with an initial period of transitory growth followed by zero growth-rates thereafter (cf. also Section 7.2).

A higher-level behavioral heterogeneity can be however conceived. Indeed, given the extreme assumptions made on the rational bounds of the agents populating our economy, the model appears to highlight a few sources of potential conflict between individual and collective rationality. For instance, what would it happen if the population of our naïve entrepreneurs is injected by 'more rational' players, which behaves on the grounds of some (appropriately defined) expectations ?

In order to illustrate this point, consider the following simple example. Assume an economy characterized by: (i) constant technological returns to scale (i.e. $\alpha=1$); (ii) no knowledge diffusion (i.e. $\rho=\infty$); (iii) no path-dependency in innovation (i.e. $\varphi=0$); (iv) all $N$ agents working at time $t=0$ on a single island ($\ell_0=1$) with co-ordinates $(x^*,y^*)$ and initial productivity $s^*=x^*+y^*$; (v) a constant positive transportation cost $\beta>0$, which explorers have to pay in each period of their search.

We will consider two different settings for what concerns behavioral assumptions, namely: (a) the population is composed of $N$ agents behaving according to the rules defined in Section 3; and: (b) a 'representative individual' (RI), with unbounded computational skills and complete information, is introduced in the population. In particular, assume that the RI knows: (i) the co-ordinates $(x^*,y^*)$; (ii) the system parameters; (iii) the model of the economy. Although the RI is aware that, on average, the initial productivity of a new island is increasing in its distance from the origin, he does not know where new islands are actually located. Hence, starting from the node $(x,y)$, it will make use of a 'rational' exploration rule which gives equal probability to the nodes $(x+1,y)$ and $(x,y+1)$. Finally, assume for simplicity that the intertemporal discount rate is zero[63].

At time $t=1$, the problem for the RI is to decide whether to continue to extract the good at time $t=2$ or start to explore. In the first case, it will

[63] Our conjecture is that the following results will hold *a fortiori* for a strictly positive discount rate.

get a per-period net output from mining equal to $Q_M = s^*$. In the second case (see Section 7.2), the expected per-period net output from exploration will be: $Q_E = [(1+\lambda)(s^* + \tau) - \beta\tau]/\tau$, where $\tau = 1/\pi$ is the expected length of exploration (or, equivalently, the expected distance between $(x^*, y^*)$ and a new island). Then, the RI will decide to remain on island $(x^*, y^*)$ if and only if $Q_M > Q_E$, i.e. iff:

$$\pi < \frac{1}{1+\lambda} - \frac{1}{s^*} + \frac{\beta}{(1+\lambda)s^*} = \pi^*(\beta, \lambda, s^*) \tag{16}$$

As one can easily check, $\pi^*(\beta, \lambda, s^*)$ is increasing in $\beta$, decreasing in $\lambda$, and increasing in $s^*$ if $\lambda > \beta - 1$ (i.e. if opportunities are large enough compared to exploration costs). Notice that if $s^* \to \infty$ the RI will always stay on $(x^*, y^*)$, while if $\lambda \to \infty$ it will always leave.

Consider now, for given values of $s^*$, the set of $(\beta, \lambda)$ satisfying (16) for some $\pi \in (0, 1)$. In such a parameter region, the RI will decide to continue to work as a 'miner' and get a constant output $Q_M = s^*$. On the contrary, any economy characterized by the same $(\beta, \lambda, \pi, s^*)$ and composed of homogeneous agents behaving as in setup (a) above (cf. Section 3), will face a rather 'poor' environment, in which there is *neither* knowledge diffusion, *nor* path-dependency in innovation, *nor* increasing returns to scale. Furthermore, let us assume that our 'naïve' agents are characterized by a very low 'willingness to explore' (i.e. $\epsilon = 0.05$). Notwithstanding all that, as Figure 33 shows, the economy is able to get a per-capita net output persistently greater than $Q_M$.

Thus, even in this very simple setting, collective growth finds its necessary condition in the presence of a number of 'irrational' individuals. Even more so, the potential conflict between individual rationality and collective welfare emerges in the general setting with unlimited notional opportunities of exploration and transportation costs born up front by the 'explorers' themselves.

Note that, as mentioned earlier, this property significantly expands upon the common result from e.g. New Growth literature that in presence of externalities or dynamic increasing returns a systematic divergence between endogenously generated growth rates and socially optimal ones (whatever the latter means...) is likely to emerge. Here, one may require indeed the presence of straightforwardly *irrational* agents in order to have self-sustaining growth at all.

# 8 Conclusions

The paper presents a simple model in which self-sustaining growth endogenously emerges, under suitable technological and behavioral conditions, as the result of imperfect coordination among stylized, boundedly-rational, heterogeneous, firms which locally interact in an open-ended technological space and are able to modify the set of their 'nearest neighbors'.

Among other properties, the model shows that the very possibility of notionally unlimited (albeit unpredictable) technological opportunities is a necessary condition for patterns of persistently fluctuating exponential growth to be generated in the economy.

In that circumstance, self-sustaining growth is attained whenever technological opportunities (as captured by both the density of 'islands' $\pi$ and the mean of Poisson jumps to radically new paradigms $\lambda$), path-dependency (i.e. the fraction of idiosyncratic knowledge, $\varphi$, that agents are able to carry over to newly discovered technologies) and globality of interactions in the information diffusion process ($-\rho$), are beyond identifiable thresholds. In that region of the parameter space, the system self-organizes through subsequent phases of development and exhibits ordered GNP time-paths characterized by smaller growth-rates volatility.

Furthermore, the overall performance of the economy appears to be monotonically and linearly increasing in all sources of growth except the degree of globality of interactions, which engenders a strong threshold effect in the average growth rates of the system. The trade-off between exploitation of the fundamentals and exploration of the unknown - generated in the model by the strong uncertainty of the relative returns from R&D and production activities - clearly emerges, however, when one investigates how growth is affected by the propensity to explore ($\epsilon$) - or equivalently by the share of sedentary agents in the population ($\mu$) - and by the cumulativeness of learning ($\alpha$). In well-defined technological regimes, indeed, the system generates self-sustaining patterns of growth and higher overall performances *only if* a suitable balance between R&D and production is achieved.

As mentioned, the model could be considered as a sort of 'reduced form' evolutionary model, with an almost exclusive emphasis upon the learning/diffusion aspects of economic evolution, while repressing the competition/selection features of market interactions. Although the limitations stemming from this assumption are quite obvious (for example, the 'microeconomics' is bound to be rather poor), the model is nonetheless able to generate GNP time-series with statistical properties which robustly replicate some of the few unquestioned stylized facts of the 'business cycle' (e.g. GNP growth autocorrelation structure, persistence of fluctuations, etc.)[64] and, at the same time, to avoid drawbacks shared by many models of endogenous growth such as scale-effects (see Section 7.1). Moreover, self-organized patterns of exponential growth are not only attained in the system without appealing to the forecasting powers of any far-sighted 'representative agent', but, even stronger than that, the economy might require *non-average* (and individually irrational) behaviors in order to achieve such self-sustained growth paths[65].

As it stands, the model seems quite well suited to account for some generic properties of knowledge-driven growth. Nevertheless, further developments come easily to mind. First, one could try to see how the results presented here are modified by the introduction of an additional 'Keynesian' coordination problem affecting interdependent demand generation mechanisms. Second, one might likewise study the relevance of adding explicit selection processes affecting the frequency in the population (i.e. the size) of different agents which are 'carriers' of different technologies. Finally, once reasonable rules of interaction *between economies* are suitably defined, one could also investigate some 'convergence'

[64] Note also, that, in principle, the above variables and parameters can find empirical (although inevitably rough) proxies. Therefore, one might not despair to test the qualitative properties generated by the model against actual data.

[65] A similar point on non-average behaviors inducing symmetry breaks in the distribution of particular features or performances of a population of agents is in Allen (1988).

issues by simultaneously simulate multi-layer systems characterized by different initial conditions and system parameters.

However, even before all that come, it seems that the foregoing work might contribute to the understanding of how endogenous learning processes, with imperfect collective adaptation and heterogeneous agents, drive growth notwithstanding (or rather *because of*) the absence of fantastically rational agents and equilibria fulfilled throughout.

# References


[1] Aghion, P. and Howitt, P. (1992), "A model of growth through creative destruction", *Econometrica*, **60**: 323-351.

[2] Allen, P.M. (1988), "Evolution, innovation and economics", in Dosi, G., Freeman, C., Nelson, R.R., Silverberg, G. and Soete, L. (Eds.), *Technical Change and Economic Theory*, London, Francis Pinter.

[3] Allen, P.M. and McGlade, J.M. (1986), "Dynamics of discovery and exploitation : the case of the scotian shelf groundfish fisheries", *Canadian Journal of Fishery and Aquatic Sciences*, **43**: 1187-1200.

[4] Amable, B. (1995), "Endogenous Growth and Cycles Through Radical Incremental Inovation", CEPREMAP, Discussion Paper, 9504.

[5] Arrow, K.J. (1962), "The economic implications of learning by doing", *Review of Economic Studies*, **29**: 155-173.

[6] Arthur, W.B. (1989), "Competing technologies, increasing returns and lock-in by historical small events: the dynamics of allocation under increasing returns to scale", *Economic Journal*, **99**: 116-131.

[7] Arthur, W.B. (1993), "On designing artificial agents that behave like human agents", *Journal of Evolutionary Economics*, **3**: 1-22.

[8] Arthur, W.B. (1994), *Increasing returns and path-dependency in economics*, Ann Arbor, Univ. of Michigan Press.

[9] Arthur, W.B., Durlauf, S.N. and Lane, D. (Eds.) (1997), *The Economy as an Evolving Complex System II*, Santa Fe Institute, Addison Wesley.

[10] Atkinson, A. and Stiglitz, J.E. (1969), "A new view of technical change", *Economic Journal*, **79**: 573-578.

[11] Axtell, R. and Epstein, J. (1994), "Modeling artificial economies of adaptive agents", IIASA, Colloquium on Evolutionary Economics.

[12] Barro, R.J. and Sala-y-Martin, X. (1995), *Economic Growth*, New York, McGraw-Hill.

