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-97-086 / December 1997

Localized Technological Change and Path-Dependent Growth

Andrea P. Bassanini (bassanini@utovrm.it)

Approved by
Giovanni Dosi (dosi@iiasa.ac.at)
Leader, *TED project*

Interim Reports on work of the International Institute for Applied Systems Analysis receive only limited review. Views or opinions expressed herein do not necessarily represent those of the Institute, its National Member Organizations, or other organizations supporting the work.

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.

Acknowledgements

The author thanks William Brock, Paul David, Klaus Desmet, Giovanni Dosi, Jess Gaspar, Pierre Gourinchas, Peter Hammond, Michael Horvath, Chad Jones, Boyan Jovanovic, Ken Judd, Yuri Kaniovski, Mordecai Kurz, Giancarlo Marini, Yaw Nyarko, Nicola Rossi, Fabiano Schivardi, participants to the 12th annual congress of the European Economic Association and seminar participants at IIASA, Stanford and Rome "Tor Vergata" for their comments. Financial support from Banca Nazionale del Lavoro is gratefully acknowledged. The paper was completed when the author was visiting the International Institute for Applied Systems Analysis, Laxenburg, Austria. The usual disclaimers apply.

Abstract

In recent years the theory of macroeconomic growth has seen an expanding literature building upon the idea that technological change is localized (technology-specific) to investigate various phenomena such as leapfrogging, take-off, and social mobility. In this paper I explore the relationship between localized technological change and dependence on history of long-run aggregate output growth. The growth model I set forth show that, subject to mild assumptions on the stochastic processes representing exogenous environment, path-dependence of aggregate output growth is a robust property of an economic system with localized learning-by-doing and diversity of technological opportunity. In general there are multiple steady states (with fast vs. slow growth properties) whose selection depends on the sequence of historical events: *a priori* all steady states are attainable but only one of them emerges as the results of history; two economies with identical fundamentals can thus evolve towards different steady states just because of the sequence of historical (possibly small) events. Superior technologies might not be adopted: This happens because the opportunity cost of abandoning an adopted technology is higher (and the probability of switching even to a dynamically superior path is lower) the higher the accumulation of experience in the adopted technology. However growth turns out to be ergodic if agents are sufficiently heterogeneous in terms of technological capabilities. The same applies if technological opportunity is very different across technologies, but this is hardly the case when there are many "young" technologies.

JEL classification numbers: O33, O41.

Contents

Localized Technological Change and Path-Dependent Growth

Andrea P. Bassanini (bassanini@utovrm.it)

1 Introduction

Knowledge accumulation is essentially a non-binary process. Usually there is no deterministic correspondence between inputs of innovative activity and its output. Partial tacitness of knowledge may induce firm-specificity. Different technologies and different products may require different set of knowledge to be mastered and skills accumulated by producing with one technology do not have necessarily a useful employment in a different production activity; this does not mean that there are no spillovers but, to some extent, they are somewhat specific. In the literature on technological innovation there is now a broad consensus that technological change is cumulative and partially localized (firm-, technology- or product-specific) and that the results of technological effort are uncertain; moreover different technologies have different degrees of technological dynamism: technological opportunity typically differ across technologies [e.g. Nelson (1981), Levin et al. (1985), Dosi (1988), Antonelli (1995) and Stoneman (1995)].

A common thrust of the same literature is that the above attributes of technological change may induce some degree of path-dependence (or dependence on history) in the innovation and diffusion patterns at the sector level. The idea is that cumulativity and specificity bring about irreversibility and self-reinforcing mechanisms that can amplify and render persistent the effect of some historical contingency [1]. This paper is concerned with the following question: do the above characteristics of technological change have implications in terms of the degree of path-dependence of macroeconomic growth as well?

As far as I know these concepts relative to technological change have almost never been simultaneously applied in the theory of aggregate output growth. The idea that technological progress is cumulative in nature can be traced back at least to the work of Solow, Kaldor, and Arrow, in the late '50s and early '60s. Uncertainty is now widely assumed even in endogenous growth models. Both these applications do not need to be further reviewed here.

At the technique level, the first discussion of the implications of localized technological change for the theory of growth is due to Atkinson and Stiglitz (1969). David (1975) provided a first formalized stochastic model along these lines. More recently this idea has been exploited by many papers studying different phenomena related to technological change such as leapfrogging in international leadership [Brezis et al. (1993), Desmet

[1] See e.g. Arthur (1988) and Dosi (1997). David (1985), Cowan (1990), Cusumano et al. (1992), Cowan and Gunby (1996) and Kirsch (1996) provide somewhat controversial empirical evidence [for a different perspective see Liebowitz and Margolis (1995) and West (1994)] of sector-level path-dependence. Due to the insurmountable difficulty of observing unexploited opportunities, the evidence is limited to the anedoctal level of historical case studies. More interesting is the evidence on international specialization patterns: a great deal of randomness, independent of comparative advantage, is observable in the historical evolution of specialization patterns [see e.g. Dosi, Pavitt and Soete (1990)].

(1997)], dynamic complementarities and take-off [Durlauf (1993,1994)], excess human capital traps [Jovanovic and Nyarko (1996)] and social mobility [Galor and Tsiddon (1997)]. Parente (1994) and Smulders and van de Klundert (1995) consider deterministic models of growth with firm-specific technological change. Grossman and Helpman (1991) consider stochastic models where there is cumulative and sector-specific technological change; however technological opportunity is the same in every sector whose number is of the order of the continuum, therefore averages prevail and aggregate variables have patterns as if the model were deterministic.

The model of Jovanovic and Nyarko (1996) is the closest in spirit to this work: They consider a model with cumulative, uncertain and technology-specific learning-by-doing. However they exclude any type of technological change that is not dependent on experience accumulated through production. As I will discuss in the third section of this paper, to some respect their results correspond to a special case of mine.

I employ here a standard concept of dependence on history or path-dependence [2]: An economic system is path-dependent if historical events affect its long-run equilibrium patterns; if there are steady states an economy has path-dependent features if occurrence of different events may select the steady state to which the system eventually converges. This means that two economies with identical fundamentals may evolve along completely different patterns that lead to different steady states.

An economic system may be affected by big events (plagues, catastrophes, wars, major innovations, etc...), with very strong and persistent effects but very low frequence, and small events (weather conditions, incremental innovations, etc...). Most of the economic modeling takes into account only the first type of events: In deterministic models with multiple steady states, history may select the initial condition from which the attainable steady state is univoquely determined [3]. However this approach is not entirely satisfactory because on the one hand it cannot represent the dynamic process which makes history relevant (everything historically relevant is over when the analyst's videocamera is switched on, and it doesn't help much to watch the crime scene when all the facts has already happened), on the other hand it may rule out important phenomena that are related to timing, potential repetition and correlation of historical events.

The other alternative is of course a stochastic approach which allows the representation of small and big events together. It is well known that a stochastic growth model with nonconvex production sets usually generates nonergodic patterns [see Majumdar et al. (1989)]. However models of this type have often undesirable features from the point of view of modeling economic growth: shocks are usually represented by means of time-homogeneous Markov chains on compact state spaces; hence optimal investment programs are usually time-homogeneous Markov chains on compact state spaces. A nonergodic aperiodic Markov chain on a compact state space has at least two ergodic sets with empty intersection. As a consequence there are some sets of initial conditions that univoquely determine the steady state. This is not that bad in abstract, but in a growth model it may mean that a country which starts with a too low level of capital will never take-off, in sharp contrast to the fact that originally every country had low capital stock levels[4]. The only two simple ways to get around this problem, without losing the tractability of the Markov structure, are either to restrict meaningfully the set of initial states to transient sets plus, in case, acceptable ergodic sets (while increasing the dimension of the state space) or to

[2] See David (1975), Becker et al. (1990), Krugman (1991) and Durlauf (1994).

[3] For a survey see Azariadis (1996).

[4] King and Robson (1993) belongs to this class. Moreover many models can be easily modified to obtain path-dependent growth [for instance Acemoglu and Zilibotti (1996)] at the cost of incurring in the same "disease".

introduce time -heterogeneity in the transition probabilities. The second route has been followed by David (1975) and Durlauf (1994). In the models of the next two sections I will follow both approaches.

Clearly we need a formal definition of path-dependence which rules out stochastic processes where the limit distribution depends only on the initial state and do not evolve with the realization of historical events[5]. We restrict ourselves to cases where at least one limit distribution actually exists. Following Arthur et al. (1987a), Durlauf (1994), Liebowitz and Margolis (1995) and Benaim and Hirsch (1996) we can formalize it as follows. Let Ω_t be the set of states ω_t of the environment at date $t \geq 1$, with Ω_t endowed with the σ-algebra $\mathcal{E}_t$. The stochastic environment is given by the probability space $(\Omega, \mathcal{F}, \mu)$ where $\Omega = \Pi_1^\infty \Omega_t$ is the space of all sequences $\omega = \omega_t$ such that $\omega_t \in \Omega_t, \forall\, t \geq 1$, $\mathcal{F} = \otimes_1^\infty \mathcal{E}_t$ is the smallest σ-algebra on Ω generated by the measurable cylindrical sets and μ is the measure on Ω . Let partial history define the filtration $\{\mathcal{F}_t = \otimes_1^t \mathcal{E}_k\}, t \geq 1$. Finally let us denote with $\{Z_t(\omega_t)\}, t \geq 1$, a sequence of random variables. We can now state the following definition:

Definition 1 *A random process $\{Z_t(\omega_t)\}$, $t \geq 1$, defined as above is said to be path-dependent if there exist $k > 0$, $A_k \in \mathcal{F}_k$ and a distribution function F such that both the following conditions hold:*

a) $\lim_{t\to\infty} \mu(Z_t \leq x \,|A_k\,) = F(x)$ *for every x of continuity of F.*

b) $\exists h \geq 0$ *and* $\mathcal{F}_h \subset \mathcal{F}_k$ *:for every* $A_h \in \mathcal{F}_h$ *such that* $A_h \supset A_k$,

then $\lim_{t\to\infty} \mu(Z_t \leq x \,|A_h\,) \neq F(x)$ *for some x of continuity of F.*

The model I set forth in the next three sections is a stylized discrete time stochastic growth model where technological change is partially endogenous and technology-specific. The number of technologies is modeled as finite, drawing from a growing literature on technological revolutions[6]. They can be interpreted as broadly defined technological paradigms[7] or macroinventions[8], whose introduction occurs at a slower pace than convergence. Endogenous technological change is represented as learning-by-doing, which is modelled as a function of the time spent in using one technology, as implicitly assumed in the literature on endogenous growth and quality ladders [e.g. Grossman and Helpman (1991), Aghion and Howitt (1992)] and explicitly assumed in some recent paper [Parente (1994), Galor and Tsiddon (1997)].

A robust finding of the model is that the required conditions to generate path-dependent growth are relatively mild. Superior technologies may not be adopted if technological change is localized. As pointed out by Jovanovic and Nyarko (1996) this happens because the opportunity costs of switching technology increase with experience in one specific technology so that, under certain conditions, the probability of switching asymptotically tends to zero. However this tendency may be balanced by technological change independent of specific experience; therefore path-dependence may not hold if technological opportunity is very different across competing technologies. Moreover if agents are sufficiently heterogeneous in their technological capabilities, growth turns out to be ergodic (the intuition

[5]That is we want to rule out processes that do not have any transient state [a simple example of multiple-equilibria stochastic models with no transient state is provided by Brock and Durlauf (1997)].

