IIASA

International Institute for Applied Systems Analysis • A-2361 Laxenburg • Austria
Tel: +43 2236 807 • Fax: +43 2236 71313 • E-mail: info@iiasa.ac.at • Web: www.iiasa.ac.at

INTERIM REPORT IR-98-012 / March 1998

Competing Technologies, International Diffusion and the Rate of Convergence to a Stable Market Structure

Andrea P. Bassanini (bassanini@utovrm.it)
Giovanni Dosi (dosi@iiasa.ac.at)

Approved by
Gordon J. MacDonald (macdon@iiasa.ac.at)
Director, *IIASA*

About the Author

Andrea Bassanini is from the Department of Economics at University ”La Sapienza”, Rome, Italy, and from the Department of Economics at Stanford University, Stanford, CA, USA; Giovanni Dosi is from the Department of Economics at University ”La Sapienza”, Rome, Italy, and from the International Institute for Applied Systems Analysis, Laxenburg, Austria.

Acknowledgements

The authors thank Klaus Desmet, Judith Gebauer, Michael Horvath, Andrea Prat, Valter Sorana, and especially Yuri Kaniovski for their comments. Financial support from International Institute for Applied Systems Analysis (IIASA), Banca Nazionale del Lavoro (BNL), Italian Research Council (CNR), and Italian Ministry of University and Scientific Research (MURST) is gratefully acknowledged. The usual disclaimers apply.

Abstract

This paper is motivated by two 'stylized facts' concerning the dynamics of diffusion of different technologies competing for the same market niche. a) A stable pattern of market sharing with no overwhelming dominant position is rarely observed in markets with network esternalities. Unbounded increasing returns to adoption are often called for an explanation of this fact. However the argument is generally based on an incorrect interpretation of the popular Brian Arthur (1989) model. As we show with a simple counterexample unbounded increasing returns are neither necessary nor sufficient to lead to technological monopolies even in a stable external environment. b) International diffusion may lead sometimes to different standards in different countries (the archetypical case is the diffusion of typewriter-computer keyboards - AZERTY vs. QWERTY) or to the diffusion of the same standard in every country (the archetypical example being VCRs - Beta vs. VHS), even without intervention of any regulatory agency. Intuitively when convergence to the same standard is not an accident of history, it is an outcome of the relative weight of international spillovers as compared to nationwide esternalities. The crucial question is: can a model that account for the former fact accomodate also the latter? In this paper, by establishing some mathematical properties of generalized urn schemes, we build on a class of competing technology dynamics models to develop an explanation for the former "fact" and to provide sufficient conditions for convergence to the same or to different technological monopolies in different countries. Our explanation for the empirical tendency to converge to technological monopoly relies on convergence rate differentials to limit market shares: We show that a market can approach a monopoly with a higher speed than it approaches any feasible limit market share where both technologies coexist. Convergence to market sharing, we conclude, is in general so slow that the environment changes before the market share trajectory becomes stable in a neighborhood of its limit. The empirical implication is that among markets with high rate of technological change and increasing returns to adoption, a prevalence of stable monopolies over stable market sharing should be expected.

JEL classification numbers: O30, O33.

Contents

Competing Technologies, International Diffusion and the Rate of Convergence to a Stable Market Structure

Andrea P. Bassanini
Giovanni Dosi

1 Introduction

In this paper we address the dynamics competing technologies. Two stylized facts are at the origin of this work:

a) A stable pattern of market sharing between competing technologies with no overwhelming dominant position has been rarely observed in markets with positive feedbacks. For example, even in the case of operating systems, which is often quoted as a case of market sharing, Apple MacIntosh has never held a market share larger than 1/5 (a partial exception being the submarket of personal computer for educational institutions). This fact has also triggered suspicion of market inefficiencies: Technological monopolies may prevail even when the survival of more than one technology may be socially optimal [Katz and Shapiro (1986), David (1992)]. Think for example to the competition between Java-based architectures and ActiveX architectures for web-based applets: Given that with any of the two paradigms the standard tasks that can be performed are different, the general impression of experts is that society would benefit from the survival of both.

b) International diffusion may sometimes lead to different standards in different countries or conversely to the diffusion of the same standard in every country. For example, while in all the English-speaking world the QWERTY keyboard represents the standard, in the French-speaking world a slightly different version (the AZERTY keyboard) is by far the more adopted one. On the contrary in the VCR market VHS is the worldwide technological leader while the original competitor, Beta, has disappeared.

In turn, from the point of view of interpretation of the processes of diffusion of new products and technologies, it is acknowledged that, in many modern markets, they are characterised by increasing returns to adoption or positive feedbacks. This has partly to do with supply-side causes: the increasing amount of knowledge and skills that are dynamically accumulated through the expansion of markets and production usually reduce the hedonic price of production or consumption goods, thus increasing the net benefit for the user of a particular technology. Boeing 727, for example, which has been on the jet aircraft market for years, has undergone constant modification of the design and improvement in structural soundness, wing design, payload capacity and engine efficiency as it accumulates airline adoption and hours of flight [Rosenberg (1982), Arthur (1989)]. Similar observations can be made for many helicopter designs [Saviotti and Trickett (1992)].

Supply-side causes of this type have received some attention in the economic literature for quite a while. However in the last fifteen years a great deal of attention has been devoted also to demand-side positive feedbacks, so-called network esternalities or (more neutrally) network effects [Katz and Shapiro (1994), Liebowitz and Margolis (1994)]. For example, telecommunication devices and networks (for instance fax machines), as a first approximation, tend not to provide any utility per se but only as a function of the number of adopters of compatible technologies with whom the communication is possible [Rohlfs (1974), Oren and Smith (1981), Economides (1996)]. The benefits accruing to a user of a particular hardware system depend on the availability of software whose quantity and variety may depend on the size of the market if there are increasing returns in software production. This is the case of VCRs, microprocessors, hi-fi devices and in general systems made of complementary products which need not be consumed in fixed proportions [Cusumano et al. (1992), Church and Gandal (1993), Katz and Shapiro (1985, 1994)]. A similar story can be told for the provision of postpurchase service for durable goods. In automobile markets, for example, the diffusion of foreign models has often been slow because of consumers' perception of a thinner and less experienced network of repair services [Katz and Shapiro (1985)]. Standardization implies also saving out of the cost of investment in complementary capital if returns from investment are not completely appropriable: In software adoption firms can draw from a large pool of experienced users if they adopt software belonging to a widespread standard, thus de facto sharing the cost of training [Farrell and Saloner (1986), Brynjolfsson and Kemerer (1996)]. Moreover product information may be more easily available for more popular brands or, finally, there may be conformity or psychological bandwagon effects [Katz and Shapiro (1985), Glaziev and Kaniovski (1991), Banerjee (1992), Arthur and Lane (1993), Bernheim (1994), Dosi and Kaniovski (1994), Brock and Durlauf (1995)].

Katz and Shapiro (1994) in their review of the literature on systems competition and dynamics of adoption under increasing returns distinguish between *technology adoption decision* and *product selection decision.*

The former refers to the choice of a potential user to place a demand in a particular market. Relevant questions in this case are the conditions for an actual market of positive size, the notional features of a "socially optimal" market size and the conditions allowing penetration of a new (more advanced) technology into the market of an already established one [Rohlfs (1974), Oren and Smith (1981), Farrell and Saloner (1985,1986), Katz and Shapiro (1992)]. For example purchasing or not a fax or substituting a compact disc player for an analogical record player are technology adoption decisions.

Conversely product selection refers to the choice between different technological solutions which perform (approximately) the same function and are therefore close substitutes. Relevant questions here are whether the market enhances variety or standardization, whether the emerging market structure is normatively desirable and what is the role of history in the selection of market structure [Arthur (1983,1989), Katz and Shapiro (1985,1986), David (1985), Church and Gandal (1993), Dosi et al. (1994)]. Choosing between VHS or Beta in the VCR market or between Word or Wordperfect in the word-processors market are typical examples of product selection decisions.

This work is concerned with the dynamics of product selection. To explain the two stylized facts recalled above we analyze properties of a fairly general and nowadays rather standard class of models of competing technologies, originally suggested by Arthur (1983) and Arthur et al. (1983) and subsequently made popular by Arthur (1989) [and further explored by Cowan (1991), Dosi et al. (1994), Dosi and Kaniovski (1994) and Kaniovski and Young (1995) among others]. This class of models will be presented in details in section 2.

Despite mixed results of some pioneering work on the dynamics of markets with network effects [e.g. Katz and Shapiro (1986)], unbounded increasing returns are commonly called for as an explanation of the emergence of technological monopolies. Usually the argument is based on the results of the model set forth by Arthur (1989). For instance Robin Cowan summarizes it in the following way:

> "If technologies operate under dynamic increasing returns (often thought of in terms of learning-by-doing or learning-by-using), then early use of one technology can create a snowballing effect by which that technology quickly becomes preferred to others and comes to dominate the market.
>
> "Following Arthur, consider a market in which two types of consumers adopt technology sequentially. As a result of dynamic increasing returns arising from learning-by-using, the payoff to adopting a technology is an increasing function of the number of times it has been adopted in the past. Important with regard to which technology is chosen next is how many times each of the technologies has been chosen in the past. Arthur shows that if the order of adopters is random (that is, the type of the next adopter is not predictable) then *with certainty one technology will claim the entire market*" [Cowan (1990) p. 543, italics added].

It will be shown in the following that *this statement does not always hold*. Unbounded increasing returns to adoption are neither necessary nor sufficient to lead to the emergence of technological monopolies. As proved in the next section, strictly speaking, Arthur's result applies only when returns are linearly increasing and the degree of heterogeneity of agents is, in a sense, small. Moreover it cannot be easily generalized further: Some meaningful counterexamples will be provided. More generally the emergence of technological monopolies depends on the nature of increasing returns with respect to the degree of heterogeneity of the population. Relatedly, given a sufficiently high heterogeneity amongst economic agents, limit market sharing may occur even in the presence of unbounded increasing returns. The bearing of our analysis, in terms of the interpretation of the empirical evidence, stems from the results presented in section three: In essence, we suggest that the observation of the widespread emergence of monopolies is intimately related to the properties of different rates of convergence (to monopoly and to market sharing respectively) more than to the properties of limit states as such. It will be shown that a market can approach a monopoly with a higher speed than it approaches any feasible limit market shares where both technologies coexist. Following a line of reasoning put forward by Winter (1986) among others, our argument proceeds by noticing that when convergence is too slow the external environment is likely to change before any sufficiently small neighborhood of the limit can be attained. The result that we obtain, based on some mathematical properties of generalized urn schemes[1], is general for this class of models. The empirical implication is that among markets with high rate of technological change and increasing return to adoption, a prevalence of stable monopolies over stable market-sharing should be observed.

How does one reconcile this result with the second "stylized fact" on the emergence of different technological monopolies in different interdependent countries? Start by noticing that a system of local monopolies is quite different from a system with prevailing worldwide market sharing (even if at high level of aggregation they may look alike). Intuitively convergence to the same standard is an outcome of the relative weight and strength of international spillovers as compared to nationwide (or regional) esternalities. For example,

[1] Throughout this paper we label the generalization of Polya urn schemes set forth by Hill et al. (1980) as generalized urn scheme. That generalization is the most popular in economics but obviously it is not the only possible one [see e.g. Walker and Muliere (1997)].

in the case of typewriter keyboards, geographical areas with the same language tend to be reflected in spillover clusters due to free “migration” of typists, similar training institutions, etc.... On the other hand, historically, gaining leadership in the European market, with the consequent bias in the related home video market, was crucial to VHS to resolve in its favor the battle for leadership in the Japanese market as well [Cusumano et al. (1992)]. However very little modeling effort has been made so far to formally explore this intuitive explanation. In the fourth section we will consider a generalization of one-market models of competing technologies in order to establish the conditions of convergence to the same or different technological monopolies in different but interrelated markets.