[13] Blanchard, O.J. and Fischer, S. (1989), *Lectures on macroeconomics*, Cambridge, MIT Press.

[14] Blanchard, O.J. and Quah, D. (1989), "The dynamic effects of aggregate demand and supply disturbances", *American Economic Review*, **79**: 655-673.

[15] Bloug, S.R. (1992), "The Relationship between Power and Level for Generic Unit Root Tests in Finite Samples", *Journal of Applied Econometrics*, **7**: 295-308.

[16] Brock, W.A. and Durlauf, S.N. (1995), "Discrete choice with social interactions I: Theory", University of Wisconsin at Madison, SSRI Working Paper, 9521.

[17] Campbell, J.Y. and Mankiw, N.G. (1987), "Are output fluctuations transitory ?", *Quarterly Journal of Economics*, **102**: 857-880.

[18] Campbell, J.Y. and Mankiw, N.G. (1989), "International evidences on the persistence of economic fluctuations", *Journal of Monetary Economics*, **23**: 319-333.

[19] Cheng, L. K. and Dinopoulos, E. (1991), "Stochastic Schumpeterian economic fluctuations", University of Florida, mimeo.

[20] Cheng, L.K. and Dinopoulos, E. (1992), "Schumpeterian growth and international business cycles", *American Economic Review*, **82**: 409-414.

[21] Chiaromonte, F. and Dosi, G. (1993), "Heterogeneity, competition and macroeconomic dynamics", *Structural Change and Economic Dynamics*, **4**: 39-63.

[22] Christiano, L.J. and Eichenbaum, M. (1989), "Unit roots in real GNP: Do we know and do we care ?, NBER, Working Paper, 3130.

[23] Clark, P.K. (1987), "The Cyclical Component of U.S. Economic Activity", *Quarterly Journal of Economics*, **102**: 797-814.

[24] Cochrane, J.H. (1988), "How big is the random walk in GNP*?*" *Journal of Political Economy*, **96**: 893-920.

[25] Cochrane, J.H. (1994), "Permanent and Transitory Components og GNP and Stock Prices", *Quarterly Journal of Economics*, **109**: 241-265.

[26] Cogley, T. and Nason, J.M. (1995), "Output Dynamics in Real-Business-Cycle Models", *American Economic Review*, **85**: 492-511.

[27] Cooper, R. and John, A. (1988), "Coordinating coordination failures in Keynesian models", *Quarterly Journal of Economics*, **103**: 441-464.

[28] David, P.A. (1975), *Technical choice, Innovation and Economic Growth*, Cambridge, Cambridge University Press.

[29] David, P.A. (1988), "Path-dependence: putting the past into the future of economics", Stanford, Institute for Mathematical Studies in Social Sciences, Technical Report, 533.

[30] David, P.A. (1992a), "Path-dependence and predictability in dynamic systems with local externalities : a paradigm for historical economics", in Foray, D. and Freeman, C. (Eds.), *Technology and the Wealth of Nations*, London, Pinter.

[31] David, P.A. (1992b), "Path-dependence in economic processes", Stanford, CEPR, Working Paper,

[32] Davidson, P. (1996), "Reality and economic theory", *Journal of Post-Keynesian Economics*, **18**: 479-508.

[33] DeJong, D.N. and Whiteman, C.H. (1991), "The case for trend-stationarity is stronger than we thought", *Journal of Applied Econometrics*, **6**: 413-421.

[34] DeJong, D.N. and Whiteman, C.H. (1994), "The forecasting attributes of trend- and difference-stationary representations for macroeconomic time-series", *Journal of Forecasting*, **13**: 279-297.

[35] Diebold, F.X. and Senhadji, A.S. (1996), "The uncertain unit root in real GNP: Comment", *American Economic Review*, **86**: 1291-1298.

[36] Dosi, G. (1982), "Technological paradigms and technological trajectories", *Research Policy*, **11**: 147-162.

[37] Dosi, G. (1988), "Sources, Procedures and Microeconomic Effects of Innovation", *Journal of Economic Literature*, **26**: 126-171.

[38] Dosi, G. (1991), "The research on innovation diffusion: an assessment", in Grubler, A. and Nakicenovic, N. (Eds.), *Diffusion of technologies and social behavior*, New York, Springer Verlag.

[39] Dosi, G. (1997), "Opportunities, Incentives and the Collective Patterns of Technological Change", *Economic Journal*, **107**: 1530-47.

[40] Dosi, G. and Egidi, M. (1991), "Substantive and procedural uncertainty. An exploration on economic behaviours in changing environments", *Journal of Evolutionary Economics*, **1**: 145-168.

[41] Dosi, G. and Fagiolo, G. (1998), "Exploring the unknown. On entrepreneurship, coordination and innovation-driven growth", in Lesourne, J. and Orléan, A. (Eds.), *Advances in Self-Organization and Evolutionary Economics*, Paris, Economica.

[42] Dosi, G. and Lovallo, D. (1997), "Rational entrepreneurs or optimistic martyrs ? Some considerations on technological regimes", in Garud, R., Nayyar, P. and Shapiro, Z. (Eds.), *Foresights and Oversights in Technological Change*, Cambridge, Cambridge University Press.

[43] Dosi, G. and Nelson, R.R. (1994), "An introduction to evolutionary theories in economics", *Journal of Evolutionary Economics*, **4**: 153-172.

[44] Dosi, G., Freeman, C. and Fabiani, S. (1994), "The process of economic development: introducing some stylized facts and theories on technologies, firms and institutions", *Industrial and Corporate Change*, **3**: 1-46.

[45] Dosi, G., Freeman, C., Nelson, R.R., Silverberg, G. and Soete, L. (Eds.) (1988), *Technical Change and Economic Theory*, London, Francis Pinter.

[46] Dosi, G., Marengo, L. and Fagiolo, G. (1996), "Learning in Evolutionary Environments", IIASA, Laxenburg, Austria, Working Paper, WP-96-124/Nov 1996.

[47] Durlauf, S.N. (1994), "Path-dependence in aggregate output", *Industrial and Corporate Change*, **3**: 149-171.

[48] Durlauf, S.N. and Quah, D.T. (1998), "The New Empirics of Economic Growth", Centre for Economic Performance, Discussion Paper, No. 384.

[49] Fagiolo, G. (1998), "Spatial interactions in dynamic decentralized economies: a review", in Cohendet, P., Llerena, P., Stahn, H. and Umbhauer, G. (Eds.), *The Economics of Networks. Interaction and Behaviours*, Berlin - Heidelberg, Springer Verlag.

[50] Fatas, A. (1995), "Endogenous growth and stochastic trends", INSEAD, Working Paper, 95/85/EPS.

[51] Foray, D. and Freeman, C. (1992), *Technology and the wealth of nations*, Paris, OECD.

[52] Freeman, C. (1982), *The Economics of Industrial Innovation*, London, Francis Pinter.

[53] Freeman, C. (1994), "The economics of technical change", *Cambridge Journal of Economics*, **18**: 463-514.

[54] Freeman, C. and Perez, C. (1988), "Structural crisis of adjustment: business cycles and investment behavior", in Dosi, G., Freeman, C., Nelson, R.R., Silverberg, G. and Soete, L. (Eds.), *Technical Change and Economic Theory*, London, Francis Pinter.

[55] Gagnon, J.E. (1988), "Short-Run Models and Long-Run Forecasts: A Note on the Permanence of Output Fluctuations", *Quarterly Journal of Economics*, **103**: 415-424.

[56] Granstrand, O. (Ed.) (1994), *The economics of technology*, Amsterdam, North Holland.

[57] Grossman, G.M. and Helpman, E. (1991a), *Innovation and growth in the global economy*, Cambridge, MIT Press.

[58] Grossman, G.M. and Helpman, E. (1991b), "Quality ladders in the theory of growth", *Review of Economic Studies*, **58**: 43-61.

[59] Grossman, G.M. and Helpman, E. (1991c), "Trade, Knowledge spillovers and growth", *European Economic Review*, **35**: 517-526.

[60] Hirschleifer, D. (1993), "The blind leading the blind : social influence, fads and informational cascades", UCLA, Los Angeles, Anderson Graduate School of Management, Working Paper, WP 24-93.

[61] Holland, J.H. (1975), *Adaptation in Natural and Artificial Systems*, Cambridge, MIT Press.

[62] Holland, J.H. and Miller, J.H. (1991), "Artificial adaptive agents in economic theory", *American Economic Review*, **81**: 365-370.

[63] Jones, C.I. (1995a), "R&D-based models of economic growth", *Journal of Political Economy*, **103**: 759-784.

[64] Jones, C.I. (1995b), "Time series tests of endogenous growth models", *Quarterly Journal of Economics*, **110**: 495-525.

[65] Jones, L.E. and Manuelli, R. (1990), "A convex model of equilibrium growth", *Journal of Political Economy*, **98**: 1008-1038.

[66] Jones, R. and Newman, G. (1994), "Economic growth as a coordination problem", University of British Columbia, Working Paper, 94/11.

[67] Jovanovic, B. (1995), "Learning and growth", NBER, Working Paper, 5383.

[68] Jovanovic, B. and McDonald, G.M. (1994), "Competitive diffusion", *Journal of Political Economy*, **102**: 24-52.

[69] Jovanovic, B. and Rob, R. (1989), "The growth and diffusion of knowledge", *Review of Economic Studies*, **56**: 569-582.

[70] Jovanovic, B. and Rob, R. (1990), "Long waves and short waves: growth through intensive and extensive search", *Econometrica*, **58**: 1391-1409.

[71] Kaldor, N. and Mirrlees, J.A. (1962), "A new model of economic growth", *Review of Economic Studies*, **29**: 174-192.

[72] Kemeny, J.G. and Snell, J.L. (1960), *Finite Markov Chains*, New York, Van Nostrand.

[73] Kirman, A.P. (1989), "The intrinsic limits of modern economic theory: the emperor has no clothes", *Economic Journal*, **99**: 126-139.

[74] Kirman, A.P. (1992), "Whom or what does the representative individual represent ?", *Journal of Economic Perspectives*, **6**: 117-136.

[75] Kirman, A.P. (1993), "Ants, rationality and recruitment", *Quarterly Journal of Economics*, **108**: 137-156.