[6]See Brezis et al. (1993), Durlauf (1993), Caselli (1996), Galor and Tsiddon (1997) and Desmet (1997).

[7]On the concept of technological paradigm see Dosi (1982,1988).

[8]See Mokyr (1990).

being that heterogeneous agents explore and develope different technologies, letting many of them survive).

The remainder of this paper is divided as follows: Next section sets forth a simple model where learning-by-doing is deterministic and stochastic perturbations are stationary. The third section introduces cumulative stochastic technological change, both endogenous and exogenous. Section four deals with heterogeneity of capabilities. Finally the last section briefly summarizes the results. A technical appendix discusses non-linear Polya processes and contains formal statement of propositions and proofs.

2 Stationary Exogenous Environment

In this section we consider a dynamic economy where time is discrete and all agents are both producers and consumers of the same homogeneous good, which can be consumed or invested. Depreciation is complete, meaning that next period capital is equal to investment this period. Every period agents can choose among two technologies (0 and 1) represented by two production functions. Production technologies are subject to idiosyncratic stochastic shocks: formally this means that, given our previously defined probability space (Ω, F, μ) , for any time t we can identify two random variables ϵ_{0t} and ϵ_{1t}, $t \geq 1$, that affect production functions 0 and 1 respectively. Current period uncertainty is resolved before choice of technology is made. Technical progress is technology-specific and occurs by historical experience: every time a technology has been chosen and production has been accomplished with that technology, the related production function shifts upwards in an exponential way. Technological opportunity may be different across technologies: the effect of learning-by-doing on production function shift may differ across technologies. Production and utility functions are assumed to be Cobb-Douglas[9].

The list of model assumptions can be formally summarized as such:

Homogeneity of agents:

H) Every agent has the same initial endowment k_1, initial production functions, preferences and is subject to the same stochastic shocks.

Exogenous stochastic environment:

S) There are technology-specific random shocks ϵ_{it}, $i = 0, 1$, serially i.i.d. and with bounded support I_1 and I_0. Moreover $\epsilon_{it} \geq \delta > 0$.

Technology:

T) At time t available production functions are described by the following equation:

$$y_{it} = e^{A_i \nu_{it}} \epsilon_{it} k_t^{\alpha}, \quad 0 < \alpha < 1, \tag{1}$$

where $t \geq 1$, $i = 0, 1$, $\nu_{it} = \sum_{k=1}^{t-1} \theta_{ik} + \gamma_i$ and θ_{ik} is the share of agents that chose technique i at time k, A_i is a technology specific coefficient which represents technological opportunity (since it determines the effect of learning-by-doing on production function shift) and γ_i are initial integer parameters (which determine the level of the initial production functions); y and k are per capita output and capital stock respectively. To simplify the notation we will write $y_{it} = Z_{it} g(k_t)$, whenever possible.

[9]Cobb-Douglas production functions are convenient to simplify the notation, because of the correspondence between Hicks-neutral, Solow-neutral, and Harrod-neutral technical change. A wider class of production and utility functions can be easily accomodated, as shown in a related paper [Bassanini (1997)].

Assumption T implies that there are full dynamic externalities, since every agent can learn from the experience of all the others. T and H imply that all the agents face the same technological choice set at any time.

Market structure:

M) There is a continuum (of measure one) of agents. Every single agent has zero measure.

Commodity space[10]*:*

C) The commodity space consists of capital stock processes $k = \{k_t\}$ and consumption processes $c = \{c_t\}$ that satisfy the following feasibility conditions: if k_1 is initial stock, the real-valued $\mathcal{F}_t$ -adapted capital stock process k and consumption process c is a feasible program if

$$\begin{array}{l} k_{t+1} + c_t \leq \max_{i \in 0,1}\{y_{it}\} \text{ a.s., } t > 0 \\ k_t, c_t \geq 0 \text{ a.s., } t > 0. \end{array} \tag{2}$$

Preferences:

U) Intertemporal preferences are represented by a time additively separable (TAS hereafter) utility function $U : \mathbf{R}_+^\infty \to \mathbf{R}$, and the discount factor is small that is:

$$U(c) = \sum_t \beta^t u(c_t), \quad 0 \leq \beta < e^{\frac{A_1 \gamma}{\alpha - 1}} \tag{3}$$

where $u : \mathbf{R}_+ \to \mathbf{R}$ is constant-relative-risk-aversion (CRRA) utility function:

$$u(c) = \frac{c^\gamma}{\gamma}, \quad 0 < \gamma < 1 \tag{4}$$

Non-triviality condition:

NT) Initial parameters are such that both techniques can be chosen with positive probability at time $t = 1$[11].

Agent's problem consists in maximizing U subject to the feasibility costraint described by assumption C. There are two choice variables: technology and investment. However, in equilibrium, because of assumption M, others' behavior and ν_{it} are taken as given, therefore in every period, since more commodities are preferred to fewer commodities, the choice of technology is not affected by choices of investment at any time and is simply reduced to:

$$i_t^* = \arg \max_{i \in \{0,1\}} \{y_{it}\}. \tag{5}$$

As a consequence $\nu_{it} = n_{it} + \gamma_i$ where n_{it} is the number of times technology i has been chosen in the past. On the other hand investment does depend on the choice of technology, but we can obtain an equivalent problem by considering $\{n_{it}\}$ as a random process independent of agents' choices and rewriting the inequalities of assumption C as:

$$\begin{array}{l} k_{t+1} + c_t \leq y_{i^*t} = Z_{i^*t} g(k_t) \text{ a.s., } t > 0 \\ k_t, c_t \geq 0 \text{ a.s., } t > 0. \end{array} \tag{6}$$

[10] Assumptions H, T and M imply that all the agents have the same commodity space and utility function, therefore we will not introduce an indexation for agents.

[11] Otherwise, given the cumulative structure of learning implied by assumption T, the model would be reduced to a standard stochastic growth model with exogeneous technical progress.

Since Z_{i^*t} is a random variable this problem becomes simply a standard problem of stochastic growth with non-stationary shocks and TAS preferences [see e.g. Joshi (1995)][12]. Hence we can simply study the choice of technique process and then, once we have obtained a sufficient characterization of that process (and consequently of the process $\{Z_{i^*t}\}$) we can use it to study the growth model subject to (6) and derive properties of the whole growth system.

Let us assume without loss of generality that $A_1 \geq A_0$ and that if $y_{1t} = y_{0t}$ then agents choose technology 1[13]. Hence agents choose $i = 1$ if and only if

$$Z_{1t} g(k_t) \geq Z_{0t} g(k_t). \tag{7}$$

More precisely:

$$e^{A_1(n_{1t}+\gamma_1)}\epsilon_{1t} \geq e^{A_0(n_{0t}+\gamma_0)}\epsilon_{0t}. \tag{8}$$

Passing to logarithms (8) becomes

$$A_1(n_{1t}+\gamma_1) - A_0(n_{0t}+\gamma_0) \geq log\epsilon_{0t} - log\epsilon_{1t}. \tag{9}$$

Let us denote with ζ_t the random variable on the right hand side. Since if ϵ_{1t} is i.i.d. and ϵ_{0t} is i.i.d., the probability of choosing i=1 at any time t is just a function of $A_1 n_{1t}$ and $A_0 n_{0t}$ and therefore the choice of technology process can be interpreted as a Markov chain $\{i_t^*, W_t\}$ on $\{0,1\} \times \mathbf{R}$ with transition probabilities given by:

$$\begin{aligned} \mu(i_t^* = 1, W_t = x + A_1 \,|W_{t-1} = x) &= F_\zeta(x) \\ \mu(i_t^* = 0, W_t = x - A_0 \,|W_{t-1} = x) &= 1 - F_\zeta(x) \end{aligned} \tag{10}$$

where F_ζ is the distribution function of the random variables ζ_t. Essentially from this observation all the results of this section follow.

The next proposition establishes that the assumptions of the model are sufficient conditions for the choice of technique process to be path-dependent.

Proposition 2.1 *The choice of technology process converges either to the non-random choice of i=1 or to the non-random choice of i=0 with probability 1. Both limits occur with positive probability. In other words, according to definition 1.1, the choice of technology process is path-dependent.*

The logic of the proof is the following: Given that the distributions of the stochastic shocks have bounded support it is always possible to find a sequence of realizations such that one technology has been chosen sufficiently more often than the other to make the productivity gap large enough that there cannot be any random shock sufficiently strong to revert the choice of technology. Therefore the set of states which allows both technologies to be chosen with positive probability is transient and, by a standard property of Markov chains on compact state space, the process will be absorbed in finite time.

This result has important implications in terms of Total Factor Productivity (TFP) growth. In this model we can simply define TFP as Z_t since:

$$\frac{\Delta y_t}{y_t} = \frac{Z_{t+1}g(k_{t+1}) - Z_t g(k_t)}{Z_t g(k_t)} = \frac{\Delta Z_t}{Z_t} + \frac{Z_{t+1}}{Z_t}\frac{\Delta g(k_t)}{g(k_t)}, \tag{11}$$

[12] Given that the auxiliary problem does not involve externalities, the resulting decentralized equilibrium (with distinct consumers and firms) would give the same solutions of the planner's problem.

[13] Any other random choice of technique, when $y_{1t} = y_{0t}$ can be easily accommodated.

which is equal to $\Delta Z_t / Z_t$ if $k_{t+1} = k_t$.
We have the following result:

Proposition 2.2 *The TFP growth rate process is path-dependent. Moreover there are two stochastic steady states with different mean and the process converges to one of them in finite time with probability 1. The difference between the means increases with the difference between the technological opportunities of the two technologies.*

The proof exploits the one-to-one correspondence between the choice of technology and the distribution of TFP growth rate that, conditional to the choice of technology, is stationary because of the stationarity of the distributions of the random shocks.

We can now pass onto considering agents' optimal choice of capital stock and income growth.

As said before, at any time t agent's optimal choice of capital stock problem is reduced to attaining the supremum of expected utility subject to (6), that is:

$$\begin{array}{l} V_t(k_t) = \sup \sum_{s=1}^{\infty} \beta^{s-1} E\left(u(c_{t+s}) \,|\mathcal{F}_t\right), \, t \geq 0, \\ \text{s.t.} \\ k_{t+s+1} + c_{t+s} \leq y_{i^* t+s} = Z_{i^* t+s} g(k_{t+s}) \text{ a.s.}, \\ k_{t+s}, c_{t+s} \geq 0 \text{ a.s.}, \end{array} \tag{12}$$

where $\{Z_{i^*t}\}$ is a non-stationary random process whose properties are now known. As it is standard in the literature on optimal stochastic growth, the program $c^* = \{c_t^*\}$ or $k = \{k_t^*\}$which attains the supremum is said to to be the optimal program. Given the assumption that $\beta < e^{\frac{A_1 \gamma}{\alpha - 1}}$, the problem can be reformulated in a way that existence and uniqueness of optimal programs can be proved by standard arguments.