2 Competing Technologies Revisited: Are Unbounded Increasing Returns Sufficient for the Emergence of Technological Monopolies?

The standard story of the class of competing technology dynamics models considered by Arthur (1983,1989) is the following.

Every period a new agent enters the market and chooses the technology which is best suited to its needs, given its preferences, information structure and the available technologies. Preferences are heterogeneous and a distribution of preferences in the population is given. Information and preferences determine a vector of payoff functions (whose dimension is equal to the number of available technologies) for every type of agent. Because of positive (negative) feedbacks, these functions depend on the number of previous adoptions. When an agent enters the market it compares the values of these functions (given its preferences, the available information, and previous adoptions) and chooses the technology which yields the maximum perceived payoff. Which "type" of agent enters the market at any given time is a stochastic event whose probability depends on the distribution of types (i.e. of preferences) in the population. Because of positive (negative) feedbacks, the probability of adoption of a particular technology is an increasing (decreasing) function of the number of previous adoptions of that technology.

The resulting stochastic process of adoptions can be easily represented as a Markov process with discrete state space [see e.g. Arthur (1983)]. Usually it is claimed that any such process can be represented and analyzed as a generalized urn scheme or Polya urn [Arthur (1988,1989)] - the advantage being that the theory of Polya urns provides general tools which, when applicable, greatly simplify the analysis of asymptotic patterns.

Consider the simplest case where two technologies, say A and B, compete for a market. Think of an urn of infinite capacity with white and black (A and B) balls. Starting with n^w white balls and n^b black ones, $n^w, n^b \geq 1$, in the urn, a ball is added into the urn at time $t = 1, 2 \ldots$. The number of balls of one colour represents the number of units of the corresponding technology that have been adopted at time t. A white ball will be added with probability $f(X(t))$. Here $f(\cdot)$ is a function (sometimes called urn function) which maps $R(0,1)$ in $[0,1]$ and $R(0,1)$ stands for the set of rational numbers in $(0,1)$; $X(t)$ is the proportion of white balls in the urn at time t. The dynamics of $X(t)$ is given by the relation

$$X(t+1) = X(t) + \frac{\xi^t(X(t)) - X(t)}{t + n^w + n^b}, \tag{1}$$

Here $\xi^t(x), t \geq 1$ are random variables independent in t such that

$$\xi^t(x) = \begin{cases} 1 & \text{with probability} \quad f(x) \\ 0 & \text{with probability} \quad 1 - f(x) \end{cases} \tag{2}$$

$\xi^t(\cdot)$ is a function of market shares dependent on the feedbacks in adoption. Designating $\xi^t(x) - E(\xi^t(x)) = \xi^t(x) - f(x)$ by $\zeta^t(x)$ we have

$$X(t+1) = X(t) + \frac{[f(X(t)) - X(t)] + \zeta^t(x)}{t + n^w + n^b}, \tag{3}$$

Asymptotic patterns of this process can be studied by analyzing the properties of the function $g(X(t)) = f(X(t)) - X(t)$ [See Dosi et al. (1994) for a review of the relevant theorems in the one dimensional case].

In the most general model, positive (negative) feedbacks in adoption should imply that the probability of adoption is an increasing (decreasing) function of the number of previous adoptions and not only of current market shares. However the general case can be handled if the basic model set forth above is amended by setting a two-argument urn function, dependent also on t, given the relationships between (a) total number of adoptions of both technologies $t + n^w + n^b$, (b) current market share $X(t)$, and (c) number of adoptions of one specific technology, $n^i(t)$, $i = A, B$, that is $n^A(t) = (t + n^w + n^b)X(t)$. Note in this respect that, from an interpretative point of view, *total numbers* and *shares* are likely to capture quite distinct economic and technological phenomena. For example, sheer collective externalities (related to e.g. the case of a common network) are easily captured by the shared dynamics, while more idiosyncratic, cumulative and type-specific processes of learning are more naturally represented as functions of varying numbers of adopters (learning-by-doing and learning-by-using being classical cases to the point).

As shown by Arthur et al. (1986, theorem 3.1) [see also Dosi et al. (1994)] the relevant theorems to study asymptotic patterns of convergence still hold provided that there exist a limit urn function $f(\cdot)$ (defined as that function f such that f_t tends to it as t tends to ∞) and the following condition is satisfied:

$$\sum_{t \geq 1} t^{-1} \sup_{x \in [0,1] \cap R(0,1)} \mid f_t(x) - f(x) \mid < \infty. \tag{4}$$

Arthur (1983,1989) considers a payoff function of the following type:

$$\Pi^i_j(n^j(t)) = a^i_j + r(n^j(t)), \tag{5}$$

where $j = A, B$, $i \in S$, S is the set of possible types [in the simplest case, considered also in the foregoing quotation from Cowan (1990), $S = \{1, 2\}$], $n^j(t)$ is the number of adoptions of technology j at time t, a^j_i identifies a baseline payoff for agents of type i from technology j (every agent is associated to a pair (a^i_A, a^i_B) which represents the network-independent component of its preferences), and r is an increasing function (common for every agent) capturing increasing returns to adoption. If, at time t, an agent of type i comes to the market, it compares the two payoff functions choosing A if and only if[2]:

$$\Pi_A(n^A) \geq \Pi_B(n^B). \tag{6}$$

that is

$$a^i_A + r(n^A(t)) \geq a^i_B + r(n^B(t)). \tag{7}$$

Suppose that which type of agent enters the market at time t is the realization of an iid random variable $i(t)$. Thus (3) implies that the agent coming to the market chooses A with probability

[2]We assume that, if there is a tie, agents choose technology A. Qualitatively, breaking the tie in a different way would not make any difference.

$$\mathcal{P}(A(t)) = F_\theta(r(n^A(t) - r(n^B(t))), \tag{8}$$

where $F_\theta(\cdot)$ denotes the distribution function of $\theta(t) = a^i_B(t) - a^i_A(t)$.

From these considerations Arthur's main theorem can be derived:

Theorem 1 [Arthur (1989), theorem 3] *If the improvement function r increases at least at rate $\epsilon > 0$ as n^j increases, the adoption process converges to the dominance of a single technology, with probability one.*

The proof of the theorem is based on theorem 3.1 of Arthur et al. (1986). In fact it is easy to check that in this case, whatever the distribution of a_j is, the limit urn function $f(\cdot)$ is a step function defined in the following way:

$$f(x) = \begin{cases} 1 & \text{if} \quad x > 1/2 \\ F_\theta(0) & \text{if} \quad x = 1/2 \\ 0 & \text{if} \quad x < 1/2 \end{cases} \tag{9}$$

By applying standard properties of generalized urn schemes [see e.g. Dosi et al. (1994), theorems 1 and 3], it is trivial to show that a process with an urn function defined as above converge to $S = \{0, 1\}$ with probability 1.

However theorem 3.1 of Arthur et al. (1986) is not applicable here because condition (4) does not hold in this case. Actually the urn functions are defined by:

$$f_t(x) = F_\theta(r[x(t + n^w + n^b)) - r[(1-x)(t + n^w + n^b)]). \tag{10}$$

Moreover for $t > K > 0$, t even, they are such that $f_t(0) = 0$, $f_t(1/2) = F_\theta(0)$, $f_t(1) = 1$ and they are continuous in a left neighborhood (which depends on t) of $1/2$; therefore $\sup_{x \in [0,1] \cap R(0,1)} \mid f_t(x) - f(x) \mid = \max\{F(0), 1 - F(0)\}$ which is constant with respect to t [3].

Even though Arthur's proof is *wrong*, the theorem is *right* and an *ad hoc* proof can be constructed by showing that $n^A(t) - n^B(t)$ is a time-homogeneous Markov chain with two absorbing barriers [Bassanini (1997), proposition 2.1, provides a complete proof along these lines]. However this result strictly depends on the fact that the function $r(\cdot)$ is asymptotically linear or more than linear. Arthur's result is not generalizable to any type of unbounded increasing returns. Both in the case of increasing returns that are diminishing at the margin and in the case of heterogeneous increasing returns it is possible to find simple examples where convergence to technological monopolies is not an event with probability 1.

Let us illustrate all this by means of two straightforward counterexamples.

Example 1 Let us assume that increasing returns have the common sense property that the marginal contribution to social benefit of, say, the 100th adopters is larger than that of, say, the 100,000th and that this contribution tends asymptotically to zero; formally this means that $\frac{d}{dn^j} f(n^j) > 0$, $\frac{d^2}{dn^{j2}} f(n^j) < 0$ and $\lim_{n^j \to \infty} \frac{d}{dn^j} f(n^j) = 0$ [this class of functions has been considered by Katz and Shapiro (1985)].

Focusing on the case set forth by Robin Cowan in the foregoing quotation, let us assume that there are only two types of agents ($i = 1, 2$) and two technologies. Recall Arthur's payoff functions (7), $\Pi_j(n^j(t)) = a^i_j + r(n^j(t))$, and assume that $r(\cdot) = s\log(\cdot)$ is a function (which is common for every agent: s is a constant) that formalizes unbounded increasing

[3] To be precise Arthur (1989) quotes also Arthur et al. (1983), though there the properties are stated only as yet-to-be-proved good sense conjectures.

returns to adoption. Agents choose technology A if and only if $\Pi_A(n^A) \geq \Pi_B(n^B)$. By taking the exponential on both sides and rearranging we have:

$$\frac{X(t)}{1-X(t)} \geq e^{\frac{1}{r}(a^i_B(t)-a^i_A(t))}. \tag{11}$$

The function of the attributes of agent's type which is on the right hand side can be considered a random variable because, as discussed above, $i(t)$ is a random variable. Moreover they are iid because $i(t)$ is iid. Denoting the random variables on the right hand side with $\varsigma(t)$, from (11) we have that the adoption process can be seen as a generalized urn scheme with urn function given by:

$$f(x) = F_\varsigma(x/(1-x)), \tag{12}$$

where $F_\varsigma(\cdot)$ is the distribution function of $\varsigma(t)$. Because $i(t)$ takes just two values $(1, 2)$, also $\varsigma(t)$ takes just two values:

$$\varsigma(t) = \begin{cases} e^{\frac{1}{r}(a^1_B(t)-a^1_A(t))} & \text{with probability } \alpha \\ e^{\frac{1}{r}(a^2_B(t)-a^2_A(t))} & \text{with probability } 1-\alpha \end{cases} \tag{13}$$

where we have assumed without loss of generality that

$$a^1_B(t) - a^1_A(t) \leq a^2_B(t) - a^2_A(t). \tag{14}$$

Thus F_ς is by construction a step function with two steps:

$$F_\varsigma(y) = \begin{cases} 0 & \text{if } y \leq e^{\frac{1}{r}(a^1_B(t)-a^1_A(t))} \\ \alpha & \text{if } e^{\frac{1}{r}(a^1_B(t)-a^1_A(t))} < y \leq e^{\frac{1}{r}(a^2_B(t)-a^2_A(t))} \\ 1 & \text{if } y > e^{\frac{1}{r}(a^2_B(t)-a^2_A(t))} \end{cases} \tag{15}$$

Therefore, taking into account (12), we have that the urn function has two steps and is defined in the following way:

$$f(x) = \begin{cases} 0 & \text{if } x \leq \frac{e^{\frac{1}{r}(a^1_B(t)-a^1_A(t))}}{1+e^{\frac{1}{r}(a^1_B(t)-a^1_A(t))}} \\ \alpha & \text{if } \frac{e^{\frac{1}{r}(a^1_B(t)-a^1_A(t))}}{1+e^{\frac{1}{r}(a^1_B(t)-a^1_A(t))}} < x \leq \frac{e^{\frac{1}{r}(a^2_B(t)-a^2_A(t))}}{1+e^{\frac{1}{r}(a^2_B(t)-a^2_A(t))}} \\ 1 & \text{if } x > \frac{e^{\frac{1}{r}(a^2_B(t)-a^2_A(t))}}{1+e^{\frac{1}{r}(a^2_B(t)-a^2_A(t))}} \end{cases} \tag{16}$$

Denoting with x^* and x^{**} the two points where the urn functions jumps, we can apply the results from from Dosi et al. (1994) and show that, if the following condition is satisfied

$$x^* < \alpha < x^{**} \tag{17}$$

then there is a set of initial conditions (that imply giving both technologies a chance to be chosen "at the beginning of history") for which market sharing is asymptotically attainable with positive probability.