[76] Kirman, A.P. (1997), "The Economy as an Evolving Network", *Journal of Evolutionary Economics*, **7**: 339-353.

[77] Kirman, A.P. (1998), "Economies with interacting agents", in Cohendet, P., Llerena, P., Stahn, H. and Umbhauer, G. (Eds.), *The Economics of Networks. Interaction and Behaviours*, Berlin - Heidelberg, Springer Verlag.

[78] Klenow, P.J. and Rodriguez-Clare, A. (1997), "Economic Growth: a Review Essay", *Journal of Monetary Economics*, **40**: 597-617.

[79] Kremer, M. (1993), "Population Growth and Technological Change: One Million B.C. to 1990", *Quarterly Journal of Economics*, **108**: 681-716.

[80] Krugman, P. (1996), *The self-organizing economy*, Cambridge, Blackwell.

[81] Kuran, T. (1988), "The tenacious past: theories of personal and collective conservatism", *Journal of Economic Behavior and Organization*, **10**: 143-171.

[82] Lane, D. (1993a), "Artificial worlds in economics : Part I", *Journal of Evolutionary Economics*, **3**: 89-107.

[83] Lane, D. (1993b), "Artificial worlds in economics : Part II", *Journal of Evolutionary Economics*, **3**: 177-197.

[84] Leijonhufvud, A. (1996), "Adaptive behavior, market processes and the computable approach", Center for Computable Economics, UCLA, mimeo.

[85] Levin, R.C., Klerovich, A., Nelson, R.R. and Winter, S.G. (1987), "Appropriating the returns from industrial research and development", *Brookings Papers on Economic Activity*, 783-820.

[86] Levinthal, D.A. and March, J.G. (1981), "A model of adaptive organizational search", *Journal of Economic Behavior and Organization*, **2**: 307-333.

[87] Levitt, B. and March, J.G. (1988), "Organizational learning", *Annual Review of Sociology*, **14**: 319-340.

[88] Lucas, R.E. (1988), "On the mechanics of economic development", *Journal of Monetary Economics*, **22**: 3-42.

[89] Lundvall, B.A. (1993), *National systems of innovation*, London, Francis Pinter.

[90] Malerba, F. and Orsenigo, L. (1993), "Technological regimes and firm behaviors", *Industrial and Corporate Change*, **2**: 54-71.

[91] March, J.G. (1991), "Exploration and Exploitation in Organizational Learning", *Organization Science*, **2** : 71-87.

[92] McGrattan, E.R. and Schmitz, J.A. (1998), "Explaining Cross-Country Income Differences", Federal Reserve Bank of Minneapolis, Staff Report, 250.

[93] Metcalfe, J. S. (1988), "Trade, technology and evolutionary change", Manchester, Manchester University, mimeo.

[94] Metcalfe, J. S. (1996), "Lectures on evolutionary theory", Manchester, Manchester University, mimeo.

[95] Nelson, C.R. and Plosser, C.I. (1982), "Trends and random walks in macroeconomic time series", *Journal of Monetary Economics*, **10**: 139-162.

[96] Nelson, R.R. (1981), "Research on productivity growth and differences", *Journal of Economic Literature*, **19**: 1029-1064.

[97] Nelson, R.R. (1995), "Recent Evolutionary Theorizing About Economic Change", *Journal of Economic Literature*, **33**: 48-90.

[98] Nelson, R.R. (1998), "The agenda for growth theory: a different point of view", *Cambridge Journal of Economics*, **22**: 497-520.

[99] Nelson, R.R. and Winter, S.G. (1977), "Simulation of Schumpeterian competition", *American Economic Review*, **67**: 271-276.

[100] Nelson, R.R. and Winter, S.G. (1982), *An Evolutionary Theory of Economic Change*, Cambridge, The Belknap Press of Harvard University Press.

[101] Orléan, A. (1992), "Contagion des opinions et fonctionnement des marchés financiers", *Revue Economiques*, **43**: 685-97.

[102] Parente, S.L. (1994), "Technology adoption, learning by doing and economic growth", *Journal of Economic Theory*, **63**: 346-369.

[103] Rebelo, S.T. (1991), "Long run policy analysis and long run growth", *Journal of Political Economy*, **99**: 500-521.

[104] Romer, D. (1996), *Advanced macroeconomics*, New York, McGraw Hill.

[105] Romer, P.M. (1986), "Increasing returns and long-run growth", *Journal of Political Economy*, **94**: 1002-1037.

[106] Romer, P.M. (1990), "Endogenous technological change", *Journal of Political Economy*, **98**: 71-102.

[107] Romer, P.M. (1994), "The origins of endogenous growth", *Journal of Economic Perspectives*, **8**: 3-22.

[108] Rosenberg, N. (1982), *Inside the Blackbox*, Cambridge, Cambridge University Press.

[109] Rosenberg, N. (1994), *Exploring the Black Box : Technology, Economics and History*, Cambridge, Cambridge University Press.

[110] Rudebusch, G.D. (1993), "The uncertain unit root in real GNP", *American Economic Review*, **83**: 264-272.

[111] Schumpeter, J.A. (1934), *The theory of economic development*, Cambridge, MA, Harvard University Press.

[112] Shackle, G.L.S. (1955), *Uncertainty in economics*, Cambridge, Cambridge University Press.

[113] Silverberg, G. and Verspagen, B. (1994), "Collective learning, innovation and growth in a boundedly rational, evolutionary world", *Journal of Evolutionary Economics*, **4**: 207-226.

[114] Silverberg, G. and Verspagen, B. (1995a), “An evolutionary model of long term cyclical variations of catching up and falling behind”, *Journal of Evolutionary Economics*, **5**: 209-227.

[115] Silverberg, G. and Verspagen, B. (1995b), “Evolutionary theorizing on economic growth”, IIASA, Laxenburg, Austria, Working Paper, WP-95-78.

[116] Silverberg, G. and Verspagen, B. (1995c), “From the artificial to the endogenous: modeling evolutionary adaptation and economic growth”, IIASA, Laxenburg, Austria, Working Paper, WP-95-08.

[117] Silverberg, G., Dosi, G. and Orsenigo, L. (1988), “Innovation, diversity and diffusion: a self-organization model”, *Economic Journal*, **98**: 1032-1054.

[118] Solow, R.M. (1960), “Investment and technical progress”, in Arrow, K.J., Korbin, S. and Suppes, P. (Eds.), *Mathematical methods in social sciences*, Stanford, Stanford University Press.

[119] Stiglitz, J.E. (1987), “Learning to learn, localized learning and technological progress”, in Dasgupta, P. and Stoneman, P. (Eds.), *Economic policy and technological performance*, Cambridge, Cambridge University Press.

[120] Stock, J.H. and Watson, M.W. (1986), “Does GNP have a unit root”, *Economic Letters*, **22**: 147-151.

[121] Stock, J.H. and Watson, M.W. (1988), “Variable Trends in Economic Time-Series”, *Journal of Economic Perspectives*, **2**: 147-174.

[122] Stoneman, P. (Ed.) (1995), *Handbook of the Economics of Innovation and Technological Change*, Oxford, Blackwell.

[123] Watson, M.W. (1986), “Univariate Detrending Methods with Stochastic Trends”, *Journal of Monetary Economics*, **18**: 49-75.

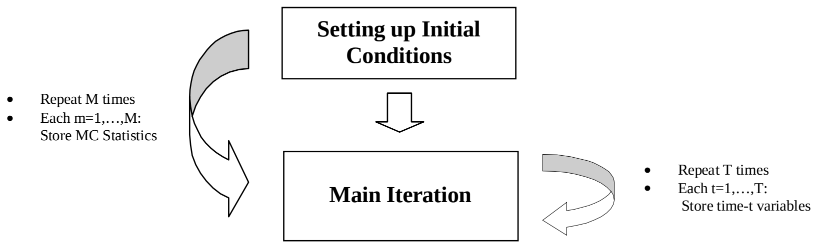
[124] Young, A. (1993), “Substitution and complementarity in endogenous innovation”, *Quarterly Journal of Economics*, **108**: 775-807.

# Appendix 1

## A1.1 System Variables of the Basic Model

| | |
|---|---|
| $\rho$ | Globality of Information Diffusion Process |
| $\varphi$ | Path-Dependency in Exploration-Discovery Processes |
| $\lambda$ | Mean of Poisson Random Variable W |
| $\pi$ | Probability that a Node is an Island |
| $\alpha$ | Returns to scale parameter in production |
| $\epsilon$ | Willingness to Explore |
| $N$ | Initial Size of the Population |
| $T$ | Econometric Sample Size |
| $M$ | Monte Carlo Sample Size |

## A1.2 Pseudo-Code of the Implemented Computer Program




Begin Proc **Initial Conditions**;
Let time period t=0;

Begin Proc **Generating Initial Islands**
    Define a square $Z$ in L$^2$ with vertices $(0, 0)$ and $(q, q)$;
    Do For each $x = 1, \ldots, q$ and $y = 1, \ldots, q$;
        Let each $(x, y) \in$ Z be an island with probability Pi;
    Call m = number of islands in Z after the assignation;
    Let island $\in \{1, \ldots$ ,m$\}$ index islands;
    Let [$x$(island),$y$(island)] be the coordinates of island 'island';
End Proc **Generating Initial Islands**;

Begin Proc **Assigning Productivity Coefficients**;
    Do For each island=1, . . . ,m;
        Let s(island) = s[$x$(island),$y$(island)] := $x$(island)+$y$(island);
End Proc **Assigning Productivity Coefficients**;