The following proposition derives strong results on convergence to multiple steady states for the growth rate of output.

Proposition 2.3 *Both capital stock and output growth rates are path-dependent, have two stochastic steady states with different mean and they converge to one of them with probability 1. The difference between the means increases with the difference between the technological opportunities of the two technologies.*

This result is simply a consequence of the properties of the TFP. In fact, knowing that in finite time the distribution of TFP growth rate becomes invariant, we can rescale the dynamic programming problem and then apply any standard ergodic theorem for stochastic growth models.

This section has shown that, subject to mild assumptions on the stochastic shocks, if technological opportunity is different across techniques *both capital stock and output growth rates are path-dependent.* Growth rates have multiple steady states, characterized by fast vs. slow growth properties. *Which type of equilibrium eventually emerges depends on the sequence of historical events.* Nothing guarantees that the dynamically more efficient path will be actually attained.

Even though the methodology employed to obtain the result cannot be applied beyond the particular specific exponential form that we have given to production function shifts as a consequence of learning-by-doing, the result is however quite robust. For instance, suppose to replace (1) with:

$$y_{it} = (A_i \nu_{it})^{\phi} \epsilon_{it} k_t^{\alpha}, \quad 0 < \alpha < 1, \phi > 0.$$

Even if absolutely unusual in the literature a functional form like this can be justified to obtain a technological opportunity which is decreasing in the time spent in that technology, thus allowing technologies to become "older" as time goes by[14]. We could repeat the same reasoning which lead from eq. (7) to eq. (9), however this time extracting the logarithm would not help because the logarithm is not a linear operator. The process could not therefore be so easily represented as a time-homogeneous Markov chain. However by elevating both sides of eq. (7) to the power of $1/\phi$ and setting $X_t = \nu_{1t}/(\nu_{1t} + \nu_{0t}) = (n_{1t} + \gamma_1)/(t + \gamma_1 + \gamma_0)$, we get the following inequality:

$$A_1 X_t / A_0 (1 - X_t) \geq (\epsilon_{0t}/\epsilon_{1t})^{1/\phi}.$$

Denoting the random variable on the right hand side with ζ_t, we have that if ϵ_{1t} is i.i.d. and ϵ_{0t} is i.i.d., the probability of choosing $i = 1$ at any time t is just a function of X_t and therefore the choice of technology process can be interpreted as a non-linear Polya process[15] with urn function f given by

$$f(X) = \mu(i_t^* = 1 \,|X_t = x) = F_{\zeta}(A_1 x / A_0 (1 - x)).$$

Proposition 2.1 would be amended in the following way:

Proposition 2.4 *The choice of technology process converges either to the non-random choice of i=1 or to the non-random choice of i=0 with positive probability. From any initial conditions both these two limits occur with positive probability. In other words, according to definition 1.1, the choice of technology process is path-dependent and has at least two steady states. If the set* $B = \{x \in \mathbf{R} : F_{\zeta}(x) = A_0 x/(A_1 + A_0 x)\}$ *contains only isolated points then the number of stochastic steady states to which the choice of technology process converges with positive probability is less than* $card(B)$.

Interestingly this proposition establishes that it is not sure that the choice of technology locks in one of the two options, but it may converge to a proper stochastic distribution with continuous reswitching of technologies according to the shocks. This possibility is due to the fact that without exponential growth it is possible that no technology forges ahead before the rate of growth becomes too small.

The other two propositions however would be maintained with appropriate minor changes.

3 Stochastic Technological Change

In the previous section we have considered an exogenous stochastic environment represented by shocks affecting production functions. However, while those shocks can be interpreted as exogenous contingencies affecting production (such as weather conditions), it would be hard to identify them as exogenous stochastic technological progress, because it would lack cumulativity. On the other hand the endogenous component of technological change (production function shifts due to learning-by-doing) is not stochastic. Therefore

[14]For a different justification of a similar functional form see Jones (1995).

[15]Non-linear Polya processes are synthetically described in Appendix A.

at least one of the attributes of technological change we highlighted in the introduction is misrepresented in that model. Moreover that model allows only for a given number of technologies: at any point in time either the introduction of a new technology would be an event with zero probability or it would contradict the assumption of rational expectations. In a sense, that model draws the picture of a situation where technological revolutions are no more possible[16].

On the contrary, in the model that we consider here, both learning-by-doing and exogenous technological change are stochastic and purely cumulative (both of them are represented as processes with one unit root). Moreover, the upper bound to the support of the distribution of both learning-by-doing and exogenous technical progress can be arbitrarily large to allow for the representation of history as the outcome of both rare big events and small (frequent) events[17]; thus this unifies both approaches described in the introduction.

Consider the dynamic stochastic economy of previous section and replace assumption T, S and NT with the following:

Technology:

T') At time t available production functions are described by the following equation:

$$y_{it} = e^{\nu_{it}} k_t^{\alpha}, \quad 0 < \alpha < 1, \tag{13}$$

where $t \geq 1$, $i = 0, 1$, $\nu_{it} = \sum_{k=1}^{t-1} (\theta_{ik}\psi_{ik} + \xi_{ik}) + \gamma_i$, θ_{ik} is the share of agents that chose technology i at time k, $\psi_{ik} \geq 0$ is a stochastic variable representing learning-by-doing for technology i at time k, $\xi_{ik} \geq 0$ is a stochastic variable representing exogenous technological change for technology i at time k, γ_i are initial integer parameters; there is no need to specify a technological opportunity coefficient since it is already included in the parameters of the distribution of the stochastic shocks; y and k are per capita output and capital stock respectively. As before, to simplify the notation we will write $y_{it} = Z_{it} g(k_t)$, whenever convenient.

Stochastic environment:

S') The stochastic variables ψ_{it} and ξ_{it}, $i = 1, 0$, are serially i.i.d. and have the following properties:

$$\begin{aligned} &\psi_{it} \in \mathbf{Z}_+ \text{ , with bounded support,} \\ &\xi_{it} \in \mathbf{Z}_+ \text{ , with bounded support,} \\ &\mu(\psi_{it} = 0) \leq L < 1/2. \\ &\mu(\psi_{it} = 1) \geq M > 0 \end{aligned} \tag{14}$$

Assumption S' is technical in nature and it is not strictly necessary but it greatly simplifies the analysis[18]. Notice that the stochastic shocks representing exogenous technological change may be zero with very high probability and very large with very low probability representing in that case rare big events.

[16] Obviously this last logical problem would not arise if we rewrite the model assuming overlapping generations instead of infinitely lived agents.

[17] To obtain asymptotically stationary growth rates we make the assumption that shocks are bounded above (assumption S' below). Technically we can relax this assumption to shocks with bounded mean and variance; however in this case the steady states should be defined only as mean-stationary growth rates.

[18] In this way we can rely on many already established results on non-linear Polya processes with multiple additions. However, if the domain of the shocks were $\mathbf{R}^+$, the analysis could be carried over anyway by resorting to the more general class of Robbins-Monro algorithms (see next section and appendix A).

Non-triviality condition:

NT') Denoting with $S(X)$ the support of a random variable X, for $i \neq j$, either $\max S(\xi_{it}) > \min S(\xi_{jt} + \psi_{jt})$ or $\gamma_i \geq \gamma_j$.

Assumption NT' allows initial situations where a technology is not competitive but it may become competitive in the future as the result of exogenous technological change. In general this framework can be used to analyze the dynamics of adoptions of different competing technologies with different technological opportunity, like the model of the previous section, or it can be used to analyze the effects of potential introduction of a new (and possibly more dynamic) technology in an economy with an already established old one. The second interpretation will be more emphasized, because it leads to interesting results when passing from a two technology model to a multiple technology model[19].

Hereafter we will call technology 1 as "new" and technology 0 as "old". Without loss of generality it will be assumed that the means of the random variables of technology 1 are larger than (or at least equal to) the means of the random variables relative to technology 0. Moreover it could be assumed that $\gamma_0 \geq \gamma_1$, to represent a situation where the alternative technology is not yet competitive but agents have rational expectations on the probability of the occurrence of the "innovation"[20]. However even when an alternative technology has been adopted there is still the possibility that a major "innovation" occurs in the old technology, reverting the pattern of adoption to the old renewed technology[21].

The analysis of the output growth patterns is greatly simplified if in addition to previous assumptions we replace assumption U with:

Preferences:

U') Intertemporal preferences are represented by a TAS utility function $U : \mathbf{R}_+^\infty \rightarrow \mathbf{R}$ with unit intertemporal elasticity of substitution that is:

$$U(c) = \sum_t \beta^t \log(c_t), \quad 0 \leq \beta < 1. \tag{15}$$

As in the model of previous section, agent's problem consists in maximizing U subject to the feasibility costraint described by assumption C. There are two choice variables: technology and investment but, in equilibrium, because of assumption M, others' behavior and ν_{it} are taken as given, therefore in every period, the choice of technology is not affected by choices of investment at any time and is simply reduced to (5). As a consequence $\theta_{i^*t} = 1$. Again we can simply study the choice of technology process and then, once we have obtained a sufficient characterization of that process (and consequently of the process $\{Z_{i^*t}\}$) we can use it to study the growth model subject to (6) and derive properties of the whole growth system.

Let us assume without loss of generality that if $y_{1t} = y_{0t}$ then agents choose technology 1. Hence agents choose $i = 1$ if and only if:

$$Z_{1t} g(k_t) \geq Z_{0t} g(k_t). \tag{16}$$

More precisely passing to logarithms:

---

[19]However, as shown below, multiple technologies make possible to combine both interpretations at the same time.

[20]Alternatively, and perhaps more precisely, we can see this setup as representing a situation where a basic new (potentially revolutionary) idea is already born, but it is not yet competitively applicable to production so that it has to be developed in academic or research laboratories.

[21]This is much less unrealistic than it may seem if we interpret the term technology in a broad way as technological paradigm. Even when a technological paradigm seems on the way of decline it may be revitalized by new discoveries [see Dosi (1988)].

$$\nu_{1t} \geq \nu_{0t} \tag{17}$$

setting $X_t = \nu_{1t}/(\nu_{1t} + \nu_{0t})$, (17) becomes:

$$X_t \geq 1/2. \tag{18}$$

Consequently the choice of technology is a previsible ($\mathcal{F}_{t-1}$-adapted) stochastic process with:

$$\mu(i_t^* = 1, |X_t = x) = \begin{cases} 1 \text{ if } x \geq 1/2 \\ \\ 0 \text{ if } x < 1/2 \end{cases}. \tag{19}$$

Therefore the dynamics of the choice of technology depends entirely on the dynamics of the relative technological progress X_t.

The dynamics of X_t is by definition

$$X_{t+1} = X_t + \frac{\psi_{1t}^* + \xi_{1t} - X_t(\psi_{1t}^* + \xi_{1t} + \psi_{0t}^* + \xi_{0t})}{\nu_{1t+1} + \nu_{0t+1}}, \tag{20}$$

where

$$\psi_{1t}^* = \begin{cases} \psi_{1t} \text{ if } X_t \geq 1/2 \\ \\ 0 \text{ if } X_t < 1/2 \end{cases}, \tag{21}$$

and ψ_{0t}^* is similarly defined.