Technically speaking [in the jargon introduced by Hill et al. (1980)] under the above condition we have five fixed points, three of which are downcrossing. The above condition imply that the ratios $(e^{a^i_A})^{1/r}/(e^{a^i_B})^{1/r}$ are sufficiently different between the two types. In other words *there might be sufficient heterogeneity among agents to counterbalance the effect of increasing returns to adoption.*

Example 2 Consider now payoff functions of this type:

$$\Pi_j(n^j) = a_j + r_j n^j, \tag{18}$$

where r_j is a bounded random variable which admits density. Such a function allows agents to be heterogeneous also in terms of the degree of increasing returns which they experience. First, let us take the case of $a_j = 0$, $j = A, B$. By deviding payoff functions by total number of adoptions we obtain as before that agents choose A if and only if:

$$\frac{X(t)}{1 - X(t)} \geq \frac{r_B}{r_A}. \tag{19}$$

Now suppose that r_A and r_B are highly correlated and both have bimodal distributions very concentrated around the two modes, so that the distribution of r_A/r_B is also bimodal and very concentrated around the two modes. If the two modes are sufficiently far away, by repeating the same argument of the example above, we obtain that the urn function, though continuous, has a graph not too different from the graph of a step function with at least three downcrossing fixed point; thus there exists a market share $x^0 \notin \{0, 1\}$ which can be asymptotically attained with positive probability. Moreover if $a_j \neq 0$ but have bounded support and admit density, then condition (4) applies and the same argument holds.

The two examples above show that the degree of increasing returns needs to be compared to the degree of heterogeneity. Unbounded increasing returns that are diminishing fast at the margin are not sufficient to generate asymptotic survival of only one technology, provided that agents are not completely homogeneous [see also Farrell and Saloner (1985,1986) for early models with homogeneous agents that leads to the survival of only one technology]. Even more interesting, when heterogeneity is so wide that there are agent-specific increasing returns, the emergence of technological monopolies is not guaranteed even with returns that are linearly increasing.

To summarize, the foregoing examples show that if preferences are sufficiently heterogeneous and/or increasing returns to adoption are less than asymptotically linear, then Arthur's result cannot be generalized and variety in the asymptotic distribution of technologies can be an outcome with positive probability.

From the point of view of empirical predictions, at first look, the foregoing results might sound, if anything, as a further pessimistic note on "indeterminacy". That is, not only "history matters" in the sense that initial small events might determine which of the notional, technologically attainable, asymptotic states the system might "choose": More troubling, the argument so far suggests that, further, the very distribution of the fine characteristics and preferences of the population of agents might determine the very nature of the attainable asymptotic states themselves. Short of empirically convincing restrictions on the distribution of agents (normally unobservable) characteristics, what we propose is instead an interpretation of the general occurrence of technological monopolies (*cum* increasing returns of some kind) grounded on the relative speed of convergence to the underlying (but unobservable) limit states.

3 Rate of Convergence in One-Dimensional Models of Competing Technologies

Consider the celebrated example of the VCR market. JVC's VHS and Sony's Beta were commercialized approximately at the same time. According to many studies [see

Cusumano et al. (1992) and Liebowitz and Margolis (1994)] none of the two standards has ever been perceived as unambiguously better and, despite their incompatibility, their features were more or less the same, due to the common derivation from the U-matic design. For these reasons the relevant decisions were likely to be sequential. First, a consumer chooses whether or not to adopt a VCR - technology adoption decision in Katz and Shapiro's terminology -, then, once the adoption decision has been made, it devotes its mind to choose which type of VCR to purchase - product selection decision - (in general it can be expected that most of the consumers buy one single item and not both). Network effects in this market come mainly from increasing returns in design specialization and production of VCR models (so that historically all firms specialized just in one single standard) on the supply side, and from increasing returns externalities and consequent availability of home video rental services on the demand side [Cusumano et al. (1992)]. Despite technical similarities between the two standards, preferences were strongly heterogeneous, due mainly to a brand-name-loyalty type of consumer behavior, which was exploited (especially by JVC) through Original Equipment Manufacturers (OEM) agreements with firms with well-established market shares in electronic durable goods.

The fact that the decision is sequential makes product selection decisions naturally dependent on market shares rather than on the absolute size of the network. Moreover the size of VCR market is sufficiently large (hundreds of millions of sold units) to make it approximable by the abstract concept of an infinite capacity market. Therefore, following the presentation given in the previous section, it is clear that the dynamics of this market can be represented by a generalized urn scheme with urn function independent of t. Heterogeneity of preferences and the degrees of increasing returns affect only the functional form of the urn function but do not jeopardize the possibility of representing the dynamics through this type of models.

Many other markets display somewhat similar characteristics (for instance spreadsheets, wordprocessors, computer keyboards, pc-hardwares, automobiles etc...); particularly in many markets product selection can be assumed to sequentially follow technology adoption decisions [for instance in the data set of Computer Intelligence InfoCorp employed by Breuhan (1996), more than 80% of the firms in the sample report using a single word processing package]. Provided that product selection is likely to be dependent on market shares rather than absolute size of installed base, the analysis that follows could go through even in the absence of a clear pattern of increasing returns to adoption. Simply enough, in the presence of constant returns to adoption, the urn function would be completely constant but the following theorems would still hold.

Theorems 2 and 3 represent the main result of this paper: their comparison allow us to make the relevant statements on the rate of convergence to technological monopoly or to a limit market share where both technologies coexist.

As above, denote with $f(\cdot)$ the urn function; the following theorem gives the results on the rate of convergence to 0 and 1.

Theorem 2 *Let* $\epsilon > 0$ *and* $c < 1$ *be such that*

$$f(x) \leq cx \;\; for \;\; x \in (0, \epsilon) \quad (f(x)) \geq 1 - c(1-x) \;\; for \;\; x \in (1-\epsilon, 1)).$$

Then for any $\delta \in (0, 1-c)$ *and* $\tau > 0$

$$\lim_{t\to\infty} \mathcal{P}\{t^{1-c-\delta} X(t) < \tau | X(t) \to 0\} = 1$$
$$\left(\lim_{t\to\infty} \mathcal{P}\{t^{1-c-\delta}[1 - X(t)] < \tau | X(t) \to 1\} = 1\right),$$

where $X(\cdot)$ *stands for the random process given by* (3).

The theorem is proved in the appendix.

As shown by Hill et al. (1980) and Arthur et al. (1987) [see also the review in Dosi et al. (1994)] the market share trajectory $X(\cdot)$ converges almost surely, as t tends to infinity, to the set of appropriately defined zeroes of the function $g(x) = f(x) - x$. However since we are not going to restrict ourselves to the case when $g(\cdot)$ is a continuous function, we need some standard definition concerning equations with discontinuous functions.

For a function $g(\cdot)$ given on $R(0,1)$ and a point $x \in [0,1]$ set

$$\begin{aligned} \underline{a}(x,g) &= \inf_{\{y_k\}\in R(0,1)} \liminf_{k\to\infty} g(y_k), \\ \bar{a}(x,g) &= \sup_{\{y_k\}\in R(0,1)} \limsup_{k\to\infty} g(y_k), \end{aligned} \tag{20}$$

where $\{y_k\}$ is an arbitrary sequence converging to x. Then the set of zeroes $A(g)$ of $g(\cdot)$ on $[0,1]$ is defined by the following relation $A(g) = \{x \in [0,1] : [\underline{a}(x,g), \bar{a}(x,g)] \ni 0\}$. Note that for a continuous $g(\cdot)$ this definition gives the roots of the equation $g(x) = 0$ in the conventional meaning.

Applying the results known for the generalized urn scheme, we can provide conditions of convergence with positive probability to subintervals and certain singleton components of $A(g)$ and also with zero probability to other singleton components of this set. The components of the first kind are called *attainable*, while those of the second kind *unattainable*. For our model of competing technologies they would mean a separation of all possible limit market shares (given by the roots of equation $g(x) = x$) into *feasible* (i.e. which realize with some positive probability) and *infeasible* (i.e. which never realize).

One particular class of attainable singleton components comprises the *stable* ones, i.e. the points where $f(x)-x$ changes its sign from plus to minus. We are going to characterize the (relative) rate of convergence to them.

From the theory of urn schemes with balls of two colors [see Hill et al. (1980)] we know that if θ is a stable root [or a downcrossing point as it is called in Hill et al. (1980)] of $f(x) - x = 0$, that is there is $\epsilon > 0$ such that for every $\delta \in (0, \epsilon)$

$$\inf_{\delta\le|x-\theta|\le\epsilon} [f(x) - x](x - \theta) < 0, \tag{21}$$

then $X(\cdot)$ converges to θ with positive probability for some initial numbers of balls in the urn.

But if in addition to (21)

$$f(x) \in (0,1) \qquad \textit{for all} \qquad x \in R(0,1), \tag{22}$$

then it converges with positive probability to θ for any initial combination of balls in the urn.

Finally if the urn function does not have touchpoints and the set $A(g)$ with $g(x) = f(x)-x$ is composed only of singleton components then almost surely the process converges to the set of stable components.

One would like to derive a counterpart of theorem 2 for the case of stable roots. But (21) is not enough to do this. Let us consider the hypotheses of that theorem in details. These have the form

$$f(x) \le cx \qquad \textit{for} \;\; x \in (0,\epsilon) \;\; (f(x) \ge 1 - c(1-x) \qquad \textit{for} \;\; x \in (1-\epsilon, 1)). \tag{23}$$

This is implies a sort of sensitivity conditions concerning the deviations of $f(x)$ from $f(0) = 0$ ($f(1) = 1$) in a neighborhood of 0(1). Hence, to ensure compatibility of the results we are going to obtain, with theorem 2 we need the following hypothesis

$$[f(x) - \theta](x - \theta) \le k(x - \theta)^2 \quad \text{for } x \in (\theta - \epsilon, \theta + \epsilon). \tag{24}$$

Here $k < 1$ is a constant and ϵ is a small positive number. Indeed this, as well as (23), implies the continuity of $f(\cdot)$ at θ and a linear bound in $|x-\theta|$ bound for the deviation of $f(x)$ from $f(\theta)=\theta$.