Begin Proc **Def. Econ. Frontier (EF) and Curr. Realized Econ. (CRE)**;
    Let $x$frontier = {max $x$(island)}, island $\in \{1, \ldots$ ,m$\}$;
    Let $y$frontier = {max $x$(island)}, island $\in \{1, \ldots$ ,m$\}$;
    Define the CRE as the Box {(0,0),($x$frontier, $y$frontier)};
End Proc **Def. Econ. Frontier (EF) and Curr. Realized Econ. (CRE)**;

```
Begin Proc Assigning Agents to Initial Islands;
    Do for each agent i={1, ,N};
        Assign at random island(i)∈{1,... ,m};
        Let status(i)=miner;
        Let agent i coordinates be x(i)=x(island), y(i)=y(island);
        Let [# of miners]=N, [#explorers]=0, [#imitators=0];
End Proc Assigning Agents to Initial Islands;

End Proc Initial Conditions;

Begin Proc Main Iteration;
Let t=t+1;
Do until t=T;

Begin Proc Updating #min., #expl., #imit., #min.(islands);
    #miners = Sum over all i such that status(i)=miner;
    #explorers = Sum over all i such that status(i)=explorer;
    #imitators = Sum over all i such that status(i)=imitator;
    Do for each island=1,... ,m;
        Compute #miners(island) = sum over all i={1,... ,N}
          such that status(i)=miner and island(i)=island;
End Proc Updating #min., #expl., #imit., #min.(islands);

Begin Proc Computing output and Islands' productivities;
    Do for each agent i={1, ,N} such that status(i)=miner;
        output(i) = s[island(i)]*[#miners(island(i))]^(Alpha-1);
    Do for each island=1,... ,m;
        output(island)=s(island)*[#miners(island)]^(Alpha);
        pr(island)= s(island)*[#miners(island)]^(Alpha-1);
    GNP = Sum[Q(island)], island=1,... ,m;
End Proc Computing output and Islands' productivities;

Begin Proc Miners Behavior;
    Do for each agent i={1, ,N} such that status(i)=miner;
        Begin Proc Assigning Memory to Miners;
            Memory(i)= Output(i);
        End Proc Assigning Memory to Miners;
        Begin Proc Miners -> Explorer change of status;
            Do with probability Epsilon;
                Let status(i)=explorer at t+1;
        End Proc Miners -> Explorer change of status;

End Proc Miners Behavior;

Begin Proc Explorers Behavior;
    Do for each agent i={1, ,N} such that status(i)=explorer;

        Begin Proc Explorer Move;
            Draw randomly one of the 4 adjacent nodes of x(i),y(i);
            Check positiviteness of new coordinates;
            Update accordingly x(i),y(i);
        End Proc Explorer Move;

        Begin Proc Checking whether explorer i is IN or out the CRE;
```

If it is IN then GO to Proc Known Island Discovery;
If it is OUT then GO to Proc New Island Discovery;
End Proc **Checking whether explorer i is IN or out the CRE**;

Begin Proc **Known Island Discovery**;
If $x$(i)=$x$(island), some island $j^* \in \{1, \ldots, m\}$ then do;
Let Status(i)= miner at t+1;
Let Island(i)= $j^*$ at t+1;
Let $x$(i)= $x$(island $j^*$), $y$(i)= $y$(island $j^*$) at t+1;
Else GO to Proc Information Diffusion;
End Proc **Known Island Discovery**;

Begin Proc **New Island Discovery**;
Do with probability pi;
Let $x$(i),$y$(i) be an island;
Let Island(i)= m+1 at t+1;
Let $x$(island m+1)= $x$(i);
Let $y$(island m+1)= $y$(i);
Let Status(i)= miner at t+1;

Begin Proc **Set Productivity Coeff. New Island**;
Draw W from a Poisson[Lambda];
Draw U from a 0-mean Rectangular Distr;
Define Dist(i)=$x$(i)+$y$(i);
Define PD(i) = Phi * Memory(i);
Let s(m+1)= (1+W)*[Dist(i)+ PD(i) + U];
End Proc **Set Productivity Coeff. New Island**;

Begin Proc **Updating EF and CRE**;
Let $x$frontier = {max $x$(island)}, island $\in \{1, \ldots, m\}$;
Let $y$frontier = {max $x$(island)}, island $\in \{1, \ldots, m\}$;
Update the CRE Box {(0,0),($x$frontier, $y$frontier)};
End Proc **Updating EF and CRE**;

Enlarge the set of currently known islands m=m+1;
End Proc **New Island Discovery**;

End Proc **Explorers Behavior**;

Begin Proc **Imitators Behavior**;
Do for each agent i={1, ,N} such that status(i)=imitator;
Call imitated(i) = index of island imitated by i;
Let dist(i, imitated(i))= dist(i, imitated(i)) - 1;
If dist(i, imitated(i))= 0 [i has reached the imitated island $j^*$] then do;
Let Status(i)= miner at t+1;
Let Island(i)= $j^*$ at t+1;
Let $x$(i)=$x$($j^*$) at t+1;
Let $y$(i)=$y$($j^*$) at t+1;
End Proc **Imitators Behavior**;

Begin Proc **Information Diffusion**;

Begin Proc **Signals Reception**;
Do for each agent i={1, ... ,N} such that status(i)$\in$ {miner, explorer};

Do for each island $=\{1,\ldots,$m$\}$ such that
#miners(island)<>0 and island <> island(i);
D(i;j)=Abs[$x$(i)-$x$(island j)]+Abs[$y$(i)-$y$(island j)];
Let agent i receive a signal from island j with probability:
int = [#miners(island)/#miners]*exp{-Rho* D(i;j)};
Compute the number si(i) of signals received by i;
Store the subset of indexes S(i)$\subseteq\{1,\ldots,$m$\}$, |S(i)|=si(i)
of islands which i has received a signal from;
Store the current pr(island), island $\in$ S(i);
End Proc **Signals Reception**;

Begin Proc **Signals Choice**;
Do for each agent i=$\{1,\ldots,$N$\}$ such that
status(i) $\in$ {miner, explorer} and si(i)>0;
Let $k_1$(i) the [si(i)+1]-vector whose first si(i) entries are
the elements of S(i) and whose [si(i)+1]-th entry is island(i);
Let $k_2$(i) the [si(i)+1]-vector whose h-th entry is pr(island),
island $\in$ S(i), h $\leq$ si(i), and whose [si(i)+1]-th entry is pr[island(i)];
Extract an index from $k_1$(i) with probabilities proportional
to the entries of $k_2$(i) and call it imitated(i);
If imitated(i) <> island(i) then do;
Let Status(i) = imitator at t+1;
Let dist(i, imitated(i)) = Abs[$x$(i)-$x$(imitated(i))] +
+ Abs[$y$(i)-$y$(imitated(i))];
Let imitated island = imitated(i);
If imitated(i) = island(i) then go to next i;
End Proc **Signals Choice;**

End Proc **Information Diffusion**;

Store all relevant time t variables;
End Proc **Main Iteration**;

# Appendix 2

## A2.1 Proof of Proposition in Section 4.1

Consider the stochastic process $\underline{M}_t = \{(M_{1t}, M_{2t}),\ t = 1, 2, \ldots\}$, where $M_{jt}$ is the r.v. "number of agents 'mining' on island $j$ at time $t$", $j \in \{1, 2\}$. Since it must hold that the total number of miners at the beginning of $t = 0$ is $N$ (i.e. $M_{10} + M_{20} = N$), we can assume as initial conditions some distribution $P_0$ on $M_{10}$ with state space $\{0, \ldots, N\}$ and let $P(\underline{M}_0) = P(M_{10}, N - M_{10}) = P_0(M_{10})$. The process $\underline{M}_t$ completely describes the economy, as the current number of imitators will simply be $E_t = N - (M_{1t} + M_{2t})$.

1. Define $L_{j,t}$, $j = 1, 2$ and $t \geq 0$, as the random variables: "number of miners working on island $j$ at (the beginning of) time $t$ who choose to imitate island $j\prime \in \{1, 2\}$, $j\prime \neq j$ ". Given stochastic independence in the two-stage procedure described in Section 3.5 - cf. in particular Eq. (10) - then it is straightforward to see that:

   (a) The probability distribution of $L_{j,t}$ conditional on $\underline{M}_t = (m_{1,t}, m_{2,t})$ is a binomial $(n, p)$ with parameters $n = m_{j,t}$ and:

$$
\begin{aligned}
p = p_{j',t} &= \\
&= \left[\frac{m_{j',t}}{m_{j,t} + m_{j',t}} e^{-\rho(s-1)}\right] \cdot \left[\frac{s_{j'}\ m^{\alpha}_{j',t}}{s_j\ m^{\alpha}_{j,t} + s_{j'}\ m^{\alpha}_{j',t}}\right] = \\
&= \frac{s_{j'}\ m^{\alpha+1}_{j',t}}{(m_{j,t} + m_{j',t})\left(s_j\ m^{\alpha}_{j,t} + s_{j'}\ m^{\alpha}_{j',t}\right)} e^{-\rho(s-1)}
\end{aligned}
\tag{A2.1}
$$

   where $(m_{j,t} + m_{j',t})$ and $s_j\ m^{\alpha}_{j,t} + s_{j'}\ m^{\alpha}_{j',t}$ are, respectively, the total number of miners and total GNP at time $t$; $p_{j',t}$ is simply the probability that a miner working at (the end of) time $t$ on island $j$ will decide to imitate island $j\prime$[66].

   (b) The random variables $L_{j\tau}|\underline{M}_{\tau}$ and $L_{j\prime\tau\prime}|\underline{M}_{\tau\prime}$ are stochastically independent for any $\tau \neq \tau\prime$, all $(j, j\prime) \in \{1, 2\}^2$, and so are $L_{jt}|\underline{M}_t$ and $L_{j't}|\underline{M}_t$ for $j \in \{1, 2\}$ and $j\prime \in \{1, 2\}$, $j\prime \neq j$.

   (c) The relation between $\underline{M}_t$ and $\underline{L}_{t-\tau} = (L_{1,t-\tau}, L_{2,t-\tau})$, $1 \leq \tau \leq t$, reads:

$$
\begin{aligned}
\Delta M_{j,t} &= M_{j,t} - M_{j,t-1} = -L_{j,t-1}, \quad t = 1, \ldots, s-1 \\
\Delta M_{j,t} &= M_{j,t} - M_{j,t-1} = L_{j\prime,t-s} - L_{j,t-1}, \quad t \geq s
\end{aligned}
\tag{A2.2}
$$

[66] Notice that in Eq. (A2.1) we have supposed that, once having received a signal, agents choose islands to imitate with probabilities proportional to $Q_t(x_j, y_j)$ and not to their productivities as assumed in Section 3.5. This scaling operation has been performed to deal with values of $\alpha \in [0, \infty]$ instead of $\alpha \in [1, \infty]$ and does not affect the results presented here. Moreover, in order to avoid a $0^0$ expression for $P(L_{j,t} = 0|m_{1,t}, m_{2,t})$ when $m_{j,t} = N$ (and then $p_{j\prime,t} = 0$), we have also assumed that $P(L_{j,t} = 0|m_{j,t} = N, m_{j\prime,t} = 0) = 1$.