Therefore the process X_t can be seen as a non-linear Polya process with multiple additions[22] whose right-hand side of the associated Ordinary Differential Equation (ODE) is:

$$h(x) = \begin{cases} E(\psi_{1t}) + E(\xi_{1t}) - xE(\psi_{1t} + \xi_{1t} + \xi_{0t}) \text{ if } x \geq 1/2 \\ \\ E(\xi_{1t}) - xE(\xi_{1t} + \psi_{0t} + \xi_{0t}) \qquad \text{otherwise} \end{cases}. \tag{22}$$

From these observations all the results of this section follow.

Next proposition establishes necessary and sufficient conditions for the choice of technology to be path-dependent.

Proposition 3.1 *If exogenous technological progress has the same expected rate for both technologies, then the choice of technology process converges to the old (i=0) or the new (i=1) technology with positive probability. From any initial conditions both limits occur with positive probability, hence the process is path-dependent. More precisely path-dependence holds if the difference between the expected rates of exogenous technological progress of the two technologies is less than the expected rate of potential learning-by-doing in the old technology. Otherwise the choice of technology converges almost surely to the new technology.*

Roughly speaking technological progress is represented here as a process which has many similar properties to a random walk with drift: At any time it is always possible to find a sequence of outcomes with positive probability that can lead one technology (either

[22] See appendix A.

old or new), which is lagging behind, to overtake the other. In general this probability will asymptotically tend to zero because the drift of the leading technology tends to be greater than that of the laggard. More technically, as discussed in Appendix A, asymptotic values for X_t have to be searched among the zeroes of the associated ODE. However, in the path-dependent case, we are not able to rule out an asymptotic pattern where the choice of technology cycles continually between the old and the new technology, because of the discontinuity of the ODE.

Essentially path-dependence arises when the new technology is not sufficiently dynamically superior to the old one. Roughly speaking, to avoid path-dependence it is necessary that the amount of technological progress that can occur *without the adoption* of the new technology is on average higher than the technological progress which is obtainable *by producing* with the old technology. Clearly this condition may apply when a certain technological paradigm is nearly exhausted, while it is hardly the case if the old technology is a sufficiently young one with still strong technological dynamics. Timing of innovations becomes a relevant factor. However this issue can be better tackled with a higher number of competing technologies. Proposition 3.4 at the end of the section will do the job.

The results on TFP and output growth are correspondingly the following:

Proposition 3.2 *If the difference between the expected rates of exogenous technological progress of the two technologies is less than the expected rate of potential learning-by-doing in the old technology the TFP growth rate process is path-dependent and there are two stochastic steady states with different mean and the process converges to one of them with positive probability. Otherwise the TFP growth rate is ergodic.*

As before, the proof exploits the one-to-one correspondence between the choice of technology and the distribution of TFP growth rate that, conditional to the choice of technology, is stationary because of the stationarity of the distributions of the random shocks.

As for capital stock and output growth the following proposition derives strong results on convergence to multiple steady states for the growth rate of output.

Proposition 3.3 *If the difference between the expected rates of exogenous technological progress of the two technologies is less than the expected rate of potential learning-by-doing in the old technology, then both capital stock and output growth rates are path-dependent, have two stochastic steady states with different mean and converge to one of them with positive probability; otherwise they are ergodic.*

This result is simply a consequence of the properties of the TFP and of the properties of the logarithmic utility. In fact, we can rescale the dynamic programming problem in a way that it can be solved as a standard ergodic growth one.

Assume now that there are three technologies, say 0,1,2. 0 is the old technology, with both lower exogenous technological progress and potential learning-by-doing. 1 and 2 are new technologies with same expected rate of exogenous technological progress but different potential learning-by-doing, say 1 has larger expected rate of potential learning-by-doing than 2. Therefore a dynamic path of higher output growth rate is associated to 1. To avoid asymptotic survival of the old technology let us also assume that the difference between the expected rates of exogenous technological progress is larger than the expected rate of potential learning-by-doing in the old technology. As expectable, these conditions are not sufficient to guarantee ergodicity; rather the reverse is true:

Proposition 3.4 *If exogenous technological progress has the same expected rate for both the new technologies and the gap with the expected rate of exogenous technological progress in the old technology is larger than the expected rate of potential learning-by-doing in the old technology, then almost surely the old technology do not survive asymptotically but the choice of technology is path-dependent and converges with positive probability to one of the new technologies. From any initial condition both limits occur with positive probability.*

The proof is technically more complicated than that of proposition 3.1, but its logic is the same.

In this section we have shown that both capital stock and output growth rate can be path-dependent only if the new technology is not sufficiently dynamically superior to the old one. However, still, if there are at least two techniques whose technological opportunity is not too diverse, growth rates have multiple steady states, characterized by fast vs. slow growth properties; the type of equilibrium that eventually emerges depends on the sequence of historical events and the dynamically more efficient path may not be actually attained.

Timing of innovation is therefore a crucial issue. The technology which is firstly introduced has positive probability of survival regardless to the fact that the other potential "new" technology may be dynamically better, provided that the latter is not "too" superior. Better technologies and growth opportunities may be ruled out by historical contingencies. In a sense this result parallels that one obtained in some diffusion models with vintage capital and vintage-specific learning-by-doing [e.g. Silverberg et al. (1988)].

Suppose finally to drop assumption NT', for instance by not allowing for exogenous technological progress ($\xi_{it} = 0$, for $i = 0, 1, 2$). In this case initial conditions determine the outcome, as if the model were deterministic. This is essentially the result of Jovanovic and Nyarko (1996). In their model the only stochastic technological change is the learning-by-doing component. Here this result arises as a special case because of the richer assumed stochastic structure of technological change.

We have established necessary and sufficient conditions for nonergodicity subject to a very strong assumption of homogeneity of agents. At this point it becomes interesting to analyze whether relaxing that assumption can restore ergodicity. This task is accomplished in the next section.

4 Heterogeneity of Technological Capabilities

In a realistic world agents typically differ in their endowment, production capabilities and preferences. Asymmetry and variety are evident in the distribution of firms' performance. Partial tacitness of knowledge and partial unavailability of information at no cost induce variety and asymmetry in the distribution of the capability to manufacture.

The dynamics of the distribution of capabilities is an economic process that deserves to be investigated. However different individuals and different organizations are naturally more geared to different technological paradigms. As a first approximation, in this section we amend the world described in the previous section by allowing agents to be heterogeneous in their "innate" ability of mastering different technologies[23]. A distribution of abilities is introduced. Given that our framework allows only for one single product and

[23] The type of heterogeneity which is introduced here is therefore similar to that introduced by Arthur (1989).

a finite number of technologies, the degree of asymmetry and variety are summarized by the same single parameter.

The underlying economic intuition of this section is the following: If agents have different natural inclinations towards different technologies, even in the presence of strong spillovers, they tend to produce with their "naturally" preferred technology provided that the alternative is at the moment not too superior. In such a way they explore a wider set of technologies, keeping alive a higher number of them. The probability of survival of many technologies in the short-run is higher the wider the heterogeneity of capabilities. If one technology is by far potentially superior to the others, from the analysis of previous sections we should expect this to be sufficient to let it dominate in the long-run. Proposition 4.1 addresses this point.

In a less aggregate world different firms are specialized in producing different goods and services with different technologies. As an aggregate approximation the type of heterogeneity which is represented here tries to capture heterogeneity in production tasks. Two general purpose competing technology can compete over different related production uses. The intuition is that the probability of survival of both technologies in the short-run is higher the more diverse the production tasks where both can be employed. By surviving in a technological niche, one technology keeps being developed and so there is still a positive probability that major discoveries occur and eventually that technology diffuses again through a wider set of uses. For instance Tell (1997) shows that the fact that direct current electric technology survived in the railway transportation market segment was the key for the rediscovery of that technology as long distance electricity carrier many years later.

Let us modify assumption H and T' in the following way:

Heterogeneity of agents:

H") Production functions differ across agents; initial endowment k_1, preferences and stochastic shocks are the same for every agent.

Technology:

T") At time t available production functions for agent $j \in [0,1]$ are described by the following equations:

$$\begin{aligned} y_{j1t} &= e^{(1-c_j)\nu_{1t}} k_{jt}^{\alpha}, \quad 0 < \alpha < 1, \\ y_{j0t} &= e^{c_j \nu_{0t}} k_{jt}^{\alpha}, \quad 0 < \alpha < 1, \end{aligned} \tag{23}$$

where c_j is a deterministic agent-specific capability coefficient whose population is assumed to be uniformly distributed between $[1/2 - a, 1/2 + a]$, $a \in [0, 1/2)$. All other variables are as defined in assumption T'.

Heterogeneity of technological capabilities is represented by c_j; a is an index of the degree of heterogeneity ($a = 0$ gives previous section's model - no heterogeneity - as a particular case).

As in the model of previous section, agent's problem consists in maximizing U subject to the feasibility costraint described by assumption C. There are two choice variables: technology and investment but, in equilibrium, because of assumption M, others' behavior and ν_{it} are taken as given, therefore in every period, the choice of technology is not affected by choices of investment at any time and is simply reduced to:

$$i_{jt}^{*} = \arg \max_{i \in \{0,1\}} \{y_{jit}\}. \tag{24}$$

Again we can simply study the choice of technology process and then, once we have obtained a sufficient characterization of that process (and consequently of the process

$\{Z_{ji^*t}\}$) we can use it to study the growth model and derive properties of the whole growth system. However this time, because of agents' heterogeneity $0 \leq \theta_{it} \leq 1$.

Following the same steps that lead to eq. (17) we have that agent j chooses technology i at time t if and only if:

$$(1 - c_j)\nu_{1t} \geq c_j \nu_{0t} \tag{25}$$

setting $X_t = \nu_{1t}/(\nu_{1t} + \nu_{0t})$, (25) becomes:

$$X_t \geq c_j. \tag{26}$$

Consequently the choice of technology is a previsible ($\mathcal{F}_{t-1}$-adapted) stochastic process with:

$$\mu(i^*_{jt} = 1, |X_t = x) = \begin{cases} 1 \text{ if } x \geq c_j \\ \\ 0 \text{ if } x < c_j \end{cases} . \tag{27}$$

Therefore the dynamics of the choice of technology depends entirely on the dynamics of the relative technological progress X_t.

The dynamics of X_t is by definition:

$$X_{t+1} = X_t + \frac{F_a(X_t)\psi_{1t} + \xi_{1t} - X_t(F_a(X_t)\psi_{1t} + \xi_{1t} + (1 - F_a(X_t))\psi_{0t} + \xi_{0t})}{\nu_{1t+1} + \nu_{0t+1}}, \tag{28}$$

where $F_a(.)$ is the distribution function of $U(1/2 - a, 1/2 + a)$.