Now we are ready to formulate the theorem concerning convergence rates to stable roots.

Theorem 3 *Let $\theta \in (0,1)$ be a stable root and (24) take place for some $\epsilon > 0$ and $k < 1$. Then for every $\delta \in (0, \min(1-k, 1/2))$ and $\tau > 0$*

$$\lim_{t\to\infty} \mathcal{P}\{t^{\min(1-k,1/2)-\delta}|X(t)-\theta| < \tau | X(t) \to \theta\} = 1.$$

This theorem can be proved by essentially the same methods as theorem 2, and we do it in the appendix.

Theorems 2 and 3 show that convergence to 0 and 1 can be much faster (almost of order $1/t$ as $t \to \infty$) than to an interior limit (which can be almost of order $1/\sqrt{t}$ only). Here t stands for the number of adoptions to the urn. That is, we are talking about relative rates (the ideal time which is considered here is the time of product selection choices). This result is however stronger than it may seem at first glance. In fact it has also implications for the patterns of product selection in "real" (empirical) time where plausibly the speed of the market share trajectory depends also on technology adoption decisions. There is much qualitative evidence and some econometric results [e.g. Koski and Nijkamp (1997)] showing that technology adoption is at the very least independent of market shares if not enhanced by increasing asymmetry in their distribution. Thus a fortiori we can conclude that there is a natural tendency of this class of processes to converge faster to 0 or 1 rather than to an interior limit. The explanation is that the variance of $\zeta^t(x)$, which characterizes the level of random disturbances in the process (3), is $f(x)(1-f(x))$. Under condition (23) this value vanishes at 0 and 1 but it does not vanish at θ, being equal to $\theta(1-\theta)$, under condition (24). Notice also that in example 1 $c = 0$ and in example 2 $c \approx 0$.

For a differentiable $f(\cdot)$ at 0(1), (23) holds with c arbitrarily close to $\frac{d}{dx}f(0)(\frac{d}{dx}f(1))$; similarly, if $f(\cdot)$ is differentiable at θ, then k in (24) can be arbitrarily close to $\frac{d}{dx}f(\theta)$; actually for a differentiable $f(\cdot)$ the conditional limit theorem for the generalized urn scheme (see Arthur et al. (1987)) gives an even better characterization of the lowest possible convergence rate. We put this result as a remark to theorem 3.

Remark *Let $\theta \in (0,1)$ be a stable root of $f(x) - x = 0$ and $f(\cdot)$ is differentiable at θ. Then for every $y \in (-\infty, \infty)$*

$$\lim_{t\to\infty} \mathcal{P}\left\{ \sqrt{t\frac{1-2\frac{d}{dx}f(\theta)}{2\theta(1-\theta)}}[X(t)-\theta] < y | X(t) \to \theta \right\} = \Phi(y),$$

where $\Phi(\cdot)$ stands for the Gaussian distribution function having zero mean and variance 1.

Note that, even if the foregoing results are known to mathematicians in the case of a differentiable $f(\cdot)$, the analysis of their economic implications, we suggest, brings novel insights into the nature of diffusion processes.

As shown in the previous section, the urn function can have any shape and there is no reason to believe that problems characterized by 0 and 1 as the only stable points are the only ones that we can expect. Therefore, in principle, an asymptotic outcome where both technologies survive should be observable with positive frequency in real markets. As discussed in the previous section the tendency to converge to market sharing or technological monopolies is an outcome induced by the relative impact of heterogeneity of preferences and increasing returns to adoption. What tendency is realized depends on which of the two prevails. Notice however that the prevalence of one of the two factors is not always predictable ex ante even for a nearly omniscient agent fully aware of all fundamentals of the economy: in the examples of the previous section both type of outcomes are possible, but which one is realized depends on the actual sequence of historical events that lead to it. In these type of models, in general, when multiple asymptotic equilibria are attainable, history plays a major role in the selection of the actual one[4].

If asymptotic patterns were observable, the results of the previous section would imply that we should observe both stable market sharing and technological monopolies. However for the interpretation of empirical stylized facts, the point where the process eventually would converge may be irrelevant. Indeed, the rate of change of the technological and economic environment can be sufficiently high that one can always observe diffusion dynamics well short of any meaningful neighborhood of the limit it would have attained under forever constant external conditions. So while it is true that a convergent process should generate a long-lasting stable pattern, the time required to generate it may be too long to actually observe it: the world is likely to change well before convergence is actually attained. In a sense these changes can be viewed as resetting the game to its starting point.

On the basis of the mathematical results of this section we have that convergence to technological monopolies tends to be much faster (in probabilistic terms) than to any stable market sharing where both technologies coexist, because of the intrinsic variability that market sharing carries over. Thus the empirical prediction of these results can be stated as follows: in market with increasing returns to adoption and a high rate of technological change we expect to observe a prevalence of unstable market sharing (persistent fluctuations in the market shares) and technological quasi-monopolies (which are stable patterns by definition) over stable patterns of market sharing. The reason for this being that technological monopolies can be easily attained in a reasonably short time, i.e. sufficiently before any significative change in the underlying basic technological paradigms [Dosi (1982)].

Our final point is that these predictions are not contradicted by the observation of the emergence of different monopolies in different related markets (e.g. different geographical areas). Of course it is trivially true that, with mutually independent markets, different trajectories could emerge in different markets as if they were different realizations of the same experiment, since the trajectory selection is often the outcome of the realization of historical events. However, in the next section we will show that the foregoing logical construction can be extended also when markets are interdependent. Not contrary to the intuition, it is the balance between local and global feedbacks which determines whether the system converges to the same or different monopolies in every market. However a system that converges to different local monopolies has qualitative features more similar to those of a "univariate" system converging to a monopoly, rather than to those of a system converging to market sharing.

[4]For a general discussion on this point see also Dosi (1997).

4 International Diffusion and the Emergence of Technological Monopolies

Suppose that two technologies, say A and B, compete for a complex market which consists of m interacting parts (which can be thought of as economic regions or even countries) of infinite capacity. At any time $t = 1, 2, \ldots$ a new agent enters one of the markets and has to adopt one unit of one technology. The type of agent and the region are randomly determined on the basis of the distribution of agents in the population. There can be positive (negative) feedbacks not only inside every single region but also across them. Thus the probability of adoption is a function of the vector of previous adoptions in all regions. For the i-th region we consider the following indicators: $x_i(1 - x_i)$ – the current market share of $A(B)$; γ_i – the frequency of additions to the region (i.e. this is the ratio of the current number of units of the technologies in the region to the total current number of units of the technologies in the market). We assume also that initially, at $t = 1$, there are $n_i^A(n_i^B)$ units of the technology $A(B)$ in the i-th region and $n_i^A, n_i^B \geq 1$. Finally n stands for the initial number of units of the technologies on the market, i.e. $n = n_1^A + n_1^B + n_2^A + n_2^B + \ldots + n_m^A + n_m^B$.

Similarly to the setup above we can conceive m urns of infinite capacity with black and white balls which correspond respectively to technologies A and B. Starting with n_i^w white balls and n_i^b black balls, $n_i^w, n_i^b \geq 1$, in the i-th urn a ball is added into one of the urns at time instants $t = 1, 2 \ldots$. It will be added with probability $f_i(\vec{X}(t), \vec{\gamma}(t))$ to the i-th urn. Also it will be white with probability $f_i^w(\vec{X}(t), \vec{\gamma}(t))$ and black with probability $f_i^b(\vec{X}(t), \vec{\gamma}(t))$. Here $\vec{f}(\cdot, \cdot)$ is a vector function which maps $R(\vec{0}, \vec{1}) \times R(S_m)$ in S_m and $\vec{f}(\cdot, \cdot) = \vec{f}^w(\cdot, \cdot) + \vec{f}^b(\cdot, \cdot)$. By $R(\vec{0}, \vec{1})$ we designate the Cartesian product of m copies of $R(0, 1)$. S_m is defined by the following relation

$$S_m = \{\vec{x} \in R^m : \ x_i \geq 0, \ \sum_{i=1}^{m} x_i = 1\}. \tag{25}$$

Here $R(0, 1)$ and $R(S_m)$ stand for the sets of rational numbers from $(0, 1)$ and, correspondingly, vectors with rational coordinates from S_m. By $\vec{X}(t)$ we designate the m-dimensional vector whose i-th coordinate $X_i(t)$ represents the proportion of white balls in the i-th urn at time t. Also $\vec{\gamma}(t)$ stands for the m-dimensional vector whose i-th coordinate $\gamma_i(t)$ is the frequency of balls' additions to the i-th urn by the time t [5]. The dynamics of $\vec{X}(\cdot)$ and $\vec{\gamma}(\cdot)$ can be described as a scheme of multiple urns with balls of two colors, in analogy with to the one-dimensional case.

Consider the process of adoptions of technology A at time t in market i, which is obviously a stochastic variable:

$$\xi_i^t(\vec{x}, \vec{\gamma}) = E[\xi_{i,1}^t(\vec{x}, \vec{\gamma}) | \xi_{i,1}^t(\vec{x}, \vec{\gamma}) + \xi_{i,2}^t(\vec{x}, \vec{\gamma}) = 1]. \tag{26}$$

Then

$$\xi_i^t(\vec{x}, \vec{\gamma}) = \begin{cases} 1 & \text{with probability} \quad g_i(\vec{x}, \vec{\gamma}) \\ 0 & \text{with probability} \quad 1 - g_i(\vec{x}, \vec{\gamma}) \end{cases} \tag{27}$$

where $g_i(\vec{x}, \vec{\gamma}) = f_i^w(\vec{x}, \vec{\gamma})[f_i(\vec{x}, \vec{\gamma})]^{-1}$ [6] .

[5] It is clear that among the frequencies of balls' additions to the m urns only $m - 1$ are independent. Consequently, to operate with independent variables only, we can study, for example, the dynamics of the first $m - 1$, expressing the last through the rest.

[6] Throughout this section we assume that $f_i(\vec{x}, \vec{\gamma}) > 0$ for all possible $\vec{x}$ and $\vec{\gamma}$.

Let us ignore the time instants t when $X_i(\cdot)$ does not undergo any changes. Then we obtain a new process $Y_i(\cdot)$ which has the same convergence properties as $X_i(\cdot)$ providing that balls are added into the i-th urn infinitely many times with probability one. We will implicitly assume the latter condition throughout this section[7]. For particular cases $Y_i(\cdot)$ turns out to be a conventional urn process, or anyhow can be studied by means of some associated urn process.