Indeed, imitators take $s$ periods before they start producing on the chosen island. Then, in the first $s-1$ periods, the number of miners working on island $j$ at time $t$ will be simply equal to $M_{j,t-1} - L_{j,t-1}$. When $t \geq s$, however, imitators who left $j\prime$ at time $t-s$ have reached $j$. So $M_{j,t} = (M_{j,t-1} - L_{j,t-1}) + L_{j\prime,t-s}$.

Equations (A2.2) imply that $\underline{M}_t$ is a Markov process, as the conditional distribution will obey:

$$P(\underline{M}_t|\underline{M}_{t-h},\ h=1,2,\dots,t) = \begin{cases} P(\underline{M}_t|\underline{M}_{t-1}), & 1 \leq t \leq s-1 \\ P(\underline{M}_t|\underline{M}_{t-1}, \underline{M}_{t-s}), & t \geq s \end{cases} \tag{A2.3}$$

while, at any time $t \geq 1$, the joint distribution will be given by:

$$P(\underline{M}_t, \underline{M}_{t-1}, \dots, \underline{M}_0) = [\prod_{h=0}^{t-s} P(\underline{M}_{t-h}|\underline{M}_{t-h-1}, \underline{M}_{t-h-s})] \cdot \tag{A2.4}$$

$$\cdot [\prod_{h=1}^{s-1} P(\underline{M}_{s-h}|\underline{M}_{s-h-1})\ ] \cdot\ P(\underline{M}_0)$$

Moreover, the events: "a miner $i$ working on $j$ at time $t$ decides to imitate island $j\prime \in \{1,2\}$, $j\prime \neq j$ " are stochastically independent, for $t \geq s$, given $(\underline{M}_{t-1}, \underline{M}_{t-s})$ and, for $1 \leq t \leq s-1$, given $\underline{M}_{t-1}$. Then:

$$P(\underline{M}_t|\underline{M}_{t-1}) = P(M_{1,t}|\underline{M}_{t-1}) \cdot P(M_{2,t}|\underline{M}_{t-1}), \qquad 1 \leq t \leq s-1 \tag{A2.5}$$

$$\begin{gathered} P(\underline{M}_t|\underline{M}_{t-1}, \underline{M}_{t-s}) = \\ = P(M_{1,t}|\underline{M}_{t-1}, \underline{M}_{t-s}) \cdot P(M_{2,t}|\underline{M}_{t-1}, \underline{M}_{t-s}), \qquad t \geq s \end{gathered}$$

To complete the proof of point (1) it must be shown that the initial condition $M_{10} + M_{20} = N$ ensures that $M_t = M_{1t} + M_{2t} \leq N$, all $t \geq 1$. To see this, consider Eqs. (A2.2) and let $(M_{10}, M_{20}) = (m_{10}, N - m_{10})$. Then, by recursion, one gets:

$$M_t = \begin{cases} N - \sum_{h=0}^{t-1} \sum_{j=1,2} L_{j,h} & 1 \leq t \leq s-1 \\ \\ M_{s-1} + \sum_{h=0}^{s-2} \sum_{j=1,2} L_{j,h} + \sum_{h=0}^{s-2} \sum_{j=1,2} L_{j,t-h-1} & t \geq s \end{cases}$$

Hence, when $1 \leq t \leq s-1$, $N \geq M_{jt} \geq M_{jt-1}$, $j = 1,2$ and $M_t \leq N$. Moreover, when $t \geq s$:

$$M_t \leq M_{s-1} + \sum_{h=0}^{s-2} \sum_{j=1,2} L_{j,h} = M_{s-1} + N - M_{s-1} = N.$$

Before proving point (2), let us compute the distributions of $(M_{j,t}|\underline{M}_{t-1})$ , $1 \leq t \leq s-1$, and of $(M_{j,t}|\underline{M}_{t-1}, \underline{M}_{t-s})$, $t \geq s$. Let $\underline{M}_{t-1} = (m_{1,t-1}, m_{2,t-1})$ and $\underline{M}_{t-s} = (m_{1,t-s}, m_{2,t-s})$. Then, by Eqs. (A2.2), for any $\underline{m}= \{(m_1, m_2) : 0 \leq m_1 \leq N,\ 0 \leq m_2 \leq N,\ m_1 + m_2 \leq N\}$, and for $j \in \{1, 2\}$, $j\prime \in \{1, 2\}$, $j\prime \neq j$:

$$P(M_{j,t} = m_j|\underline{M}_{t-1}) = P(L_{j,t-1} = m_{j,t-1} - m_j|\underline{M}_{t-1}), \qquad 1 \leq t \leq s-1 \tag{A2.6}$$

$$\begin{gathered} P(M_{j,t} = m_j|\underline{M}_{t-1}, \underline{M}_{t-s}) = \\ = P(H_{j,j\prime}(t-1, t-s) = m_j - m_{j,t-1}|\underline{M}_{t-1}, \underline{M}_{t-s}), \qquad t \geq s \end{gathered}$$

where $H_{j,j\prime}(t-1, t-s) = L_{j\prime,t-s} - L_{j,t-1}$ is a random variable which, conditional on some feasible $\underline{M}_{t-1} = (m_{1,t-1}, m_{2,t-1})$ and $\underline{M}_{t-s} = (m_{1,t-s}, m_{2,t-s})$, has support $\{-m_{j,t-1}, \dots, -1, 0, +1, \dots, m_{j\prime,t-s}\}$ and is (conditionally) distributed as a difference between a $BIN(m_{j\prime,t-s}, p_{j,t-s})$ and a $BIN(m_{j,t-1}, p_{j\prime,t-1})$. Hence:

$$\begin{gathered} Prob\{M_{j,t} = k \mid \underline{M}_{t-1} = (m_{1,t-1}, m_{2,t-1}),\ \underline{M}_{t-s} = (m_{1,t-s}, m_{2,t-s})\} = \\ = (1 - p_{j,t-s})^{m_{j\prime,t-s}} (1 - p_{j\prime,t-1})^{m_{j,t-1}} \cdot \\ \cdot \sum_h \binom{m_{j,t-1}}{h} \binom{m_{j\prime,t-s}}{k+h} \left(\frac{p_{j\prime,t-1}}{1 - p_{j\prime,t-1}}\right)^h \left(\frac{p_{j,t-s}}{1 - p_{j,t-s}}\right)^{k+h} \end{gathered} \tag{A2.7}$$

where the sum is taken over $h = \{max\{0, -k\}, \dots, min\{m_{j,t-1}, m_{j\prime,t-s} - k\}\}$ and $p_{j,t}$'s are as in (A2.1).

2. When no 'best practice' is present in the economy, Eqs. (A2.2) and (A2.5) boil down to:

$$\begin{gathered} \Delta M_{j,t} = M_{j,t} - M_{j,t-1} = L_{j\prime,t-1} - L_{j,t-1} \\ P(\underline{M}_t|\underline{M}_{t-1}, \underline{M}_{t-s}) = P(M_{1,t}|\underline{M}_{t-1}) \cdot\ P(M_{2,t}|\underline{M}_{t-1}) \end{gathered} \tag{A2.8}$$

for all $t \geq 1$. As $M_{1t} + M_{2t} \equiv N$, any $t \geq 1$, the behavior of the system for all $t \geq 0$ is described by the scalar r.v. $M_t = M_{1t}$. For a given initial distribution $P(M_0)$, the stochastic process $\{M_t, t \geq 1\}$ is a stationary Markov chain with discrete state space $\{0, 1, \dots, N-1, N\}$, as (A2.8) implies that $P(M_t) = P(M_{t-1}) \cdot\ P(M_t|M_{t-1})$, $t \geq 1$. By (A2.6), we know that: $p(m\prime|m) = P(M_t = m\prime|M_{t-1} = m) = P(H_{12}(t-1) = m\prime - m)$; and that $H_{12}(t-1)$ - conditionally on $M_{t-1} = m$ - is the difference between a $BIN(N-m, p_1)$ and a $BIN(m, p_2)$, where:

$$(p_1, p_2) = \left(\frac{m^{\alpha+1}}{N \cdot (m^\alpha + (N-m)^\alpha)}, \frac{(N-m)^{\alpha+1}}{N \cdot (m^\alpha + (N-m)^\alpha)}\right)$$

Hence $p(m\prime|m) = 1$ if and only if $(m\prime, m) \in \{(0, 0), (N, N)\}$, while $p(m\prime|m) = 0$ if and only if $m \in \{0, N\}$ and $m\prime \in \{1, \dots, N-1\}$. After a standard row/column permutation, the $(N+1) \times (N+1)$ transition matrix $\mathbf{P}$ gets the form:

$$\mathbf{P} = \begin{pmatrix} \mathbf{I}_{2\times 2} & \mathbf{0}_{2\times(N-1)} \\ \mathbf{R}_{(N-1)\times 2} & \mathbf{Q}_{(N-1)\times(N-1)} \end{pmatrix}$$

where the $m$-th row of $\mathbf{R}$ is $\underline{r}_m = [p(0|m), p(N|m)] = [p_1^{N-m}, p_2^m]$, $m \in \{1, \ldots, N-1\}$, and the generic $(m, m\prime)$-entry of $\mathbf{Q}$, for $(m, m\prime) \in \{1, \ldots, N-1\} \times \{1, \ldots, N-1\}$, is given by (see A2.7):

$$p(m\prime|m) = (1-p_1)^{N-m}(1-p_2)^m \cdot$$
$$\cdot \sum_{h=\max\{0,m-m'\}}^{\min\{m,N-m'\}} \binom{m}{h}\binom{N-m}{m'-m+h}\left(\frac{p_2}{1-p_2}\right)^h \left(\frac{p_1}{1-p_1}\right)^{m'-m+h}$$

Hence the Markov chain $M_t$ is not irreducible (because $\mathbf{P}$ is not regular) and the process displays two absorbing states $X_R = \{m_+, m_-\} = \{0, N\}$. The latter are the only recurrent states while $X_T = \{1, \ldots, N-1\}$ are all transient. This implies that the process will be absorbed almost surely. Finally, consider the $(N-1) \times 2$ matrix $\mathbf{B}$ whose generic row is $[b_0(m), b_N(m)]$, $m \in \{1, \ldots, N-1\}$ and $b_{m^*}(m)$ is the probability of being absorbed by the recurrent state $m^* \in \{0, N\}$ starting at time 0 from $m$. Then, standard results - cf. Kemeny and Snell (1960) - allows us to conclude that: $\mathbf{B} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{R}$.