Therefore the process X_t can be seen as a stochastic approximation algorithm with stochastic step[24] whose right-hand side of the associated ODE is:

$$h(x) = \begin{cases} E(\psi_{1t}) + E(\xi_{1t}) - xE(\psi_{1t} + \xi_{1t} + \xi_{0t}) \text{ if } x \geq 1/2 + a \\ \\ \begin{array}{c} F_a(x)E(\psi_{1t}) + E(\xi_{1t}) + \\ -xF_a(x)[E(\psi_{1t}) - E(\psi_{0t})] + \\ -xE(\xi_{1t} + \psi_{0t} + \xi_{0t}) \end{array} \text{ if } 1/2 - a < x < 1/2 + a \\ \\ E(\xi_{1t}) - xE(\xi_{1t} + \psi_{0t} + \xi_{0t}) \quad \text{otherwise} \end{cases} \tag{29}$$

In the previous section we obtained that the asymptotic properties of the choice of technology depended on the difference between expected rates of potential technical change. Here we are interested in two questions: Is there any threshold in the degree of heterogeneity such that when trespassed it makes the system ergodic? How this threshold depends on the difference between expected rates of technological change? Before addressing these questions, let us introduce an index $d(x)$ of *relative difference between expected rates of technological change* when $X_t = x$:

$$d(x) = \frac{F_a(x)E(\psi_{1t}) + E(\xi_{1t}) - (1 - F_a(x))E(\psi_{0t}) - E(\xi_{0t})}{F_a(x)E(\psi_{1t}) + E(\xi_{1t}) + (1 - F_a(x))E(\psi_{0t}) + E(\xi_{0t})} \tag{30}$$

Next proposition contains the answer to the questions we have just formulated.

Proposition 4.1 *For any technological change distributions there exists a degree of heterogeneity that makes the system ergodic. More precisely there exists a threshold for the*

[24]See Duflo (1996).

index of the degree of heterogeneity such that above that threshold the choice of technology is an ergodic process, provided that the degree of heterogeneity is not too large. This threshold is a linear decreasing function of the index of relative difference between expected rates of technological change when $X_t = 0$.

The logic of the proof can be roughly summarized by the following argument: A positive degree of heterogeneity makes the associated $h(.)$ continuous and almost linear in the interval $[1/2 - a, 1/2 + a]$. For any given expected values relative to technological change, it is sufficient to take a large enough to avoid the associated deterministic equation having zeroes in the interval $[0, 1/2 - a]$ but small enough to have a zero in the interval $[1/2 + a, 1]$.

5 Concluding Remarks

In this paper we have discussed the relationship between localized technological change and path-dependent features of aggregate output growth. The growth model set forth in this paper show that, subject to mild assumptions on the stochastic processes representing the exogenous environment, path-dependence of aggregate output growth is a robust property of an economic system with localized learning-by-doing and diversity of technological opportunity. In general there are multiple steady states whose selection depends on the sequence of historical events: a priori all steady states are attainable, but only one of them will eventually emerge as the outcome of history. These results are the consequence of the fact that opportunity costs of switching technology increase with experience accumulation in a specific technology.

Path-dependence may be mitigated when the rate of technological change independent of experience is high: If one technology is by far dynamically superior to the others, growth turns out to be ergodic. Anyway we have noticed that this is hardly the case when there are many "young" technologies with different technological opportunity.

More interestingly a sufficient degree of heterogeneity of production capabilities of agents restores ergodicity, the intuition being that heterogeneity weakens the trade-off between exploration and exploitation of different technologies which is brought about by technological spillovers.

Another potential factor which may reduce path-dependence that deserves to be investigated is the effect of aggregation in multi-sector economies: It is trivial to show that the limit of an economy with an increasing number identical sectors and neither inter-sector linkages nor inter-sector spillovers has ergodic patterns even with path-dependent sector dynamics. On the other hand preliminary results of a related paper [Bassanini (1997)] show that if spillovers are strong enough path-dependence may hold.

Shocks are assumed to be serially independent. This is quite a strong assumption, because many exogenous phenomena that can affect productivity in a single period are typically serially correlated. Furthermore no cost of switching from one technique to the other is explicitly introduced, because of complete depreciation of capital stock. However the main theorems of section 2 still hold if we assume positively autocorrelated shocks. Moreover a more general framework to allow for serially correlated shocks and the presence of switching costs can be handled by resorting to some more general classes of adaptive algorithms that can be studied with stochastic approximation techniques[25]. Intuitively, however, if depreciation is sufficiently high relatively to the discount factor the results should hold.

[25] See Duflo (1996).

6 Appendix A: Non-Linear Polya Processes

Let Ω_t be the set of states ω_t of the environment at date $t \geq 1$, with Ω_t endowed with the σ-algebra $\mathcal{E}_t$. The stochastic environment is given by the probability space (Ω, $\mathcal{F}$, μ) where $\Omega = \Pi_1^\infty \Omega_t$ is the space of all sequences $\omega = \omega_t$ such that $\omega_t \in \Omega_t, \forall\ t \geq 1$, $\mathcal{F} = \otimes_1^\infty \mathcal{E}_t$ is the smallest σ-algebra on Ω generated by the measurable cylindrical sets and μ is the measure on Ω . Let partial history define the filtration $\{\mathcal{F}_t = \otimes_1^t \mathcal{E}_k\}, t \geq 1$. Finally let us denote with $\{Z_t(\omega_t)\}, t \geq 1$, a sequence of random variables. We can define a sequence of random variables $\{\xi_t(\omega_t)\}$ such that $\forall t \geq 1$, $\xi_t(\omega_t) = 1$ if $\omega_t \in A_t^1$ and 0 otherwise. Let n_t be the number of times we obtained $\{\xi_t = 1\}$, $1 \leq n \leq t$, in a single realization of the process, γ_0, γ_1 and γ be fixed initial integer parameters with $\gamma_0 + \gamma_1 = \gamma$ and $X_t = (n_t + \gamma_1)/(t + \gamma)$. We can now state the following definition:

Definition A.1 *A random process $\{\xi_t(\omega_t)\}$, $t \geq 1$, defined as above is a two-color non-linear Polya process*[26] *if*

$$\mu(\xi_t = 1\,|\mathcal{F}_{t-1}\,) = \mu(\xi_t = 1\,|X_{t-1}\,)$$

and

$$\mu(\xi_t = 1\,|X_{t-1} = x\,) = f(x),$$

where$f(.) : R[0,1] \to [0,1]$ and $R[0,1]$ is the set of rational numbers in $[0,1]$. The function f is called urn function[27].

Since by definition we have

$$X_{t+1} = X_t + (t + 1 + \gamma)^{-1}[\xi_{t+1} - X_t] \tag{31}$$

the expectation of X_{t+1} conditional upon partial history up to time t can be written as

$$E(X_{t+1}\,|X_t\,) = X_t + (t + 1 + \gamma)^{-1}[f(X_t) - X_t] \tag{32}$$

then the fundamental result of the theory of Polya processes is that the system converges almost surely to the set of the appropriately defined[28] zeroes of the function:

$$h(X) = f(X) - X. \tag{33}$$

We will call (36) the right-hand side of the associated Ordinary Differential Equation (ODE)[29]. Formally:

Theorem A.1 [Arthur et al. (1983)] *The sequence $\{X_t\}$ defined as above converges almost surely to the set*

$$B = \left\{ x \in [0,1] : 0 \in \left[\liminf_{y \to x} h(y), \limsup_{y \to x} h(y) \right] \right\}.$$

[26] Extension to any finite number of colors is straightforward and will be explicitly done only for processes with general increments.

[27] Notice that X_t is just the frequence of the event $\{\xi_t = 1\}$ corrected by the initial parameters γ_0, γ_1 and γ. Therefore definition A.1 requires that X_t is a sufficient statistics for $\{\xi_t = 1\}$.

[28] The qualification "appropriately defined" is necessary because the urn function may not be continuous.

[29] For a general definition of the associated ODE see Benveniste et al. (1990).

These results can then be extended to show that convergence occurs with positive probability only to a subset of the set of zeroes of the urn function.. An isolated point $x \in B$ is called *stable* if for every small enough $\epsilon_1, \epsilon_2 > 0$

$$h(y)(y - x) < \delta(\epsilon_1, \epsilon_2) < 0, \tag{34}$$

for every $y \in R(0,1)$ such that $\epsilon_1 \leq |y - x| \leq \epsilon_2$; it is called *unstable* if for every small enough $\epsilon > 0$

$$h(y)(y - x) > 0, \tag{35}$$

for every $y \in R(0,1)$ such that $0 < |y - x| \leq \epsilon$.

Theorem A.2 [Arthur et al. (1983)] *Let $x \in B$ be a stable point in $(0,1)$ such that $x \in (a, b)$, $a < b$; if $f(y) > 0$ for $y \in (a, x)$ and $f(y) < 1$ for $y \in (x, b)$, then $\mu(\lim_{t\to\infty} X_t = x) > 0$ for every $X_1 \in (a, b)$.*

Theorem A.3 [Arthur et al. (1983)] *If $f(\gamma_1/(\gamma + t)) < 1$ for $t > 0$ and $\sum_{t>0} f(\gamma_1/(\gamma + t)) < +\infty$ then $\mu\,(\lim_{t\to\infty} X_t = 0) > 0$; also if $f((\gamma_1 + t)/(\gamma + t)) > 0$ for $t > 0$ and $\sum_{t>0}[1 - f((\gamma_1 + t)/(\gamma + t))] < +\infty$ then $\mu\,(\lim_{t\to\infty} X_t = 1) > 0$.*

Clearly, because of almost sure convergence, if all the assumptions of theorem A.3 hold, both the processes identified by X_t and ξ_t are path-dependent stochastic processes. Conditions of non-convergence to unstable points can be also given [see Hill et al. (1980) and Dosi et al. (1994)].

Let us now consider processes with multiple additions. Let us define two sequences of random variables $\xi_{it}(\omega_t)$, $i = 0, 1$, such that $\forall t \geq 1, \xi_{it}(\omega_t) \in \mathbf{Z}_+$ and $n_{it} = \sum_{k=1}^{t} \xi_{ik}(\omega_t)$; let γ_0, γ_1 and γ be fixed initial integer parameters with $\gamma_0 + \gamma_1 = \gamma$ and $X_t = (n_{1t} + \gamma_1)/(n_{0t} + n_{1t} + \gamma)$. We can state the following definition:

Definition A.2 *A random process $\{\vec{\xi_t}(\omega_t)\}$, $t \geq 1$, defined as above is a two-color non-linear Polya process with multiple additions if, $\forall t \geq 1$, $j_i \in \mathbf{Z}_+$, $i = 0, 1$,*

$$\mu(\xi_{it} = j_i\,|\mathcal{F}_{t-1}) = \mu(\xi_{it} = j_i\,|X_{t-1})$$

and

$$\mu(\xi_{it} = j_i\,|X_{t-1} = x) = q(x, j_i)$$

where $q(.,.) : R[0,1] \times \mathbf{Z}_+ \to [0,1]$.