To implement this idea, introduce $\tau_i(k)$ – the moment of the k-th addition of a ball into the i-th urn, i.e.

$$\begin{array}{l} \tau_i(1) = \min t : \xi^t_{i,1}(\vec{X}(t), \vec{\gamma}(t)) + \xi^t_{i,2}(\vec{X}(t), \vec{\gamma}(t)) = 1, \\ \tau_i(k) = \min t > \tau_i(k-1) : \\ \xi^t_{i,1}(\vec{X}(t), \vec{\gamma}(t)) + \xi^t_{i,2}(\vec{X}(t), \vec{\gamma}(t)) = 1, k \geq 2. \end{array} \tag{28}$$

Designate $X_i(\tau_i(k))$ by $Y_i(k)$ and $\xi_i^{\tau_i(k)}(\vec{X}(\tau_i(k)), \vec{\gamma}(\tau_i(k)))$ by $\zeta_i^k(Y_i(k))$. Then for $Y_i(\cdot)$ we have the following recursion

$$Y_i(k+1) = Y_i(k) + \frac{1}{k+G_i}[\zeta_i^k(Y_i(k)) - Y_i(k)], \quad k \geq 1, \quad Y_i(1) = \frac{n_i^w}{G_i}. \tag{29}$$

Note the $Y_i(\cdot)$ indeed carries all information about changes of $X_i(\cdot)$. By definition $Y_i(k) = X_i(\tau_i(k))$, but between $\tau_i(k)$ and $\tau_i(k+1)$ the process $X_i(\cdot)$ preserves its value. Consequently, should we know that $Y_i(k)$ converges with probability one (or converges with positive probability to a certain value, or converges to a certain value with zero probability, i.e. does not converge) as $k \to \infty$, the same would be true for $X_i(t)$ as $t \to \infty$. (We will systematically use this observation below without further explicit mention).

The next theorem provides sufficient conditions for convergence of $Y_i(\cdot)$ (and consequently $X_i(\cdot)$) to 0 and 1 with positive (zero) probability.

Theorem 4 *Let $g_i(\cdot) : R(0,1) \to [0,1]$ be a function such that for all possible $\vec{x} \in R(\vec{0}, \vec{1})$ and $\vec{\gamma} \in R(S_m)$*

$$g_i(\vec{x}, \vec{\gamma}) \leq g_i(x_i). \tag{30}$$

Designate by $Z_i(\cdot)$ a conventional urn process with $g_i(\cdot)$ as the urn-function and n_i^w, n_i^b as the initial numbers of balls. Then $\mathcal{P}\{Y_i(k) \to 0\} > 0$ ($\mathcal{P}\{Y_i(k) \to 1\} = 0$) if $\mathcal{P}\{Z_i(k) \to 0\} > 0$ ($\mathcal{P}\{Z_i(k) \to 1\} = 0$). Also, when

$$g_i(\vec{x}, \vec{\gamma}) \geq g_i(x_i), \tag{31}$$

the statement reads: if $\mathcal{P}\{Z_i(k) \to 1\} > 0$ ($\mathcal{P}\{Z_i(k) \to 0\} = 0$), then $\mathcal{P}\{Y_i(k) \to 1\} > 0$ ($\mathcal{P}\{Y_i(k) \to 0\} = 0$).

The theorem is proved in the appendix.

The next statement gives slightly more sophisticated conditions of convergence and nonconvergence to 0 and 1 of the process (29) and, consequently, $X_i(\cdot)$.

Theorem 5 *Set $U_i(\vec{x}) = \sup_{\vec{\gamma}} g_i(\vec{x}, \vec{\gamma})$, $L_i(\vec{x}) = \inf_{\vec{\gamma}} g_i(\vec{x}, \vec{\gamma})$. If there is $\epsilon > 0$ such that $L_i(\vec{x}) \geq x_i$ ($U_i(\vec{x}) \leq x_i$) for $x_i \in (0, \epsilon)$ ($x_i \in (1-\epsilon, 1)$), then $\mathcal{P}\{X_i(t) \to 0\} = 0$ ($\mathcal{P}\{X_i(t) \to 1\} = 0$) for any initial combination of balls in the urn. Let $g_i(\vec{x}, \vec{\gamma}) < 1$ ($g_i(\vec{x}, \vec{\gamma}) > 0$) for all $\vec{x} \in R(\vec{0}, \vec{1})$ and $\vec{\gamma} \in R(S_m)$. Also let $\epsilon > 0$, $c < 1$ and a function $f_i(\cdot)$ be such that $U_i(\vec{x}) \leq f_i(x_i) \leq c x_i$ ($L_i(\vec{x}) \geq f_i(x_i) \geq 1 - c(1 - x_i)$) for $x_i \in (0, \epsilon)$*

[7] This assumption, which in the spirit of the population models does not appear to be too strong, is needed in order to obtain asymptotic results.

$(x_i \in (1-\epsilon, 1))$. *Then* $\mathcal{P}\{X_i(t) \to 0\} > 0$ $(\mathcal{P}\{X_i(t) \to 1\} > 0)$ *for any initial numbers of balls in the* i*-th urn.*

The theorem is proved in the appendix.

Theorem 5 gives sufficient conditions of convergence with positive probability to 0 and 1. The assumptions of the theorem, anyway, are compatible with the case of independent urns (independent markets) as a particular case. More generally they represent conditions of weak feedbacks across markets (across regions) as compared to the extent of intramarkets (intraregion) spillovers: Local feedbacks are so important that a single market may converge to one technology, say A, even though all the related markets converge to the other one. In a sense, for the function $g_i(\vec{x}, \vec{\gamma})$, the most important argument is x_i.

Conversely let us now consider the case of strong positive cross-regional feedbacks: Strong positive feedbacks can be characterized in terms of the urn function as follows: We can say that spillovers are strong if $\exists \vec{\delta}, \vec{\chi}, \vec{\delta} > \vec{0}, \vec{\chi} < 1$, such that $\vec{\delta} < g(\vec{x}, \vec{\gamma}) < \vec{\chi}$ when $x_j = 0$ and $x_k = 1$ for some $k \neq i$. In other words strong positive spatial feedbacks are such that adoptions of one type in other regions can be a partial substitute for adoptions of that type in the same region: even if a region has always chosen technology $A(B)$ in the past, technology $B(A)$ can still be chosen in the future, if it has been frequently chosen in at least another region.

Let us define a set $A \subset [0,1]^m$ as *reachable* if there exists t such that $\mathcal{P}\{\vec{X}(t) \in A \big| \vec{X}(0)\} > 0$. We can derive the following result:

Theorem 6 *If* $g(\vec{0}, .) = \vec{0}$ *and* $\forall i = 1, \ldots, m$ $\exists \eta_i > 0$ *and* $c_i \in [0,1)$, *such that* $0 < g_i(\vec{x}, .) \le c_i x_i$, $\forall \vec{x} \in \prod_{i=1}^m [0, \eta_i]$ *and* $\prod_{i=1}^m [0, \eta_i]$ *is reachable, then* $\mathcal{P}\{X_i(t) \to 0\} > 0$. *If* $g(\vec{1}, .) = \vec{1}$ *and* $\forall i = 1, \ldots, m$ $\exists \eta_i > 0$ *and* $c_i \in (0,1)$, *such that* $1 > g_i(\vec{x}, .) \ge 1 - c_i(1 - x_i)$, $\forall \vec{x} \in \prod_{i=1}^m [1-\eta_i, 1]$ *and* $\prod_{i=1}^m [1-\eta_i, 1]$ *is reachable, then* $\mathcal{P}\{X_i(t) \to 1\} > 0$. *Also if* $g_i(\vec{x}, .) \ge \alpha > 0$ *in a neighborhood of every* $\vec{y}$ *such that* $y_i = 0, y_k = 1$ *for some* $k \neq i$, *then* $\mathcal{P}\{X_k(t) \to 1, X_i(t) \to 0\} = 0$ *and if* $g_i(\vec{x}, .) \le \beta < 1$ *in a neighborhood of every* $\vec{y}$ *such that* $y_i = 1, y_k = 0$ *for some* $k \neq i$, *then* $\mathcal{P}\{X_i(t) \to 1, X_k(t) \to 0\} = 0$.

The theorem is proved in the appendix.

If $\vec{0} < g(\vec{x}, \vec{\gamma}) < \vec{1}$ when $\vec{x} \in [R(0,1)]^m$ then any neighborhood of $\vec{x} = \vec{0}$ and $\vec{x} = \vec{1}$ are reachable and therefore complete worldwide monopoly of technology A or B may emerge with positive probability. Moreover theorem 6 tells us that, if feedbacks are strong, asymptotically either one technology emerges everywhere, or the other does, or market sharing is the only other possible outcome.

Subject to these conditions of weak or strong spillovers we can provide a generalization of theorem 2 to study the rate of convergence to technological monopolies.

Theorem 7 *Let* $\epsilon, \eta_i > 0$ *and* $c, c_i < 1$ *be such that either the assumptions of theorem 5 or those of theorem 6 hold. Denote* $\max_i\{c_i\}$ *with* c. *Then for any* $\delta \in (0, 1-c)$ *and* $\tau > 0$

$$\lim_{n\to\infty} \mathcal{P}\{n^{1-c-\delta} Y_i(n) < \tau | Y_i(t) \to 0\} = 1$$
$$\Big(\lim_{n\to\infty} \mathcal{P}\{n^{1-c-\delta}[1 - Y_i(n)] < \tau | Y_i(t) \to 1\} = 1\Big),$$

where $Y_i(\cdot)$ *stands for the random process given by* (29).

The line of the proof is identical to that of theorem 2, from which it can be obtained by substituting $Y_i(\cdot)$ for $X(\cdot)$.

Again, as in section 3, an important observation, which theorems 7 provides, is that convergence to 0 and 1 can be much faster (almost of order $1/t$ as $t \to \infty$) than to an interior limit (which can be almost of order $1/\sqrt{t}$ only). The explanation, at least when $g_i(\vec{x}, \vec{\gamma}) = g_i(x_i)$, is that the variance of $\zeta_i^k(x)$, which characterizes the level of random disturbances in the process (29), is $g_i(x)[1 - g_i(x)]$. This value vanishes at 0 and 1 but it does not vanish at θ, being equal to $\theta(1-\theta)$. So all the remarks that concluded section 3 can be repeated.

5 Conclusions

This paper has reassessed the empirical evidence on prevalence of technological monopolies over market-sharing in the dynamics of competing technologies. First, we have argued that the dominant explanation in the literature, namely that unbounded increasing returns can be identified as the factor responsible for this pattern, does not always hold. Arthur's results - we have shown - hold only when increasing returns to adoption are linear or more than linear and the degree of heterogeneity of agents is small. The presented counterexamples suggest that asymptotic patterns of the dynamics of competing technologies depend on the relative impact of (unbounded) increasing returns and the degree of heterogeneity of the population of adopters. Second, given all this, we have proposed that in a market with high technological dynamism, no interesting predictions can be made by simply looking at theoretical asymptotic patterns: If convergence is too slow the environment changes before the limit can be actually approached. Conversely, developing upon some mathematical properties of Polya urns, we have shown that convergence to technological monopolies tends to be (in probabilistic terms) much faster than to a limit where both technologies coexist, the empirical implication being that in markets with high turnover of basic technologies, a prevalence of technological monopolies over stable market sharing is likely to be observed. Third, we have shown that the fact that a system of different local monopolies often appears does not contradict the previous argument: Even though at high level of aggregation a system of different monopolies looks like a stable market sharing, we have shown that it has the same rate-of-convergence properties of a single technological monopoly.

6 Appendix

Proof of theorem 2. Consider only the first case – convergence to 0. Without loss of generality we can assume that $\mathcal{P}\{X(t) \to 0\} > 0$. Indeed, on the one hand, the theorem, being a statement about the convergence rate to 0, does not make any sense if $X(\cdot)$ does not converge to 0, on the other hand convergence with positive probability will be guaranteed if we additionally require that $f(x) < 1$ for all possible $x \in R(0,1)$ [see Dosi et al. (1994), theorem 4].