3. When $s > 1$, $\underline{M}_t$ is no longer a stationary Markov process, as transition probabilities for $1 \leq t \leq s-1$ differ from those for $t \geq s$. In particular, given $\underline{M}_{t-1} = (m_{1,t-1}, m_{2,t-1})$ and $\underline{M}_{t-s} = (m_{1,t-s}, m_{2,t-s})$, then if $1 \leq t \leq s-1$ (A2.9):

$$\Delta M_{1,t}|\underline{M}_{t-1} \overset{d}{\sim} -BIN(m_{1,t-1}, \frac{s m_{2,t-1}^{\alpha+1} e^{-\rho(s-1)}}{(m_{1,t-1}+m_{2,t-1})\left(m_{1,t-1}^{\alpha}+s m_{2,t-1}^{\alpha}\right)})$$

$$\Delta M_{2,t}|\underline{M}_{t-1} \overset{d}{\sim} -BIN(m_{2,t-1}, \frac{m_{1,t-1}^{\alpha+1} e^{-\rho(s-1)}}{(m_{1,t-1}+m_{2,t-1})\left(m_{1,t-1}^{\alpha}+s m_{2,t-1}^{\alpha}\right)})$$

while, for $t \geq s$ (A2.10):

$$\Delta M_{1,t}|\underline{M}_{t-1}, \underline{M}_{t-s} \overset{d}{\sim} BIN(m_{2,t-s}, \frac{m_{1,t-s}^{\alpha+1} e^{-\rho(s-1)}}{(m_{1,t-s}+m_{2,t-s})\left(m_{1,t-s}^{\alpha}+s m_{2,t-s}^{\alpha}\right)}) +$$
$$-BIN(m_{1,t-1}, \frac{s m_{2,t-1}^{\alpha+1} e^{-\rho(s-1)}}{(m_{1,t-1}+m_{2,t-1})\left(m_{1,t-1}^{\alpha}+s m_{2,t-1}^{\alpha}\right)})$$

$$\Delta M_{2,t}|\underline{M}_{t-1}, \underline{M}_{t-s} \overset{d}{\sim} BIN(m_{1,t-s}, \frac{s m_{2,t-s}^{\alpha+1} e^{-\rho(s-1)}}{(m_{1,t-s}+m_{2,t-s})\left(m_{1,t-s}^{\alpha}+s m_{2,t-s}^{\alpha}\right)}) +$$
$$-BIN(m_{2,t-1}, \frac{m_{1,t-1}^{\alpha+1} e^{-\rho(s-1)}}{(m_{1,t-1}+m_{2,t-1})\left(m_{1,t-1}^{\alpha}+s m_{2,t-1}^{\alpha}\right)})$$

where $\overset{d}{\sim}$ stands for 'is distributed as'. To show that $\underline{M}_t$ will almost surely be absorbed by the states $\underline{m}_+ = (0, N)$ and $\underline{m}_- = (N, 0)$, let us start to note that, as $M_{1,t}+M_{2,t} \leq N$ for all $t \geq 1$, then if $M_{1,t-s} = N$ (or symmetrically $M_{2,t-s} = N$) for some $t-s \geq 0$, then $M_{2,t-s} = 0$ (or symmetrically $M_{1,t-s} = 0$). Second, Eqs. (A2.2) imply that (i) $P(\Delta M_{j,t} \leq 0|\underline{M}_{t-1}) = 1$, $j = 1, 2$ when $1 \leq t \leq s-1$; and (ii) if $j\prime \in \{1, 2\}$, $j\prime \neq j$, then for $j = 1, 2$ and $1 \leq t \leq s-1$:

$$M_{j,t} = M_{j,0} - [L_{j,0} + \ldots + L_{j,t-1}],$$

while for $j = 1, 2$ and $t \geq s$ :

$$M_{j,t} = M_{j,0} - [L_{j,0} + L_{j,1} + \ldots + L_{j,t-1}] + [L_{j\prime,0} + L_{j\prime 1} + \ldots + L_{j\prime,t-s}].$$

Hence, if $t \geq 2s - 1$:

$$\begin{aligned} M_{1,t} &= M_{1,t-s} - [L_{1,t-s} + L_{2,t-s}] - [L_{1,t-1} + \ldots + L_{1,t-s+1}] + \\ &\quad +[L_{2,t-s-1} + \ldots + L_{2,t-2s+1}] \\ M_{2,t} &= M_{2,t-s} - [L_{1,t-s} + L_{2,t-s}] - [L_{2,t-1} + \ldots + L_{2,t-s+1}] + \\ &\quad +[L_{1,t-s-1} + \ldots + L_{1,t-2s+1}] \end{aligned} \tag{A2.11}$$

Consequently, if for some $t$, $M_{1,t-s} = N$ (and so $M_{2,t-s} = 0$), then $L_{j,t} = 0$ for $t = t-2s+1, \ldots, t-1$ and $j = 1, 2$. This shows that, also in the case $s > 1$, the states $\underline{m}_+ = (0, N)$ and $\underline{m}_- = (N, 0)$ are absorbing. In fact, by (A2.11), if $t$ is sufficiently large, then for all $\tau \geq 1$ (A2.12):

$$Prob\{(M_{1,t-s+\tau}, M_{2,t-s+\tau}) = (N, 0)|(M_{1,t-s}, M_{2,t-s}) = (N, 0)\} = 1$$

$$Prob\{(M_{1,t-s+\tau}, M_{2,t-s+\tau}) = (0, N)|(M_{1,t-s}, M_{2,t-s}) = (0, N)\} = 1$$

Intuitively, if all miners have to be working on island 1 at time $t - s$, then nobody could leave any islands during the $s$ time periods before $t-s$ and during the s time period after $t - s$. All other states are transient, as one can easily see by directly computing probabilities in (A2.7), for every 4-tuple $(m_{1,t-1}, m_{2,t-1}, m_{1,t-s}, m_{2,t-s}) \in \{0, 1, \ldots, N\}^4$ such that $m_{1,t-1} + m_{2,t-1} \leq N$ and $m_{1,t-s} + m_{2,t-s} \leq N$.

### A2.2 The Average Behavior of the Economy

Equations (A2.6) and (A2.7) jointly imply that, for any $j \in \{1, 2\}$ and $j\prime \in \{1, 2\}$, $j\prime \neq j$:

$$\begin{aligned} \mathbf{E}(M_{j,t}|\underline{M}_{t-1}) &= m_{j,t-1} - m_{j,t-1}p_{j\prime,t-1},\ 1 \leq t \leq s-1 \\ \mathbf{E}(M_{j,t}|\underline{M}_{t-1}, \underline{M}_{t-s}) &= m_{j,t-1} + m_{j\prime,t-s}p_{j,t-s} - m_{j,t-1}p_{j\prime,t-1},\ t \geq s \end{aligned} \tag{A2.13}$$

Let us then define the (conditional) expected share of worker on island $j$ as $x_{j,t} = m_{j,t}/N$, $j \in \{1,2\}$. Assume also that $N$ is large enough, so that $x_{j,t} \in Z[0,1]$. Substituting (A2.1) into (A2.13) yields, for $1 \leq t \leq s-1$:

$$\begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} = \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} - e^{-\rho(s-1)} \begin{bmatrix} \frac{s x_{1,t-1} x_{2,t-1}^{\alpha+1}}{(x_{1,t-1}+x_{2,t-1})(x_{1,t-1}^{\alpha}+s x_{2,t-1}^{\alpha})} \\ \frac{x_{1,t-1}^{\alpha+1} x_{2,t-1}}{(x_{1,t-1}+x_{2,t-1})(x_{1,t-1}^{\alpha}+s x_{2,t-1}^{\alpha})} \end{bmatrix}, \tag{A2.14a}$$

and, for $t \geq s$:

$$\begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} = \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} + \tag{A2.14b}$$

$$+e^{-\rho(s-1)} \begin{bmatrix} \frac{x_{1,t-s}^{\alpha+1} x_{2,t-s}}{(x_{1,t-s}+x_{2,t-s})(x_{1,t-s}^{\alpha}+s x_{2,t-s}^{\alpha})} - \frac{s x_{1,t-1} x_{2,t-1}^{\alpha+1}}{(x_{1,t-1}+x_{2,t-1})(x_{1,t-1}^{\alpha}+s x_{2,t-1}^{\alpha})} \\ \frac{s x_{1,t-s} x_{2,t-s}^{\alpha+1}}{(x_{1,t-s}+x_{2,t-s})(x_{1,t-s}^{\alpha}+s x_{2,t-s}^{\alpha})} - \frac{x_{1,t-1}^{\alpha+1} x_{2,t-1}}{(x_{1,t-1}+x_{2,t-1})(x_{1,t-1}^{\alpha}+s x_{2,t-1}^{\alpha})} \end{bmatrix}.$$

Given some fixed initial conditions $x_0 = x_{10} \in Z[0,1]$ - and $x_{20} = 1 - x_{10}$ - the system of non-linear difference equations (A2.14) describes, for all $t \geq 0$, the behavior of the expected number of miners on islands 1 and 2, conditional on their past choices.

In the following, we will briefly characterize the limit behavior of $\underline{x_t} = (x_{1t}, x_{2t})$ as a function of both system parameters $(s, \alpha, \rho)$ and initial conditions $x_0$.

*A2.2.a The case s=1*

If $s = 1$, then $x_{2t} = 1 - x_{1t} = 1 - x_t$, all $t \geq 0$. Therefore (A2.14a,b) collapse into the following 1-dimensional nonlinear difference equation:

$$x_t = x_{t-1} + x_{t-1}(1 - x_{t-1}) \frac{x_{t-1}^{\alpha} - (1 - x_{t-1})^{\alpha}}{x_{t-1}^{\alpha} + (1 - x_{t-1})^{\alpha}} = h(x_{t-1}; a) \tag{A2.15}$$

for $a \geq 0$ and initial condition $x_0 \in [0,1]$.