The right-hand side of the associated ODE in this case becomes:

$$h(X) = E(\xi_{1t}\,|X_{t-1} = X) - X\,[E(\xi_{0t} + \xi_{1t}\,|X_{t-1} = X)]\,. \tag{36}$$

The generalization of theorem A.1 is the following:

Theorem A.4 [Arthur et al. (1987b)] *Assume that*

$$q(., 0) \leq L < 1/2$$

and

$$\left[E((\xi_{0t}+\xi_{1t})^2\,|X_{t-1}=X)\right]\le K<+\infty;$$

the sequence $\{X_t\}$ *defined as above converges almost surely to the set*

$$B=\left\{x\in[0,1]:0\in\left[\liminf_{y\to x}h(y),\limsup_{y\to x}h(y)\right]\right\}.$$

For compactness we will give the generalization of theorem A.2 after having defined the multiple-color non-linear Polya process. Let us define N sequences of random variables $\{\xi_{it}(\omega_t)\}$, $i=1,2,...,N$ such that $\forall t\ge 1$, $\xi_{it}(\omega_t)\in\mathbf{N}\cup\{0\}$ and $n_{it}=\sum_{k=1}^{t}\xi_{ik}(\omega_t)$; let γ_i be fixed initial integer parameters and $X_{kt}=(n_{kt}+\gamma_k)/\sum_i(n_{it}+\gamma_i)$, $k=1,2,...,N-1$. We can state the following definition:

Definition A.3 *A random process* $\{\vec{\xi}_t(\omega_t)\}$, $t\ge 1$, *defined as above is a N-color non-linear Polya process with multiple additions if,* $\forall t\ge 1$, $\vec{j}\in\mathbf{Z}_+^N$,

$$\mu(\vec{\xi}_t=\vec{j}\,|F_{t-1})=\mu(\vec{\xi}_t=\vec{j}\,\middle|\vec{X}_{t-1})$$

and

$$\mu(\vec{\xi}_t=\vec{j}\,\middle|\vec{X}_{t-1}=\vec{x})=q(\vec{x},\vec{j}),$$

where $q(.,.):(R[0,1])^{N-1}\times\mathbf{Z}_+^N\to[0,1]$.

The right-hand side of the associated ODE ($N-1$ equations) is now:

$$h(\vec{X})=E(\vec{\xi}_t\,\middle|\vec{X}_{t-1}=\vec{X})-\vec{X}\left[E(\sum_{i=1}^{N}\xi_{it}\,\middle|\vec{X}_{t-1}=\vec{X})\right].\tag{37}$$

To simplify the notation, given that there is no ambiguity, the superscript arrow for vectors will be dropped hereafter. The generalization of theorems A.1 and A.4 is the following:

Theorem A.5 [Arthur et al. (1987b)] *If all the following conditions hold true:*

(i) $q(.,0)\le L<1/2$,

(ii) $\left[E\left((\xi_t)^2\,|X_{t-1}=X\right)\right]\le K<+\infty$,

(iii) There is a function $F:\mathbf{R}^{N-1}\to\mathbf{R}$ *which satifies the Lipschitz condition over an open set containing*

$$T_{N-1}=\left\{x\in[0,1]^{N-1}:\sum\nolimits_i x_i\le 1,\ i=1,2,...,N-1\right\}$$

and such that for any $z\in T_{N-1}\backslash B^h(T_{N-1})$

$$\inf_{g\in A^h(z)}\lim_{t\downarrow 0}\frac{\left|F(z+tg\,\|g\|^{-1})-F(z)\right|}{t}=\eta(z)>0,$$

where $B^h(z)$ *and* $A^h(z)$ *are convex hulls containing* $B_h(z)$ *and* $A_h(z)$ *respectively,* $A_h(z)=\{g:\exists y_k,\neq z,y_k\in Int(T_{N-1}),y_k\to z,h(y)\to g\}$, $B_h(z)=A_h(z)\cup\{h(z)\}$ *if* $z\in Int(T_{N-1})$ *and* $B_h(z)=A_h(z)$ *otherwise, and* $\|.\|$ *is the euclidean norm in* $\mathbf{R}^{N-1}$,

(iv) $F(B^h(T_{N-1}))$ *does not contain non-degenerate segments,*
then almost surely $\lim_{t\to\infty} \left\| X_t - B^h(T_{N-1}) \right\| = 0$.

We can now give the generalization of theorem A.2. Unfortunately for processes with general increments no conditions of non-convergence to unstable points exists when the function q is not continuous with respect to the first variable. Since the existing one is of no use for us we will not present it here[30].

Theorem A.6 [Arthur et al. (1988)] *Let* $x \in B^h(T_{N-1}) \cap T_{N-1}$ *be an isolated stable point, that is, for every small enough* $\epsilon_1, \epsilon_2 > 0$ *there exists symmetric positive definite matrix* Q *such that* $Qh(y)(y-x) < \delta(\epsilon_1, \epsilon_2) < 0$, *for every* $y \in Int(T_{N-1})$ *such that* $\epsilon_1 \leq \|y - x\| \leq \epsilon_2$ *; if all the following conditions hold true:*
(i) $q(.,j)$ *are continuous in the first argument for every* $j \in \mathbf{Z}_+^N$,
(ii) $q(.,0) \leq L < 1/2$,
(iii) $\left[E((\xi_t)^2 \,|X_{t-1} = X)\right] \leq K < +\infty$,
(iv) $\forall j \in \mathbf{Z}_+^N$, $q(x,j) \in (0,1) \Rightarrow q(z,j) \in (0,1)$, $\forall z \in Int(T_{N-1})$,
then $\mu(\lim_{t\to\infty} X_t = x) > 0$ *for every* X_1.

As a final important remark we can observe that if we generalize Polya processes to real increments the conclusions of theorem A.4 still hold provided that we suitably modify the assumption on $q(.,0)$. This happens because the fact that increments are integer never enters in the proof of that theorem provided by Arthur et al. (1987b). We put this as a remark:

Remark A.1 *If* $\xi_t(\omega_t) \in \mathbf{R}_+^N$ *and all the following conditions hold true:*
(i) $\exists \psi > 0$ *such that* $\mu(\max_{0\leq i\leq N-1}\{\xi_{it}\} < \psi \,|X_{t-1} = x) \leq L < 1/2$,
(ii) $\left[E((\xi_{0t} + \xi_{1t})^2 \,|X_{t-1} = X)\right] \leq K < +\infty$,
then the sequence $\{X_t\}$ *defined as above converges almost surely to the set*

$$B = \left\{ x \in [0,1] : 0 \in \left[\liminf_{y\to x} h(y), \limsup_{y\to x} h(y) \right] \right\} .$$

Many diffusion models employ the above machinery to formally represent path-dependent processes of technique or product diffusion when there are two or more competing technologies. The standard story of these models is the following: every period a new agent enters the market and has to choose the technology which is best suited to her needs, given her information structure and the available technologies. Because of network externalities (positive or negative) the actual or expected performance of both technologies depends on existing adoption shares, therefore the stochastic process of adoptions can be represented as a non-linear Polya process [see e.g. Arthur (1989) and Dosi et al. (1994)]. Many of these models can be seen as growth models as well, where the engine is the growth of the population of active agents in the market. However the dependence of the growth process on this demographic engine makes these models little appealing as growth models. Even worse, as noticed by David and Foray (1994) in a wider context, these models have other undesirable characteristics: first, every agent makes her choice only once for all, without any possibility of revision; second, micro-decisions are influenced by positive or

[30] See Arthur et al. (1988).

negative feedbacks from a macro-state of the system (the current frequency distribution of adoptions), ruling out the possibility of local interactions.

Anyway all these defects (from the point of view of a growth model) concern only most of existing applications of non-linear Polya processes and do not depend on the very nature of these processes. Dosi and Kaniovski (1994) and Bassanini and Dosi (1997) apply multiple-color non-linear Polya processes to study local interactions. Arthur (1993) employes non-linear Polya processes to develop an economic model representing the learning process of a single boundedly rational agent facing an uncertain environment.

Summarizing, a non-linear Polya process represents the evolution of a stochastic system where the probability of next events depends on the frequency distribution of previous events. In other words the probability of next events depends symmetrically on all previously occurred events. This fact makes it a natural candidate to represent economic growth driven by technological capability accumulation (in the presence of some discrete choice of technology): the cumulative nature of the latter makes the evolution of the system depend symmetrically on all previous increments of capabilities. Moreover this framework may apply to growth models without technological change provided that the evolution of the system depend on some cumulative process.

7 Appendix B: Propositions and Proofs

Proposition 2.1 *$\forall\omega \in \Omega$ a.e. $T(\omega) < +\infty$ such that $t > T(\omega)$ either $i_t^* = 0$ or $i_t^* = 1$. Moreover $\exists T > 0$ such that $\mu(i_t^* = 0, \forall t > T) > 0$ and $\mu(i_t^* = 1, \forall t > T) > 0$.*

Proof By assumption $\exists L > 0$ such that $\epsilon_{it} < L$ a.s., thus the process $\{i_t^*, W_t\}$ is absorbed in $i_t^* = 1$ if $\exists T > 0$ such that $W_T > logL - log\delta$. It suffices that $\exists T > (logL - log\delta - W_0)/A_1$ such that $i_t^* = 1$ $\forall t > T$ with positive probability. By assumption NT $0 < F_\zeta(W_0) < 1$, hence $\mu(i_t^* = 1, \forall t \leq T) = \prod_{t \leq T} F_\varsigma(W_t) > (F_\varsigma(W_0))^T > 0$. In the same way we can proceed for $i_t^* = 0$, so the second statement of the proposition is proved. Now notice that $\forall t > 0$ if $W_t > W_0$ then $W_{t+[(logL - log\delta - W_0)/A_1]+1} > logL - log\delta$ with probability $P_t > (F_\varsigma(W_0))^{[(logL - log\delta - W_0)/A_1]+1} > 0$. Similarly we can prove that if $W_t < W_0$ then $\exists s > 0$ such that $W_{s+t} < \log\delta - \log L$ with probability $P_t > (1 - F_\varsigma(W_0))^s > 0$; hence the set $\{W_t : \log\delta - \log L < Wt < \log L - \log\delta\}$ is transient and therefore, by a standard property of Markov chains, the first statement follows.

To facilitate the connection between statements in the main text and statements in the appendix we have to clarify what we mean with stochastic steady state when the process is nonergodic. In fact by definition there is only one limit distribution for every initial state and if the process is nonergodic there are infinite invariant distributions; hence the concept risks to be meaningless. We consider as stochastic steady state the invariant distributions defined on the single irreducible subchains on the ergodic sets since there is only one invariant distribution for every subchain. For example, proposition 2.1 shows that in finite time the TFP growth rate process enters into an ergodic set: thereafter the distribution of the TFP growth rate is stationary; in a sense it has reached the steady state.

Proposition 2.2 *For almost every $\omega \in \Omega$, $\exists T(\omega) < +\infty$ such that $\exists A_{T(\omega)} \in \mathcal{F}_{T(\omega)} \backslash \mathcal{F}_{T(\omega)-1}$, $\mu(A_{T(\omega)}) > 0$, $t > T(\omega)$, $\forall A_t \in \mathcal{F}_t$, $A_t \subseteq A_{T(\omega)}$, $\forall r, s > t$ $\mu\left(\frac{\Delta Z_s}{Z_s} \leq x \,|A_t\right) = \mu\left(\frac{\Delta Z_r}{Z_r} \leq x \,|A_t\right)$ and this distribution is stationary with $E\left(\frac{\Delta Z_s}{Z_s} \,|A_t\right) = e^{A_i} - 1$. Moreover*

$\mu\left(E\left(\frac{\Delta Z_s}{Z_s}\,|A_t\right) = e^{A_i} - 1\right) > 0$, $i = 0, 1$.