Let $Z(\cdot)$ be a conventional urn process with cx as the urn-function and the same initial numbers of balls. Then

$$EZ(t+1) \leq [1 - \frac{1-c}{t+n}]EZ(t), \;\; t \geq 1,$$

and consequently

$$EZ(t) \leq Z(1) \prod_{j=1}^{t-1} (1 - \frac{1-c}{j+n}) = Z(1) t^{c-1} [1 + o_t(1)],$$

where $o_t(1) \to 0$ as $t \to \infty$. Hence from Chebychev's inequality

$$\mathcal{P}\{t^{1-c-\delta} Z(t) < \tau\} \to 1 \ \ as \ \ t \to \infty \tag{32}$$

for every $\delta \in (0, 1-c)$ and $\tau > 0$.

For arbitrary $\sigma \in (0, \epsilon)$ and $v > 0$ there is N depending on these variables such that

$$\mathcal{P}\{\{X(t) \to 0\} \Delta \{X(t) \leq \sigma, \ t \geq N\}\} < v,$$

where $A \Delta B = (A \backslash B) \cup (B \backslash A)$. Also since $Z(t) \to 0$ with probability 1 as $t \to \infty$, we can choose this N so large that

$$\mathcal{P}\{\{X(t) \to 0\} \Delta \{X(t) \leq \sigma, \ Z(t) \leq \sigma, \ t \geq N\}\} < v. \tag{33}$$

To prove the theorem it is enough to show that

$$\lim_{t \to \infty} \mathcal{P}\{t^{1-c-\delta} X(t) < \tau, \ X(t) \to 0\} = \mathcal{P}\{X(t) \to 0\},$$

or, taking into account that v in (33) can be arbitrary small, that

$$\begin{aligned} &\lim\nolimits_{t \to \infty} \mathcal{P}\{t^{1-c-\delta} X(t) < \tau, \ X(t) \leq \sigma, \ Z(t) \leq \sigma, \ t \geq N\} = \\ &\mathcal{P}\{X(n) \leq \sigma, \ Z(t) \leq \sigma, \ t \geq N\}. \end{aligned} \tag{34}$$

Due to lemma 2.2 of Hill et al. (1980), $Z(\cdot)$ majorizes $X(\cdot)$ on the event $Z(t) \leq \sigma, \ t \geq N$, providing that these processes start from the same point. Hence

$$\mathcal{P}\{X(t) \leq \sigma, \ Z(t) \leq \sigma, \ t \geq N\} \geq$$

$$\limsup_{t \to \infty} \mathcal{P}\{t^{1-c-\delta} X(t) < \tau, \ X(t) \leq \sigma, \ Z(t) \leq \sigma, \ t \geq N\} \geq$$

$$\liminf_{t \to \infty} \mathcal{P}\{t^{1-c-\delta} X(t) < \tau, \ X(t) \leq \sigma, \ Z(t) \leq \sigma, \ t \geq N\} =$$

$$\liminf_{t \to \infty} E\mathcal{P}\{t^{1-c-\delta} X(t) < \tau, \ X(t) \leq \sigma, \ Z(t) \leq \sigma, \ t \geq N | X(N)\} \geq$$

$$\begin{aligned} \liminf_{t \to \infty} E\mathcal{P}\{\{t^{1-c-\delta} Z(t) \ &< \ \tau, \ X(t) \leq \sigma, \\ Z(t) \ &\leq \ \sigma, \ t \geq N | Z(N) = X(N)\} | X(N)\} = \end{aligned}$$

$$E\mathcal{P}\{X(t) \leq \sigma, \ Z(t) \leq \sigma, \ t \geq N | X(N)\} =$$

$$\mathcal{P}\{X(t) \leq \sigma, \ Z(t) \leq \sigma, \ t \geq N\},$$

i.e. (34) holds true. q.e.d.

Proof of theorem 3. The proof is based on the following lemmas[8]:

Lemma 8 *Let $f(.)$ be an urn function such that*

$$[f(x) - x](x - \theta) \geq k(x - \theta)^2$$

[8]We are much indebted to Yuri Kaniovski for suggesting us the following proof; obviously all caveats apply.

for some $k < 1$and $\theta \in (0,1)$. Then $\lim_{t\to\infty} d_t = \lim_{t\to\infty} E(X(t)-\theta)^2$, where

$$d_t = \begin{cases} (n+t)^{-1} & if \quad 2(1-k)-1>0 \\ (n+t)^{-1}\log(n+t) & if \quad 2(1-k)-1=0 \\ const(n+t)^{-2(1-k)} & if \quad 2(1-k)-1<0 \end{cases}$$

Proof Consider the process (3) and write $n = n^w + n^b$, then:

$$\begin{array}{l} E[(X(t+1)-\theta)^2\,|X(t)] = (X(t)-\theta)^2 + \\ +\frac{2}{N+t}[f(X(t)-X(t)] + \frac{1}{N+t}f(X(t))[1-f(X(t))]. \end{array}$$

Setting $\Delta_t = E(X(t)-\theta)^2$, from the assumptions ofthe lemma, taking into account that $f(X(t)[1-f(X(t))] \leq 1$, we have

$$\Delta_{t+1} \leq \Delta_t \left[1 - \frac{2(1-k)}{n+t}\right] + \frac{1}{(n+t)^2}.$$

Thus

$$\Delta_{t+1} \leq \Delta_1 \prod_{i=1}^{t}\left[1-\frac{2(1-k)}{n+i}\right] + \sum_{i=1}^{t}\frac{1}{(n+i)^2}\prod_{j=i+1}^{t}\left[1-\frac{2(1-k)}{n+j}\right].$$

Since

$$\prod_{j=i+1}^{t}\left[1-\frac{2(1-k)}{n+j}\right] = e^{-2(1-k)\sum\limits_{j=i+1}^{t}\frac{1}{(n+j)^2}}(1+\alpha_{it}) =$$

$$= e^{-2(1-k)[\log(n+t)-\log(n+i)}(1+\beta_{it}) = \left(\frac{n+i}{n+t}\right)^{2(1-k)}(1+\gamma_{it}),$$

where α_{it}, β_{it} and γ_{it} are small terms, then

$$\begin{array}{rcl} \Delta_{t+1} & \leq & \Delta_1 \left(\frac{n+i}{n+t}\right)^{2(1-k)}(1+\gamma_{it}) + \\ & & +(n+t)^{-2(1-k)}\sum_{i=1}^{t}(n+i)^{-2+2(1-k)}(1+\gamma_{it}). \end{array}$$

Since

$$\sum_{i=1}^{t}(n+i)^{-2+2(1-k)} \approx \begin{cases} \frac{1}{2(1-k)-1}(n+t)^{2(1-k)-1} & if \quad 2(1-k)-1>0 \\ \log(n+t) & if \quad 2(1-k)-1=0 \\ const & if \quad 2(1-k)-1<0 \end{cases}$$

we have that

$$\Delta_{t+1} \approx \begin{cases} \frac{1}{2(1-k)-1}(n+t)^{-1} & if \quad 2(1-k)-1>0 \\ (n+t)^{-1}\log(n+t) & if \quad 2(1-k)-1=0 \\ const(n+t)^{-2(1-k)} & if \quad 2(1-k)-1<0 \end{cases},$$

which implies the statement of the lemma. q.e.d.

Lemma 9 *For every $\delta \in (0, \min\{1-k, 1/2\})$, subject to the assumption of lemma 8,*

$$t^{\min\{1-k,1/2\}-\delta}(X(t)-\theta) \to 0$$

in probability.

Proof From Chebychev's inequality and lemma 8:

$$\mathcal{P}\left\{t^{\min\{1-k,1/2\}-\delta}\left|X(t)-\theta\right| \geq \epsilon\right\} \leq \frac{t^{\min\{2-2k,1\}-2\delta}E(X(t)-\theta)^2}{\epsilon^2} \to 0$$

as $t \to \infty$. q.e.d.

Set

$$\tilde{f}(x)\begin{cases} f(\theta-\epsilon) & if \quad x<\theta-\epsilon \\ f(x) & if \quad \theta-\epsilon<x<\theta+\epsilon \\ f(\theta+\epsilon) & if \quad x>\theta+\epsilon \end{cases},$$

$$\tilde{X}(t+1)=\tilde{X}(t)+\frac{\tilde{\xi}^t(\tilde{X}(t))-\tilde{X}(t)}{t+n},$$

$$\tilde{X}(1)=X(1).$$

Then with probability 1 $\tilde{X}(t)\to\theta$ as $t\to\infty$. Also by lemma 9,

$$\mathcal{P}\left\{t^{\min\{1-k,1/2\}-\delta}\left|\tilde{X}(t)-\theta\right|<\tau\right\}\to 0 \;\; as \;\; t\to\infty. \tag{35}$$

As for theorem 2 we can ignore the case when $X(t)$ does not converge with positive probability. Set

$$Y(t)=t^{\min\{1-k,1/2\}-\delta}\left|X(t)-\theta\right|,$$
$$\tilde{Y}(t)=t^{\min\{1-k,1/2\}-\delta}\left|\tilde{X}(t)-\theta\right|.$$

We have to show that as $t\to\infty$

$$\mathcal{P}\left\{Y(t)>\tau,\; X(s)\to\theta\right\}\to 0. \tag{36}$$

For every $\sigma>0$ there is a $t(\sigma)$ such that

$$\mathcal{P}\left\{\{X(s)\to\theta\}\,\Delta\,\{|X(t)-\theta|<\epsilon, t\geq t(\sigma)\}\right\}\leq\sigma.$$

Since σ can be arbitrariliy small, (36) holds if

$$\mathcal{P}\left\{Y(t)>\tau, |X(s)-\theta|<\epsilon, s\geq t(\sigma)\right\}\to 0.$$

However

$$\begin{aligned}&\mathcal{P}\left\{Y(t)>\tau, |X(s)-\theta|<\epsilon, s\geq t(\sigma)\right\}= \\ &=\sum_{y\in S_{t(\sigma)}}\mathcal{P}\left\{Y(t)>\tau, |X(s)-\theta|<\epsilon, s\geq t(\sigma)\,|X(t(\sigma))=y\right\}\mathcal{P}\left\{X(t(\sigma))=y\right\},\end{aligned} \tag{37}$$

where $S_t=\tilde{S}_t=\left\{\frac{X(1)(n+i)}{n+t-1}, 0\leq i\leq t-1\right\}$ is the set of values that $X(t)$ and $\tilde{X}(t)$ can attain (not necessarily with positive probability). Notice that for any $y\in S_{t(\sigma)}$

$$\mathcal{P}\left\{Y(t)>\tau, |X(s)-\theta|<\epsilon, s\geq t(\sigma)\,|X(t(\sigma))=y\right\}=$$
$$=\mathcal{P}\left\{\tilde{Y}(t)>\tau, \left|\tilde{X}(s)-\theta\right|<\epsilon, s\geq t(\sigma)\left|\tilde{X}(t(\sigma))=y\right.\right\}.$$

This follows from the fact that $\tilde{f}(x)$ and $f(x)$ are the same for $x\in[\theta-\epsilon,\theta+\epsilon]$. However for every $y\in S_{t(\sigma)}$

$$\mathcal{P}\left\{\tilde{Y}(t)>\tau, \left|\tilde{X}(s)-\theta\right|<\epsilon, s\geq t(\sigma)\right\}\leq\mathcal{P}\left\{\tilde{Y}(t)>\tau\right\}\to 0$$

as $t \to \infty$ by (35). Thus (37) is a sum of a finite number - namely $t(\sigma)$ - of terms each converging to zero. This completes the proof. q.e.d.