It is easy to see that the following facts about $h$ are true: (i) $h : Z[0,1] \rightarrow [0,1]$, all $\alpha \geq 0$; (ii) $h(0;\cdot) = 0$, $h(1;\cdot) = 1$, $h(\frac{1}{2};\cdot) = \frac{1}{2}$; (iii) $h(0;\cdot)$ is $C^2$ in $Z[0,1]$; (iv)$\partial h/\partial x_{t-1} \geq 0$ all $x_{t-1} \in [0,1]$; (v) $x_t - h(x_{t-1};\alpha) \geq 0$ for $x_{t-1} \leq \frac{1}{2}$ and $x_t - h(x_{t-1};\alpha) \geq 0$ for $x_{t-1} \geq \frac{1}{2}$; and, finally, (vi) $\partial^2 h/\partial x^2 \geq 0$, $x \leq \frac{1}{2}$ and $\partial^2 h/\partial x^2 \leq 0$, $x \geq \frac{1}{2}$. Hence, the nonlinear difference equation (A2.15) exhibits three fixed points $\{0, \frac{1}{2}, 1\}$, but only $\{0, 1\}$ are stable. The economy will converge to the point $x_+ = 1$ [conversely, $x_- = 0$] if $x_0 > \frac{1}{2}$ [conversely, if $x_0 < \frac{1}{2}$], while no dynamics will arise only if $x_0 = \frac{1}{2}$.

The parameter $\alpha$ will affect the absolute value of the rate of convergence:

$$\delta_t = \frac{|\Delta x_t|}{x_{t-1}} = (1 - x_{t-1})\frac{|x_{t-1}^{\alpha} - (1 - x_{t-1})^{\alpha}|}{x_{t-1}^{\alpha} + (1 - x_{t-1})^{\alpha}}.$$

Indeed, $\delta_t$ is strictly increasing in $\alpha$ for $x_{t-1} \geq \frac{1}{2}$ (i.e. when $x_t$ is increasing toward 1) and strictly decreasing for $x_{t-1} \leq \frac{1}{2}$ (i.e. when $x_t$ is decreasing toward 0). When $\alpha = 0$ (no production), then $x_t = h(x_{t-1}; 0) = x_{t-1}$ and any initial condition is stable. When $0 < \alpha < 1$ the production exhibits decreasing returns to scale at the islands' level, so that the absolute rate of convergence is very low. When the production technology is linear ($\alpha = 1$), then $k(x_{t-1}; 1) = (1 - x_{t-1})(2x_{t-1} - 1)$. As $\alpha$ increases then the shape of the function $h$ becomes more curved. Finally, as $a \to \infty$, $h$ converges to:

$$h^*(x_{t-1}) = \lim_{\alpha \to \infty} h(x_{t-1}; a) = \begin{cases} x_{t-1}^2, & 0 \leq x_{t-1} < \frac{1}{2} \\ 2x_{t-1} - x_{t-1}^2, & \frac{1}{2} \leq x_{t-1} \leq 1 \end{cases}$$

which is not defined at $x_{t-1} = \frac{1}{2}$. Indeed, $h^*(\frac{1}{2}^-) = \frac{1}{4}$ and $h^*(\frac{1}{2}^+) = \frac{3}{4}$. In this case, the behavior of the system in undetermined as long as one starts with initial conditions $x_0 = \frac{1}{2}$.

*A2.2.b The case s>1*

If $s > 1$, the two-dimensional nonlinear difference equation (A2.14) must be analyzed numerically. First, it is easy to see (by inspection) that (A2.14) still displays two stable fixed points, i.e. $\{(1, 0), (0, 1)\}$. However, there exist also an interior unstable fixed point $\{x^*(s, \alpha, \rho), 1 - x^*(s, \alpha, \rho)\}$, which is a function of the system parameters. No other limit behavior arises, as bifurcation diagrams show.

For any $s = 2, 3, \ldots$, the function $x^*(s, \alpha, \rho)$ partitions the two-dimensional parameter space $Z[0, 1] \times \Re_+$ into an 'efficient' region (i.e. such that $\lim_{t\to\infty} x_{1t} = 0$) and an 'inefficient' one (i.e. such that $\lim_{t\to\infty} x_{1t} = 1$). Figure A2.1 shows the projection of the bifurcation diagram onto the plane $(\alpha, x_0)$: for any $\alpha$, the curves display the value of $x_0$ such that $\Delta x_t = 0$, all $t \geq 1$. As the intuition suggests, the inefficient region shrinks as $\rho$ increases for $\alpha$ given $s$ - Panel (a) - and as $s$ increases for a given $\rho$ - Panel (b). In Figure A2.2 the function $x^*(s, \alpha, \rho)$ is plotted in the case $s = 2$ (its shape does not change as $s$ varies). Notice that, as $\alpha \to \infty$, $x^*(s, \alpha, \rho)$ converges to $\frac{1}{2}$, as in the case $s = 1$. Indeed, when returns to scale are infinitely strong, imitation solely drives the dynamics ($s$ has no effects) and the system either converges to island 1 or 2 depending on $x_0 > \frac{1}{2}$ or $x_0 < \frac{1}{2}$. Conversely, when $\alpha \to 0$ (no production), $x^*(s, \alpha, \rho) \to 1$, as, with no imitation, any technological gap $s > 0$ will imply convergence toward the efficient technology. Finally, $x^*(s, \alpha, \rho)$ appears to be increasing both in $\alpha$ and $\rho$, but the influence of the latter is very weak.

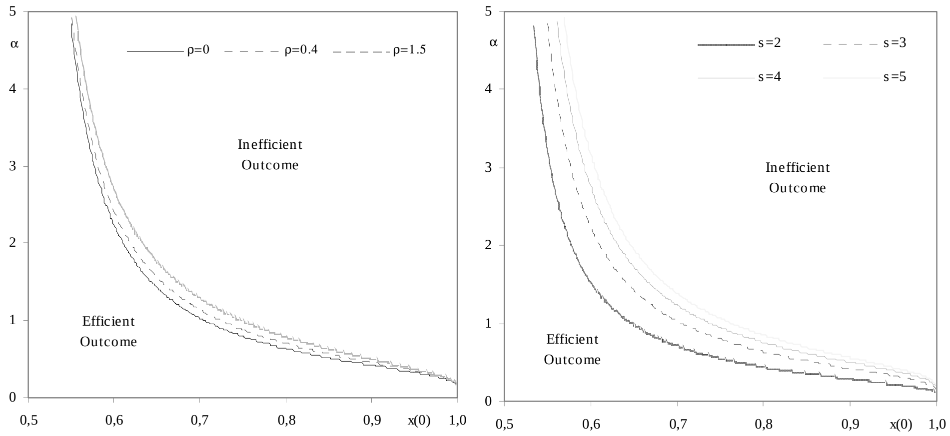



**Figure A2.1**: No Exploration Model. Bifurcation Diagrams for Dynamical Systems governing the Average Behavior of the System as a Function of the Proportion of Initial Miners on the Inefficient Island ($x_0$) and Returns to Scale ($\alpha$). Left Panel (a): Increasing the Locality of Information Diffusion (s=3). Right Panel (b) Increasing the Technological Gap between Islands (s=0.1)(North-East Portion: Convergence toward Island 1; South-West Portion: Convergence toward Island 2)

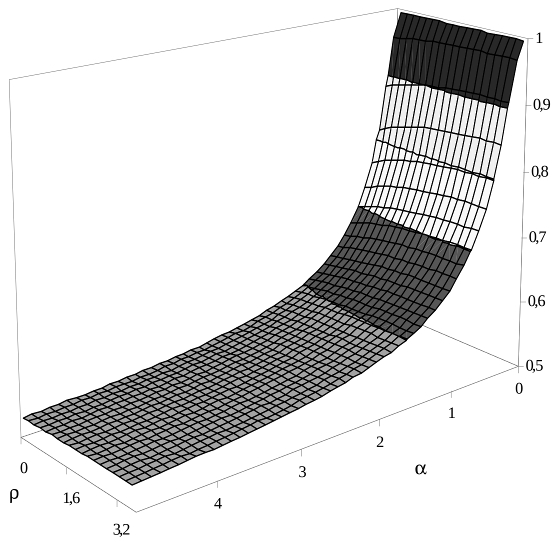



**Figure A2.2:** No Exploration Model. Numerical Analysis of the Unstable Fixed Points of the Average Behavior. The Function $x^*(s, \alpha, \rho)$ for $s = 2$.

# Appendix 3

All Montecarlo (MC) studies presented in Section 5, 6 and 7 refer to the following type of experiment. Let $\underline{\xi}(\omega) = [\xi(\omega, 0), \xi(\omega, 1), \ldots, \xi(\omega, t), \ldots, \xi(\omega, T)]'$ be some time-series generated by the model under the parametrization $\omega \in \Omega \equiv \{(\rho, \varphi, \lambda, \pi, \alpha, \epsilon, N, T) \in \Re^3_+ \times [0,1]^3 \times \aleph^2\}$ and consider the $(T+1) \times M$ matrix $\Xi(\omega)$ whose columns are made by $M$ independent replication of $\underline{\xi}(\omega)$. Then, given some statistics $S(\underline{\xi}(\omega))$, we are interested in assessing how the moments of the MC distribution $\Sigma(\omega) = \{S(\underline{\xi}_m(\omega)),\ m = 1, \ldots, M\}$, especially mean and variance, depend on $\omega$.