Proof From proposition 2.1 we know that i $_t^*$ is absorbed in finite time to either i $_t^* = 0$ or i $_t^* = 1$. Both limits occur with positive probability. Then notice that, by definition,

$$\begin{aligned} &\mu\left(\frac{\Delta Z_s}{Z_s} \le x\,|i^*_{s+1} = i, i^*_s = i\right) = \\ &= \mu\left(\frac{e^{A_i(n_{is}+1+\gamma_i)}\epsilon_{is+1}}{e^{A_i(n_{is}+\gamma_i)}\epsilon_{is}} \le x+1\,|i^*_{s+1} = i, i^*_s = i\right) = \\ &= \mu\left(\frac{e^{A_i}\epsilon_{is+1}}{\epsilon_{is}} \le x+1\right), \end{aligned}$$

which is stationary because of the i.i.d assumption. Moreover from the i.i.d. assumption we have trivially $E\left(\frac{\Delta Z_s}{Z_s}\,|i^*_{s+1} = i, i^*_s = i\right) = \frac{e^{A_i}E(\epsilon_{is+1})}{E(\epsilon_{is})} - 1 = e^{A_i} - 1$.

Proposition 2.3 *For almost every $\omega \in \Omega$, there exists a distribution function F such that $\forall \epsilon > 0$ and $\forall x$ of continuity of F, $\exists T(\omega) < +\infty$, $\exists A_{T(\omega)} \in \mathcal{F}_{T(\omega)} \backslash \mathcal{F}_{T(\omega)-1}$, $\mu(A_{T(\omega)}) > 0$, such that $\forall t > T(\omega)$, $\forall A_t \in \mathcal{F}_t$, $A_t \subseteq A_{T(\omega)}$, $\exists r > t$ $\forall s > r$ $\left|\mu\left(\frac{\Delta k_s}{k_s} \le x\,|A_t\right) - F(x)\right| < \epsilon$ and $\left|\mu\left(\frac{\Delta y_s}{y_s} \le x\,|A_t\right) - F(x)\right| < \epsilon$ with $\int \log x dF(x) = \frac{A_i}{1-\alpha}$. Moreover*

$$\mu\left(\int \log x dF(x) = \frac{A_i}{1-\alpha}\right) > 0, \quad i = 0, 1.$$

Proof From proposition 2.1 we know that $\exists T(\omega) < +\infty$ such that $\forall t > T(\omega)$ either $i^*_t = 0$ or $i^*_t = 1$ and that both limits may occur with positive probability. Let us assume without loss of generality $i^*_t = 1$ $\forall t > T(\omega)$. Then rewrite (12) in the following way:

$$\begin{aligned} &V(k_t) = \sup \sum_{s=1}^{\infty} \beta^{s-1} e^{\frac{A_1\gamma}{1-\alpha}(\nu_{1T}+t+s-T)} E\left(\frac{\hat{c}^{\gamma}_{t+s}}{\gamma}\,|\mathcal{F}_t\right), \\ &\text{s.t.} \\ &\hat{k}_{t+s+1} + \hat{c}_{t+s} = e^{\frac{A_1\gamma}{1-\alpha}} \epsilon_{i^*t+s} e^{A_{i^*}\nu_{i^*t+s} - A_1(\nu_{1T}+t+s-T)} \hat{k}^{\alpha}_{t+s} \text{ a.s.} \\ &e^{\frac{A_1\gamma}{1-\alpha}(\nu_{1T}+t+s-1-T)} k_{t+s} = \hat{k}_{t+s},\ c_{t+s} e^{\frac{A_1\gamma}{1-\alpha}(\nu_{1T}+t+s-T)} = \hat{c}_{t+s}. \end{aligned} \tag{38}$$

Notice that the constraints of this problem $\forall t > T(\omega)$ are simply:

$$\begin{aligned} &\hat{k}_{t+1} + \hat{c}_t = e^{\frac{A_1\alpha}{\alpha-1}} \epsilon_{1t} \hat{k}^{\alpha}_t \text{ a.s.}, \\ &\hat{k}_t > 0, \hat{c}_t > 0 \text{ a.s.}. \end{aligned} \tag{39}$$

Therefore (38) is nothing else than a standard optimal stochastic growth problem, and we can therefore use any standard proof of the ergodic theorem[31] for this problem to show that $\hat{k}_t$ converges to a unique stationary distribution independent of $\hat{k}_T$. This fact implies that also $\frac{\hat{k}_{t+1}}{\hat{k}_t}$ converges to a stationary distribution and so do both capital stock growth rate and output growth rate. Moreover we have that:

$$\lim_{s\to\infty} E\left(log\frac{\Delta k_{t+s}}{k_{t+s}}\,|A_t\right) = \lim_{s\to\infty} E\left(log\frac{\hat{k}_{t+s+1}}{\hat{k}_{t+s}} e^{\frac{A_1}{1-\alpha}} - 1\,|A_t\right) = \frac{A_1}{1-\alpha}, \tag{40}$$

and

[31] e.g. Mirman (1973). We can also use the technically simpler one by Brock and Mirman (1972), provided that we modify the original argument to correct for a non-general statement in the proof of the main theorem [see Bassanini (1996)].

$$\lim_{s\to\infty} E\left(log\frac{\Delta y_{t+s}}{y_{t+s}}\,|A_t\right) = \lim_{s\to\infty} E\left(log\frac{\hat{y}_{t+s+1}}{\hat{y}_{t+s}}e^{\frac{A_1}{1-\alpha}} - 1\,|A_t\right) = \frac{A_1}{1-\alpha}. \quad (41)$$

Proposition 2.4 *$\exists T > 0$ such that $\mu(i_t^* = 0, \forall t > T) > 0$ and $\mu(i_t^* = 1, \forall t > T) > 0$. Moreover if the set $B = \{x \in \mathbf{R} : F_\varsigma(x) = A_0x/(A_1 + A_0x)\}$ is contains isolated points then there exists a collection of distribution functions $\{F_j\}$ such that $\forall\epsilon > 0$, $\forall Fj$, $\exists T > 0$, $\forall t > T$, $\exists A_t \in \mathcal{F}_t \backslash \mathcal{F}_{t-1}$, $\mu(A_t) > 0$, $\forall A_t' \in \mathcal{F}_t$, $A_t' \subseteq A_t$, $\exists r > t$ $\forall s > r$ $|\mu\left(i_s^* \leq x\,|A_t\right) - F_j(x)| < \epsilon$ $\forall x$ of continuity of F_j.*

Proof The first statement can be proved as for proposition 2.1. The second statement follows from the definition of almost sure convergence and theorems A.1, A.2 and A.3, taking into account that the urn function is discontinuous only at unstable points because $F(.)$ is non-decreasing.

Proposition 3.1 *If $E(\xi_{1t} - \xi_{0t}) < E(\psi_{0t})$ then*

$$\mu\left(\lim_{t\to\infty} i_t^* = 0\right) > 0$$

and

$$\mu\left(\lim_{t\to\infty} i_t^* = 1\right) > 0.$$

If $E(\xi_{1t} - \xi_{0t}) > E(\psi_{0t})$ then

$$\mu\left(\lim_{t\to\infty} i_t^* = 1\right) = 1.$$

The proof is based on the following lemmas:

Lemma B.1 *Assume that two non-linear Polya processes with multiple additions but bounded increments generate the sequences X_t and X_t' with probability of increments $q(.,.)$ and $q'(.,.)$ $(q(.,.), q'(.,.) : (R[0,1])^{N-1} \times (\mathbf{N} \cup \{0\})^N \to [0,1])$; assume that q and q' agree almost everywhere in an open neighborhood O of a point $x \in (R[0,1])^{N-1}$; assume that for every $y \in O$ such that $\mu(X_t' = y) > 0$ then $\exists s < +\infty$ such that $\mu(X_s = y,\ n_{is} = n_{it}'$, $i = 0, 1, ..., N-1) > 0$; then $\mu(\lim_{t\to\infty} X_t' = x) > 0$ only if $\mu(\lim_{t\to\infty} X_t = x) > 0$.*

Proof Because of almost sure convergence $\mu(\lim_{t\to\infty} X_t' = x) > 0$ implies that and $\exists k_1 < +\infty$ and $\exists y \in O$ such that $\mu(\bigcap_{k=k_1}^{\infty} X_k' \in O \Big| X_{k_1}' = y) > 0$ and $\mu(X_{k_1}' = y) > 0$. Since $\mu(X_s = y,\ n_{is} = n_{ik_1}',\ i = 0, 1, ..., N-1) > 0$, then $\mu(\bigcap_{k=s}^{\infty} X_k \in O) > 0$, that is $\mu(\lim_{t\to\infty} X_t = x) > 0$.

Lemma B.2 If *$E(\xi_{1t} - \xi_{0t}) < E(\psi_{0t})$ then the process X_t converges almost surely to the set $B = \{x_1, 1/2, x_2\}$ where $x_1 < 1/2 < x_2$; moreover $\mu\left(\lim_{t\to\infty} X_t = x_1\right) > 0$ and $\mu\left(\lim_{t\to\infty} X_t = x_2\right) > 0$. If $E(\xi_{1t} - \xi_{0t}) > E(\psi_{0t})$ then $\mu\left(\lim_{t\to\infty} X_t = x_2\right) = 1$ where $x_2 > 1/2$.*

Proof If $E(\xi_{1t} - \xi_{0t}) < E(\psi_{0t})$, $h(X)$ has three appropriately defined zeroes in $[0, 1]$

corresponding to the elements of the set B; then the process X_t converges almost surely to the set B by theorem A.4. Consider a process X'_t whose probability of increments q' were equal to q in a neighborhood O of x_1 and whose associated deterministic equation is:

$$h'(X) = \begin{cases} h(x) & \text{if } x < 1/2 \\ E(\xi_{1t}) - xE(\xi_{1t} + \psi_{0t} + \xi_{0t}) & \text{otherwise} \end{cases} \quad . \tag{42}$$

By theorem A.6 X'_t converges to x_1 with positive probability. Moreover by definition we can write also $\exists \bar{k}_1 < +\infty$ such that $\forall k_1 > \bar{k}_1$ $\exists y \in O$ such that $\mu(\bigcap_{k=k_1}^{\infty} X'_k \in O \Big| X'_{k_1} = y) > 0$ and $\mu(X'_{k_1} = y) > 0$. Assumption S guarantees that if we take k_1 large enough $\exists s < +\infty$ such that $\mu(X_s = y,\, n_{is} = n'_{it},\, i = 0, 1, ..., N-1) > 0$. Since we can apply the same argument to a process $X''t$ for the convergence to x_2, by lemma B.1 the first statement follows. Since if $E(\xi_{1t} - \xi_{0t}) > E(\psi_{0t})$ then $h(X)$ has only one zero x_2 in $[0, 1]$ such that $x_2 > 1/2$, then by theorem A.4 the second statement follows.

Proof of Proposition 3.1 Notice that for any $x < 1/2$ there is an open neighborhhood such that $i^*_t = 0$; for any $x > 1/2$ there is an open neighborhood such that $i^*_t = 1$. Then apply lemma B.2.