Proof of theorem 4. The theorem is a straightforward consequence of the following lemma which generalizes lemma 2.2 from the paper of Hill et al. (1980):

Lemma 10 *Assume that we have a scheme of multiple urns, given by a set of the functions $f_i^w(\cdot,\cdot)$ and $f_i^b(\cdot,\cdot)$, $i = 1, 2, \ldots, m$. Let for some i a function $g_i(\cdot)$ be such that* (30) *or* (31) *holds true. Then there is a probability space, where the process* (29) *and a conventional urn process $Z_i(\cdot)$ can be realized and $Y_i(k) \le Z_i(k)$ or $Y_i(k) \ge Z_i(k)$ with probability 1 for $k \ge 1$ depending upon whether* (30) *or* (31) *holds. The process $Z_i(\cdot)$ has $g_i(\cdot)$ as urn-functions and $n_i^w, : n_i^b$ as initial numbers of balls.*

Proof Fix a probability space with $r_t, t \ge 1$, a sequence of independent random variables having the uniform distribution on $[0, 1]$. For given $\vec{x} \in R(\vec{0}, \vec{1})$ and $\vec{\gamma} \in R(S_m)$ introduce a partition of $[0, 1]$ by the points

$$t_0 = 0,\ t_1 = f_1^w(\vec{x}, \vec{\gamma}),\ t_2 = f_1(\vec{x}, \vec{\gamma}),$$

$$t_3 = f_1(\vec{x}, \vec{\gamma}) + f_2^w(\vec{x}, \vec{\gamma}),\ t_4 = f_1(\vec{x}, \vec{\gamma}) + f_2(\vec{x}, \vec{\gamma}),\ \ldots\ ,$$

$$t_{2m-1} = f_1(\vec{x}, \vec{\gamma}) + f_2(\vec{x}, \vec{\gamma}) + \ldots + f_{m-1}(\vec{x}, \vec{\gamma}) + f_m^w(\vec{x}, \vec{\gamma}),\ t_{2m} = 1.$$

Set

$$\xi_{i,1}^n(\vec{x}, \vec{\gamma}) = \chi_{\{r_n \in (t_{2(i-1)}, t_{2i-1})\}},\ \xi_{i,2}^n(\vec{x}, \vec{\gamma}) = \chi_{\{r_n \in (t_{2i-1}, t_{2i})\}},$$

where χ_A stands for the indicator of the event A. If $\tau_i(\cdot)$ are defined as above, set

$$\tilde{\zeta}_i^k(x_i) = \chi_{\{r_{\tau_i(k)} \in (t_{2(i-1)}, t_{2(i-1)} + g_i(x_i) f_i(\vec{x}, \vec{\gamma}))\}} \tag{38}$$

and

$$Z_i(k+1) = Z_i(k) + \frac{1}{k + G_i}[\tilde{\zeta}_i^k(Z_i(k)) - Z_i(k)],\ \ k \ge 1,\ \ Z_i(1) = \frac{n_i^w}{G_i}.$$

Hence $Y_i(\cdot)$ and $Z_i(\cdot)$ are given on the same probability space. If with $\tilde{F}_k^i$ we denote the σ-algebra generated by $r_t, : t \le \tau_i(k)$, then $E[\tilde{\zeta}_i^k(Z_i(k)) | \tilde{F}_k^i] = g_i(Z_i(k))$. Hence $Z_i(\cdot)$ is a conventional urn process with $g_i(\cdot)$ as urn-function.

Notice that

$$\zeta_i^k(x_i) = \chi_{\{r_{\tau_i(k)} \in (t_{2(i-1)}, t_{2(i-1)} + g(\vec{x}, \vec{\gamma}) f_i(\vec{x}, \vec{\gamma}))\}},$$

which, from (38), implies that $\zeta_i^t(x_i) \le \tilde{\zeta}_i^t(x_i)$ or $\zeta_i^t(x_i) \ge \tilde{\zeta}_i^t(x_i)$ with probability 1 depending whether (30) or (31) holds. Now to accomplish the proof it is enough to check that

$$Y_i(t+1) = \frac{(t + G_i - 1) Y_i(t) + \zeta_i^t(Y_i(t))}{t + G_i},$$

$$Z_i(t+1) = \frac{(t + G_i - 1) Z_i(t) + \tilde{\zeta}_i^t(Z_i(t))}{t + G_i}.$$

The lemma is proved. q.e.d.

Proof of theorem 5. We need the following lemma:

Lemma 11 *Let $X(\cdot)$ be a conventional urn process with $f(\cdot)$ as urn-function. If there is $\epsilon > 0$ such that*

$$f(x) \geq x \; for \; x \in (0, \epsilon) \quad (f(x) \leq x \; for \; x \in (1-\epsilon, 1)),$$

then $\mathcal{P}\{X(t) \to 0\} = 0$ ($\mathcal{P}\{X(t) \to 1\} = 0$)for any initial numbers of balls. Also, if $f(x) < 1$ ($f(x) > 0$) for $x \in (0,1)$ and there is $\epsilon > 0$ such that

$$f(x) < x \; for \; x \in (0, \epsilon) \quad (f(x) > x \; for \; x \in (1-\epsilon, 1)),$$

then $\mathcal{P}\{X(t) \to 0\} > 0$ ($\mathcal{P}\{X(t) \to 1\} > 0$) for any initial numbers of balls.

Proof Set that all conventional urn processes appearing here start from the same numbers of balls in the urn. Let $f(x) > x$ for $x \in (0, \epsilon)$. Set $g(x) = \max(f(x), x)$. Define $Y(\cdot)(Z(\cdot))$ a conventional urn process corresponding to the urn-function x ($g(x)$). Since $x \leq g(x)$, then due to lemma 2.2 from Hill et al. (1980) one has $Y(t) \leq Z(t)$, $t \geq 1$. Consequently $\mathcal{P}\{Z(t) \to 0\} \leq \mathcal{P}\{Y(t) \to 0\}$. But $Y(\cdot)$ is a Polya process, i.e. it converges a.s. to a random variable with a beta distribution. The limit, having a density with respect to the Lebesgue measure, takes every particular value from $[0,1]$ with probability 0. Hence $\mathcal{P}\{Y(t) \to 0\} = 0$ and, consequently, $\mathcal{P}\{Z(t) \to 0\} = 0$. But the urn-functions $f(\cdot)$ and $g(\cdot)$ agree in $(0, \epsilon)$, which due to lemma 4.1 from Hill et al. (1980) implies that $\mathcal{P}\{X(t) \to 0\} = 0$. Let $f(x) < x$ for $x \in (0, \epsilon)$ and $f(x) < 1$ for all $x \in (0,1)$. Set $g(x) = \max(f(x), x/2)$. Then $f(x) \leq g(x)$ and due to arguments similar to those given above $\mathcal{P}\{X(t) \to 0\} \geq \mathcal{P}\{Z(t) \to 0\}$, where $Z(\cdot)$ stands for a conventional urn process corresponding to $g(\cdot)$. Finally let us prove that $\mathcal{P}\{Z(t) \to 0\} > 0$. Put $d(x) = \min(g(x), g(\epsilon/2))$. The equation $d(x) - x = 0$ has the only root 0. Hence there is a conventional urn process corresponding to $d(\cdot)$ which converges to 0 with probability 1. Since $g(x) \in (0,1)$ and $d(x) \in (0,1)$ for all $x \in (0,1)$, this implies that $\mathcal{P}\{Z(t) \to 0\} > 0$ for any initial numbers of balls, because of lemma 4.1 of Hill et al. (1980).

Other cases can be handled by similar arguments. q.e.d.

Due to the aforementioned relationship between convergence of $X_i(\cdot)$ and $Y_i(\cdot)$, it is enough to establish the corresponding facts for $Y_i(\cdot)$.

The first statement follows by considering a conventional urn process with

$$d_i(y) = \inf_{\vec{x}:\, x_i = y} L_i(\vec{x}) \quad (d_i(y) = \sup_{\vec{x}:\, x_i = y} U_i(\vec{x}))$$

as the urn-function and applying lemma 11 and theorem 4.

Since convergence to 1 with positive probability can be studied by the same means, let us prove convergence with positive probability to 0 only. Let $Z_i(\cdot)$ be a conventional urn process having

$$d_i(x) = \begin{cases} f(x) & if \quad x < \epsilon/2 \\ f(\epsilon/2) & if \quad x \geq \epsilon/2 \end{cases}$$

as urn-function and starting from the same numbers of balls. Then

$$\mathcal{P}\{Z_i(t) \to 0\} = 1. \tag{39}$$

Set $l_i(t) = n_i^w(n_i^w + n_i^b + t - 1)^{-1},\ t \geq 1$. Since we assume that $g_i(\vec{x}, \vec{\gamma}) < 1$ for all possible $\vec{x}$ and $\vec{\gamma}$, the process $Y_i(\cdot)$ can move with a positive probability to the left from any point. Hence

$$\mathcal{P}\{Y_i(t) = l_i(t)\} > 0 \ \ for \ \ t \geq 1. \tag{40}$$

For any t such that $l_i(t) < \epsilon/2$ introduce $\mu_i(t)$ as the first instant after t such that $Z_i(\cdot)$ exits from $(0, \epsilon/2)$ providing that $Z_i(t) = l_i(t)$. Due to (39):

$$\mathcal{P}\{\mu_i(t) = \infty\} \to 1 \ \ as \ \ t \to \infty. \tag{41}$$

But due to lemma 10 $Y_i(n) \leq Z_i(n)$ for $t \leq n < \mu_i(t)$ providing that $Y_i(t) = Z_i(t) = l_i(t)$. Thus, taking into account (39) and (41), we get:

$$\mathcal{P}\{Y_i(n) \to 0 | Y_i(t) = l_i(t)\} \geq \mathcal{P}\{Y_i(n) \to 0,\ \mu_i(t) = \infty | Y_i(t) = l_i(t)\} \geq$$

$$\mathcal{P}\{Z_i(n) \to 0,\ \mu_i(t) = \infty | Z_i(t) = l_i(t)\} \to 1 \ \ as \ \ t \to \infty.$$

Therefore to accomplish the proof it is enough to refer to (A9). q.e.d.

Proof of theorem 6. We need the following lemma:

Lemma 12 *Consider two multiple urn processes $\vec{X}(.)$ and $\vec{Y}(.)$ which agree in a neighborhood N of a point $\theta \in (R[0,1])^m$. Then there exists an urn process $\vec{Z}(.)$ with the same urn function as $\vec{Y}(.)$ such that $\mathcal{P}\{\vec{Z}(t) \to \theta\} > 0$ if $\mathcal{P}\{\vec{X}(t) \to \theta\} > 0$. Also consider a neighborhood N_0 (N_1) of* 0 (1); *if it is reachable by $X_i(.)$, given the initial number of balls, and if the urn functions never take the values* 0 (1) *in that neighborhood* (0 (1) *excluded*), *then $\mathcal{P}\{\vec{Z}(t) \to \vec{0}\} > 0$ from every initial number of balls only if $\mathcal{P}\{\vec{X}(t) \to \vec{0}\} > 0$* ($\mathcal{P}\{\vec{Z}(t) \to \vec{1}\} > 0$ *only if* $\mathcal{P}\{\vec{X}(t) \to \vec{1}\} > 0$).