The goal of such exercise is to roughly estimate how the moments of interest (computed with respect to the true but unknown data generating process) changes as one moves across the parameter space. Consider, for instance, the case in which $S(\underline{q}_m(\omega)) = g_m(\omega) = [(q_{m,T}/q_{m,0})^{1/(T+1)} - 1]$, i.e. the AGR of the economy over $T+1$ periods. To understand how different behavioral and system parameters affect the average performance of the economy, one can employ, by the analogy principle, the MC sample mean:

$$\overline{g}_M(\omega) = M^{-1} \sum_{m=1}^{M} g_m(\omega)$$

to estimate the underlying relation: $\omega \longmapsto \mathbf{E}[g_m(\omega)]$; and the MC sample variance:

$$\overline{\sigma}^2_{g_M}(\omega) = M^{-1} \sum_{m=1}^{M} [g_m(\omega)]^2 - [\overline{g}_M(\omega)]^2$$

to study both the reliability of $\overline{g}_M(\omega)$ as an estimator of $\mathbf{E}[g_m(\omega)]$ and, more importantly, to assess how system parameters affect the variability of the economy's performance across independent simulations.

In general, all MC experiments have been carried on over $M = 10000$ independent simulations. This choice of $M$ has been suggested by two related observations. First, MC distributions become sufficiently symmetric for $M > 1000$ (see Panel (a) of Figure A3.1 for an example with AGRs) so that one may definitely employ to estimate $\mathbf{E}[g_m(\omega)]$ in the majority of parametric setups. However, when the system is fueled with high opportunities (i.e. both $\lambda$ and $\pi$ very large), AGR distributions (as well as those of other statistics) display some asymmetry with heavy right tails, due to the increasing likelihood of exceptional discoveries. In all those cases, both the median and the mean of the MC distributions have been computed and plotted against system parameters. Yet, all the main results presented in the paper are not affected by the choice of the statistics when $M >> 1000$. Second, recursive analyses have been performed to assess if, and how fast, MC moments (of any order) converge toward a stable value. For a sufficiently large sample of the relevant regions in the parameter space, sample moments of MC distributions over the first $M^*$ simulations (where $M^* = M_0, M_0 + 1, \ldots$ ) have been plotted against $M^*$. In all cases, one observes convergence after a number of MC replications well below $M = 5000$, cf. Figure A3.2 for an example with the MC AGRs first four moments in a high-opportunities setup.

Finally, as long as AGRs experiments are concerned, it is worth noting that all results presented in the paper are computed by replacing $q_{m,T}$ by:

$$\tilde{q}_{m,T}(\upsilon) = \upsilon^{-1} \sum_{\tau=0}^{\upsilon-1} q_{m,T-\tau}$$

in the expression $g_m(\omega) = [(q_{m,T}/q_{m,0})^{1/(T+1)} - 1]$. This has been done in order to reduce the dependence of $g_m(\omega)$ on $T$. Usually, we set $\upsilon = 10$.

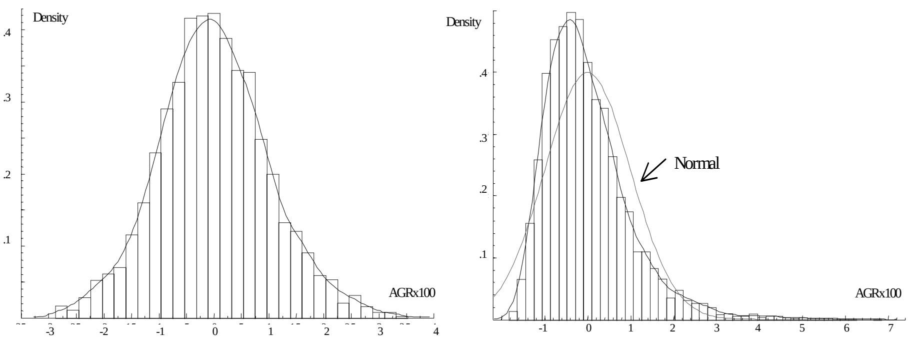



Figure A3.1: Two Examples of Montecarlo Distributions of AGR across M=1000 simulations. Panel (a) on the left: Low Opportunities ($\lambda = 1, \pi = 0.1$). Panel (b) on the right: High Opportunities ($\lambda = 5, \pi = 0.4$).

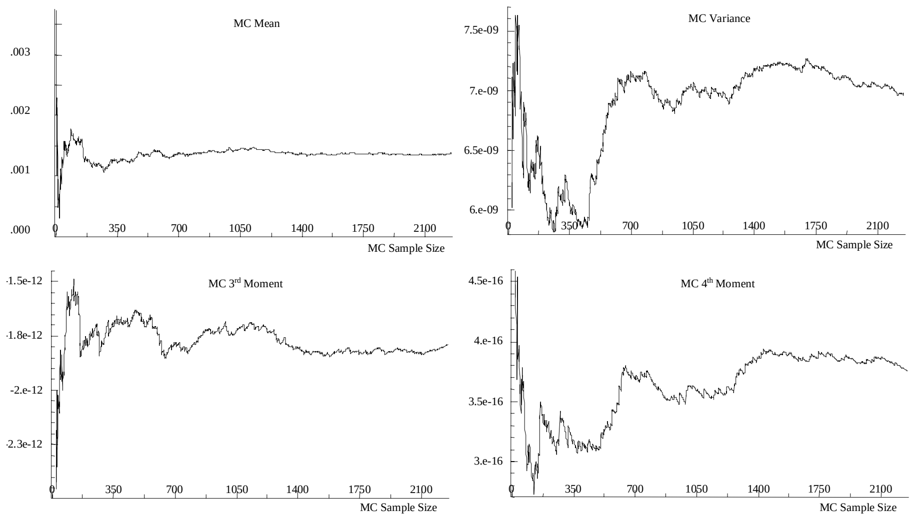



Figure A3.2: An Example of Convergence of Montecarlo Moments (of AGR) in a High Opportunity Setup. Recursive Moments computed for subsequent samples [0,M]), Parameters: $\lambda$=5, $\pi$=0.4, $\rho$=0.1, $\epsilon$=0.1, $\alpha$=1.4, $\varphi$=0.4, $N$=100, $T$=1000.

# Appendix 4

Assume that the change in log of GNP $\Delta q_t$ follows a stationary $ARMA$ process. Then $\Delta q_t$ will admit a $MA(\infty)$ representation of the form: $\Delta q_t = A(L)\nu_t$, where $A(L) = 1 + A_1 L + A_2 L^2 + \ldots$, $L$ is the lag operator and $\nu_t$ is white noise. Following Cochrane (1988) and Campbell and Mankiw (1987, 1989), we computed estimates of the following persistence measures:

$$V^k = \frac{1}{k+1}\frac{Var(q_{t+k+1} - q_t)}{Var(q_{t+1} - q_t)} = \left[1 + 2\sum_{j=1}^{k}\left(1 - \frac{j}{k+1}\right)\rho_j\right],$$

where $\rho_j$ is the $j$-th autocorrelation coefficient of $\Delta q_t$; and

$$A(1) = 1 + \sum_{j=1}^{\infty} A_j$$

Notice that if $\{q_t\}$ were trend-stationary, then $A(1) = 0$, while $V^k$ approaches zero as $k$ grows. Conversely, if $\{q_t\}$ followed a random walk, then $A_j = 0$, all $j \geq 2$, so that $A(1) = V^k = 1$, all $k$. Hence, if $\{q_t\}$ were even more persistent than a random walk, both $A(1)$ and $V^k$ would exceed unity. Note also that:

$$A(1) = \sqrt{\frac{V}{1 - R^2}}$$

where $V = lim_{k\to\infty} V^k \equiv 1 + 2(\rho_1 + \rho_2 + \ldots)$ and $R = 1 - Var(\nu_t)/Var(\Delta q_t)$.

Estimation of $V^k$ and $A(1)$ can be done non-parametrically employing sample estimates of the autocorrelation function, i.e. $r_j = \hat{\gamma}(j)/\hat{\gamma}(0)$, where

$$\hat{\gamma}(j) = \frac{1}{T}\sum_{t=j+1}^{T}(\Delta q_t - \overline{\Delta q_t})(\Delta q_{t-j} - \overline{\Delta q_t})$$

and $\overline{\Delta q_t}$ is the sample mean. We employed both first differences of log(GNP) and GNP growth rates $\{h_t\}$ as the basis for computing estimates of output growth autocorrelation functions without any significant differences. Notice that both $\{\Delta q_t\}$ and $\{h_t\}$ appear to be stationary around a mean very close to zero (i.e. $\overline{\Delta q_t} \cong \overline{h_t} \cong 0$). The results presented in Table 4 are for $\{h_t\}$.

An estimate of $V^k$ (consistent for $V$ if $k$ is large) is found simply by replacing population auto-correlations with sample counterparts (once having corrected by a downward bias), i.e.:

$$\hat{V}^k = \frac{T-k}{T}\left[1 + 2\sum_{j=1}^{k}\left(1 - \frac{j}{k+1}\right)r_j\right]$$

while $A(1)$ must be estimated non-parametrically (for large $k$) by

$$\hat{A}^k(1) = \sqrt{\frac{\hat{V}^k}{1 - r_1^2}}.$$

Notice also that since $r_1^2$ underestimates $R^2$, $\hat{A}^k(1)$ tends to underestimate $A(1)$ for large $k$. Also, the standard error of $\hat{V}^k$ is equal to:

$$S.E.(\hat{V}^k) = V^k \left[\frac{3T}{4(k+1)}\right]^{-\frac{1}{2}}.$$

Then, if one would like to test whether the data come from some $ARMA(p,q)$ process, one can plug the true value of $V^k$ (computed under the null). In our analyses, however, we employed the standard deviations of the Montecarlo distribution of estimates over $M = 1000$ replications. Both Montecarlo standard deviations and theoretical standard errors are increasing with $V^k$ (or with its estimate).

Campbell and Mankiw (1989) provide Montecarlo studies on 90% critical values of $V^k$ and $\hat{A}^k(1)$ for different data generation processes and $k = 20, 40, 60$. Even though one should be aware that it is generally hard to distinguish between different representations for $\{q_t\}$ on the basis of a single non-parametric estimate, a comparison of our estimates with the corresponding 90% percentiles of an $AR(2)$ process, leads to rejection of all stationary processes with larger root less than 0.9.

In particular, when interactions are global, path-dependency is large, the likelihood of radical innovations is high and the density of islands in the lattice is small, the values of $V^k$ for $\{h_t\}$ fit quite well the case where $\{q_t\}$ is generated by an $AR(2)$ process with the larger root in the interval $[0.9, 1.0]$ and the smaller one around 0.5.