Proof From almost sure convergence we have that $\exists t > 0$ and two vectors $\vec{n}, \vec{t} \in \mathbf{Z}_+^m$ such that

$$\mathcal{P}\left\{\begin{array}{l} \vec{X}(s) \to \theta, X_i(t) = \frac{n_i}{t_i + G_i}, i = 1, ..., m, \\ \sum\limits_{i=1}^{m}(t_i + G_i) = t + G, \vec{X}(s) \in N, s > t \end{array}\right\} > 0,$$

which implies

$$\mathcal{P}\left\{\begin{array}{l} \vec{X}(s) \to \theta, \vec{X}(s) \in N, s > t \\ \Big| X_i(t) = \frac{n_i}{t_i + G_i}, i = 1, ..., m \end{array}\right\} > 0. \tag{42}$$

Take $Z_i(0) = \frac{n_i}{t_i + G_i}$. Since on N $\vec{X}(.)$ and $\vec{Y}(.)$ agree, from (42) we have

$$\mathcal{P}\left\{\vec{Z}(t) \to \theta, \vec{Z}(t) \in N, t > 0\right\} > 0,$$

from which the first statement of the lemma follows.

Since N_0 is reachable $\exists \bar{t} > 0$ and $\exists (\vec{y}, \vec{n}, \vec{t}) \in N_0 \times \mathbf{Z}_+^{2m} : \frac{n_i}{t_i + G_i} = y_i,\ i = 1, ..., m$, $\sum\limits_{i=1}^{m}(t_i + G_i) = \bar{t} + G$ such that

$$\mathcal{P}\left\{\vec{X}(\bar{t}) = \vec{y}\right\} > 0. \tag{43}$$

Take n_i, $t_i + G_i$ as initial conditions of the process $\vec{Z}(.)$. Again from almost sure convergence we have that $\mathcal{P}\{\vec{Z}(t) \to \vec{0}\} > 0$ from every initial number of balls implies that, $\exists \vec{z} \in N_0$ and $\exists T > 0$ such that

$$\mathcal{P}\left\{\vec{Z}(s) \to \vec{0}, \vec{Z}(s) \in N_0, s \geq T, \vec{Z}(T) = \vec{z} \in N_0\right\} > 0.$$

Given that the urn functions never reach 0 or 1 in N_0, $\vec{z}$ can be reached from $\vec{y}$ without leaving N_0 (through an appropriate sequence of 0 first and 1 afterwards), we have also

$$\mathcal{P}\left\{\vec{Z}(s) \to \vec{0}, \vec{Z}(s) \in N_0, \vec{Z}(0) = \vec{y} \in N_0\right\} > 0,$$

which implies

$$\mathcal{P}\left\{\vec{Z}(s) \to \vec{0}, \vec{Z}(s) \in N_0 \middle| \vec{Z}(0) = \vec{y} \in N_0\right\} > 0. \tag{44}$$

Given that on N_0 $\vec{X}(.)$ and $\vec{Y}(.)$ agree, we can choose the process $\vec{Z}(t)$ defined on the common probability space in such a way that $\vec{X}(s + \bar{t}) = \vec{Z}(s)$ for any $s < \tilde{t} = \min\left\{t : \vec{Z}(t) \notin N_0\right\}$. From (44) we have

$$\mathcal{P}\left\{\vec{X}(s) \to \vec{0}, \vec{X}(s) \in N_0 \middle| \vec{X}(\bar{t}) = \vec{y} \in N_0\right\} > 0,$$

which, taking into account (43), implies that $\exists t > 0$ such that

$$\mathcal{P}\left\{\vec{X}(s) \to \vec{0}, \vec{X}(s) \in N_0, s \geq t\right\} > 0.$$

Convergence to $\vec{1}$ can be studied with similar arguments. q.e.d.

To prove the first statement of the theorem take a process $\vec{Y}(.)$ with urn function:

$$d(\vec{x}, .) = \begin{cases} g(\vec{x}, .) & if\ \vec{x} \in \prod_{i=1}^{m}[0, \eta_i] \\ 0 & otherwise \end{cases},$$

the apply lemma 12 and lemma 10. The second statement can be proved with similar arguments. For the third statement take a process $\vec{Y}(.)$ with urn function:

$$d(\vec{x}, .) = \begin{cases} g(\vec{x}, .) & if\ \|\vec{x} - \vec{y}\| < \epsilon \\ \alpha & otherwise \end{cases},$$

then apply lemma 12 and lemma 10. The last statement can be proved with similar arguments. q.e.d.

7 References

Arthur, W.B., On Competing Technologies and Historical Small Events: The Dynamics of Choice under Increasing Returns, IIASA Working Paper WP-83-90 [Reprinted in W.B. Arthur (1994) *Increasing Returns and Path-Dependence in the Economy*, Ann Arbor: University of Michigan Press], 1983.

Arthur, W.B., Self-Reinforcing Mechanisms in Economics, in Anderson, P.W., K.F. Arrow and R. Pines (eds.), *The Economy as an Evolving Complex System*, New York: Addison-Wesley, 1988.

Arthur, W.B., Competing Technologies, Increasing Returns and Lock-In by Historical Events, *Econ. J.*, **99**, 116–131, 1989.

Arthur, W.B., Y.Ermoliev and Y.Kaniovski, Generalized Urn Problem and Its Applications, *Cybernetics*, **19**, 61-71, 1983.

Arthur, W.B., Y.Ermoliev and Y.Kaniovski, Strong Laws for a Class of Path-Dependent Urn Processes, *Proceedings of the International Conference on Stochastic Optimization, Lecture Notes on Control and Information Sciences,* **81**, 187-300, 1986.

Arthur, W.B., Y.Ermoliev and Y.Kaniovski, Path-Dependent Processes and the Emergence of Macro-Structure, *Eur. J. of Operation Research*, **30**, 294-303, 1987a.

Arthur, W.B., Y.Ermoliev and Y.Kaniovski, Nonlinear Urn Processes: Asymptotic Behavior and Applications, IIASA Working Paper WP-87-85, 1987b.

Arthur, W.B., Y.Ermoliev and Y.Kaniovski, Nonlinear Adaptive Processes of Growth with General Increments: Attainable and Unattainable Components of Terminal Set, IIASA Working Paper WP-88-86, 1988.

Arthur, W.B. and D.Lane, Information Contagion, *Structural Change and Econ. Dynamics,* **4**, 81-104, 1993.

Banerjee, A., A Simple Model of Herd Behavior, *Quart. J. Econ.,* **107**, 797-817, 1992.

Bassanini, A.P., Localized Technological Change and Path-Dependent Growth, IIASA Interim Report IR-97-086, 1997.

Bernheim, B.D., A Theory of Conformity, *J. Pol. Econ.,* **102**, 841-77, 1994.

Breuhan, A.L., Innovation and the Persistence of Technological Lock-In, unpublished manuscript, 1996.

Brock, W.A. and S.N.Durlauf, Discrete Choice with Social Interactions I: Theory, NBER Working Paper #5291, 1995.

Brynjolfsson, E. and C.F.Kemerer, Network Externalities in Microcomputer Software: An Econometric Analysis of the Spreadsheet Market, *Management Science*, **42**, 1627-47, 1996.

Church, J. and N.Gandal, Complementary Network Externalities and Technological Adoption, *Int. J. of Industrial Organization,* **11***,* 239-60, 1993.

Cowan, R., Nuclear Power Reactors: A Study in Technological Lock-In, *J. of Econ. History,* **50**, 541-67, 1990.

Cowan, R., Tortoises and Hares: Choice among Technologies of Unknown Merit, *Econ. J.,* **101**, 801-14, 1991.

Cusumano, M.A., Y.Milonadis and R.S.Rosenbloom, Strategic Maneuvering and Mass-Market Dynamics: The Triumph of VHS over Beta, *Business History Rev.,* **66**, 51-94, 1992.

David, P., Clio and the Economics of QWERTY, *AEA Papers and Proceedings*, **75**, 332-7, 1985.

David, P., Heroes, Herds and Hystheres in Technological Theory: Thomas Edison and the Battle of Systems Reconsidered, *Industrial and Corporate Change*, **1**, 129-80, 1992.

Dosi, G., Technological Paradigms and Technological Trajectories, *Research Policy,* **11**, 142-67, 1982.

Dosi, G., Opportunities, Incentives and the Collective Patterns of Technological Change, *Econ. J.,* **107**, 1530-47, 1997.

Dosi, G., Y.Ermoliev and Y.Kaniovski, Generalized Urn Schemes and Technological Dynamics, *J. Math. Econ.*, **23**, 1-19, 1994.

Dosi, G. and Y.Kaniovski, On Badly Behaved Dynamics, *J. of Evolutionary Econ.*, **4**, 93-123, 1994.

Economides, N., The Economics of Networks, *Int. J. of Industrial Organization*, **14**, 673–99, 1996.

Farrell, J. and G.Saloner, Standardization, Compatibility and Innovation, *Rand J. of Econ.,* **16**, 70-83, 1985.

Farrell, J. and G.Saloner, Installed Base and Compatibility: Innovation, Product Preannouncements, and Predation, *Amer. Econ. Rev.,* **76**, 940-55, 1986.

Glaziev, S. and Y.Kaniovski, Diffusion of Innovations under Conditions of Uncertainty: A Stochastic Approach, in N.Nakicenovic and A.Gruebler (eds.), *Diffusion of Technologies and Social Behavior,* Berlin: Springer, 1991.

Hill, B.M., D.Lane and W.Sudderth, A Strong Law for Some Generalized Urn Processes, *Annals of Probability*, **8**, 214-26, 1980.

Kaniovski, Y. and H.P.Young, Learning Dynamics in Games with Stochastic Perturbations, *Games and Econ. Behavior,* **11**, 330-63, 1995.

Katz, M.L. and C.Shapiro, Network Externalities, Competition , and Compatibility, *Amer. Econ. Rev.*, **75**, 424–40, 1985.

Katz, M.L. and C.Shapiro, Technology Adoption in the Presence of Network Externalities, *J. Pol. Econ.*, **94**, 822–41, 1986.

Katz, M.L. and C.Shapiro, Product Introduction with Network Externalities, *J. of Industrial Econ.*, **40**, 55–84, 1992.

Katz, M.L. and C.Shapiro, Systems Competition and Network Effects, *J. of Econ. Perspectives*, **8**, 93–115, 1994.

Koski, H. and P.Nijkamp, The Installed Base Effect: Some Empirical Evidence from the Microcomputer Market, unpublished manuscript, 1997.

Liebowitz, S.J. and S.E.Margolis, S.E., Network Externality: An Uncommon Tragedy, *J. of Econ. Perspectives,* **8**, 133–150, 1994.

Oren, S. and S.Smith, Critical Mass and Tariff Structure in Electronic Communications Markets, *Bell J. of Econ.*, **12**, 467–87, 1981.

Rohlfs, J., A Theory of Interdependent Demand for a Communication Service, *Bell J. of Econ.*, **5**, 16–37, 1974.

Rosenberg, N., *Inside the Black Box*, Cambridge: Cambridge University Press, 1982.

Saviotti, P.P. and A.Trickett, The Evolution of Helicopter Technology, 1940-1986, *Economics of Innovation and New Technology*, **2**, 111-30, 1992.

Walker, S. and P.Muliere, Beta-Stacy Processes and a Generalisation of the Polya-urn Scheme, *Annals of Statistics*, **25**, 1762-80, 1997.

Winter, S.G., Comments on Arrow and Lucas, *J. of Business,* **59**, S427-34, 1986.