Proposition 3.2 *If $E(\xi_{1t} - \xi_{0t}) < E(\psi_{0t})$ there exist two distribution functions F_i, $i = 1, 0$, with $\int \log x dF_i(x) = E(\xi_{it} + \psi_{it})$ such that $\forall \epsilon > 0$, $\forall F_i$, $\exists T > 0$, $\forall t > T$, $\exists A_t \in \mathcal{F}_t \backslash \mathcal{F}_{t-1}$, $\mu(A_t) > 0$, $\forall A'_t \in \mathcal{F}_t$, $A'_t \subseteq A_t$, $\exists r > t$ $\forall s > r$ $\left|\mu\left(\frac{\Delta Z_s}{Z_s} \leq x \,| A_t\right) - F_i(x)\right| < \epsilon$, $\forall x$ of continuity of F_i. If $E(\xi_{1t} - \xi_{0t}) < E(\psi_{0t})$ then $\forall \epsilon > 0$, $\forall t > 0$, $\exists A_t \in \mathcal{F}_t \backslash \mathcal{F}_{t-1}$, $\mu(A_t) > 0$, $\forall A'_t \in \mathcal{F}_t$, $A'_t \subseteq A_t$, $\exists r > t$ $\forall s > r$ $\left|\mu\left(\frac{\Delta Z_s}{Z_s} \leq x \,| A_t\right) - F_1(x)\right| < \epsilon$ $\forall x$ of continuity of F_1.*

Proof From Lemma B.2 we know that, if $E(\xi_{1t} - \xi_{0t}) < E(\psi_{0t})$, $\forall x_j$, $j = 1, 2$ $\forall \epsilon > 0$, $\forall \delta > 0$, $\exists T > 0$, $\forall t > T$ $\mu\left(\bigcap_{s=t}^{\infty} [|X_s - x_j| < \delta \,|(|X_t - x_j|) < \delta]\right) > 1 - \epsilon$ and any neighborhood of x_j can be reached with positive probability at any time t. For a sufficiently small δ we know that either $i^*_t = 0$ for any x in the neighborhood or $i^*_t = 1$ for any x in the neighborhood. Then notice that, by definition,

$$\begin{aligned} &\mu\left(\tfrac{\Delta Z_s}{Z_s} \leq x \,| i^*_{s+1} = i, i^*_s = i\right) = \\ &= \mu\left(\tfrac{e^{\nu_{is+1} + \xi_{is+1} + \psi_{is+1}}}{e^{\nu_{is+1}}} \leq x + 1 \,| i^*_{s+1} = i, i^*_s = i\right) = \\ &= \mu\left(e^{\xi_{is} + \psi_{is}} \leq x + 1\right), \end{aligned}$$

which is stationary because of the i.i.d assumption. Moreover from the i.i.d. assumption we have trivially $E\left(\log \frac{\Delta Z_s}{Z_s} \,| i^*_{s+1} = i, i^*_s = i\right) = E(\xi_{is+1} + \psi_{is+1})$. Taking $A_t = \{\omega \in \Omega : |X_t - x_j| < \delta\}$ then the first statement follows. Similarly the second statement can be proved.

Proposition 3.3 *If $E(\xi_{1t} - \xi_{0t}) < E(\psi_{0t})$ there exist two distribution functions F_i, $i = 1, 0$, with $\int \log x dF_i(x) = E(\xi_{it} + \psi_{it})/1 - \alpha$. such that $\forall \epsilon > 0$, $\forall F_i$, $\exists T > 0$, $\forall t > T$, $\exists A_t \in \mathcal{F}_t \backslash \mathcal{F}_{t-1}$, $\mu(A_t) > 0$, $\forall A'_t \in \mathcal{F}_t$, $A'_t \subseteq A_t$, $\exists r > t$ $\forall s > r$ $\left|\mu\left(\frac{\Delta k_s}{k_s} \leq x \,| A_t\right) - F_i(x)\right| < \epsilon$ and $\left|\mu\left(\frac{\Delta y_s}{y_s} \leq x \,| A_t\right) - F_i(x)\right| < \epsilon$ $\forall x$ of continuity of F_i. If $E(\xi_{1t} - \xi_{0t}) < E(\psi_{0t})$ then $\forall \epsilon > 0$, $\forall t > 0$, $\exists A_t \in \mathcal{F}_t \backslash \mathcal{F}_{t-1}$, $\mu(A_t) > 0$, $\forall A'_t \in \mathcal{F}_t$, $A'_t \subseteq A_t$, $\exists r > t$ $\forall s > r$ $\left|\mu\left(\frac{\Delta k_s}{k_s} \leq x \,| A_t\right) - F_1(x)\right| < \epsilon$ and $\left|\mu\left(\frac{\Delta y_s}{y_s} \leq x \,| A_t\right) - F_1(x)\right| < \epsilon$ $\forall x$ of continuity of F_1.*

The proof is based on the following lemma:

Lemma B.3 *Subject to assumptions* T', S' *and* U', *an optimal program for* (12) *implies* $k_{t+1} = \alpha\beta Z_{i^*t} k_t^{\alpha}$.

Proof Consider the Bellman functional equation for this problem:

$$\begin{aligned} & v(k_t, Z_{i^*t}) = \\ & = \sup_{k_{t+1}} \left\{ \log(Z_{i^*t} k_t^{\alpha} - k_{t+1}) + \beta E \left(v(k_{t+1}, Z_{i^*t+1}) \left| \mathcal{F}_t \right. \right) \right\}, \\ & t > 0, \quad s.t.(6). \end{aligned} \tag{43}$$

It is a textbook exercise to show that an optimal policy for this problem is $k_{t+1} = \alpha\beta Z_{i^*t} k_t^{\alpha}$. If we show that

$$\lim_{t\to\infty} \beta^{t+s} E \left(v(k_{t+s}, Z_{i^*t+s}) \left| \mathcal{F}_t \right. \right) = 0 \tag{44}$$

and that for any feasible program c such that (44) does not hold there exist a feasible program c' such that (44) holds and

$$\sum_{t=1}^{\infty} \beta^{t-1} E\left(\log(c_t)\right) \geq \sum_{t=1}^{\infty} \beta^{t-1} E\left(\log(c_t)\right), \tag{45}$$

then we can invoke a standard theorem of existence of optimal programs in the case of unbounded returns [see Stokey and Lucas (1989), theorem 9.12] to show that $k_{t+1} = \alpha\beta Z_{i^*t} k_t^{\alpha}$ attains the supremum also in (12). Notice that:

$$\begin{aligned} & v(k_t, Z_{i^*t}) = \tfrac{1}{1-\alpha\beta} \log(1-\alpha\beta) + \\ & + \tfrac{1}{1-\alpha\beta} \left(\alpha \log(k_t) + \sum_{j=0}^{\infty} \beta^j E\left(\log(Z_{i^*t+j}) \left| \mathcal{F}_t \right. \right) \right), t > 0. \end{aligned} \tag{46}$$

Hence, since by definition:

$$0 < \min_{i\in\{0,1\}} \{\gamma_i\} \leq \log(Z_{i^*t}) \leq t \max(\xi_{1t} + \psi_{1t} + \xi_{0t} + \psi_{0t}) + \Sigma\theta_{i0} \tag{47}$$

then for any $t > 0$:

$$v(k_t, Z_{i^*t}) \geq \frac{\log(1-\alpha\beta)}{1-\alpha\beta} + \frac{1}{1-\alpha\beta} \left(\alpha \log(k_t) + \frac{1}{1-\beta} \min_{i\in\{0,1\}} \{\gamma_i\} \right) \tag{48}$$

and

$$\begin{aligned} & v(k_t, Z_{i^*t}) \leq \tfrac{\log(1-\alpha\beta)}{1-\alpha\beta} + \tfrac{1}{1-\beta\alpha} \alpha \log(k_t) + \\ & + \tfrac{1}{1-\beta\alpha} \sum_{j=0}^{\infty} \beta^j ((t+j) E(\xi_{1t} + \psi_{1t} + \xi_{0t} + \psi_{0t}) + \Sigma\gamma_i). \end{aligned} \tag{49}$$

By iterating the optimal policy as before we have simply:

$$\log(k_t) = \sum_{s=1}^{t-1} \alpha^{s-1} \log(\alpha\beta) + \alpha^{t-1} \log(k_1) + \sum_{s=1}^{t-1} \alpha^{t-s+1} \log(Z_{i^*s}). \tag{50}$$

Plugging (50) into (48) and (49) and taking into account (47) again, it follows that $\sum_{s=0}^{\infty} \beta^{t+s} E\left(v(k_{t+s}, Z_{i^*t+s}) \left| \mathcal{F}_t \right. \right)$ is bounded above and below and therefore (44) hold.

Now consider an arbitrary feasible program (c, k) for which (44) does not hold. Since we know from (47) that $\sum_{s=0}^{\infty} \beta^{t+s} \sum_{j=0}^{\infty} \beta^j E\left(\log(Z_{i^*t+s+j})\,|\mathcal{F}_t\right)$ is bounded above and below, then the fact that (44) does not hold implies that $\sum_{s=0}^{\infty} \beta^{t+s} \log(k_{t+s})$ is unbounded. However the fact that $\log(k_t) \leq \alpha^{t-1}\log(k_1) + \sum_{s=1}^{t-1} \alpha^{t-s-1}(s-1)(\max(\xi_{1t} + \psi_{1t} + \xi_{0t} + \psi_{0t}) + \Sigma\gamma_i)$ implies that $\sum_{s=0}^{\infty} \beta^{t+s} \log(k_{t+s})$ is bounded above, therefore $\sum_{s=0}^{\infty} \beta^{t+s} \log(k_{t+s}) = -\infty$. From (12) we know that:

$$\sum_{t=1}^{\infty} \beta^{t-1} E\left(\log(c_t)\right) \leq \sum_{t=1}^{\infty} \beta^{t-1} \alpha \log(k_t) + \sum_{t=1}^{\infty} \beta^{t-1} E\left(\log(Z_{i^*t})\right), \tag{51}$$

hence (45) follows.

Proof of Proposition 3.3 Similarly to what we did for the proof of proposition 4, rewrite k_t as $e^{\frac{\nu_{i^*t-1}}{\alpha-1}} k_t = \hat{k}_t$. Therefore from lemma B.3 we have:

$$\begin{aligned} \log(\hat{k}_{t+s}) = & \sum_{u=1}^{s} \alpha^{u-1}\log(\alpha\beta) + \alpha^s \log(\hat{k}_{t+s}) + \\ & + \tfrac{1}{1-\alpha} \sum_{u=1}^{s} \alpha^{s-u+1}(\xi_{i^*t+u-1} + \psi_{i^*t+u-1}). \end{aligned} \tag{52}$$

Subject to the same conditions that make the choice of technology and the TFP growth rate converge as in proposition 3.1 and 3.2 we have that also $\log(\hat{k}_{t+s})$ converges; hence, under the same conditions $\log(\hat{k}_{t+s+1}) - \log(\hat{k}_{t+s}) = \log \frac{\hat{k}_{t+s+1}}{\hat{k}_{t+s}}$ converges and, by definition also $\log \frac{k_{t+s+1}}{k_{t+s}}$ and $\log \frac{y_{t+s+1}}{y_{t+s}}$. Moreover:

$$\begin{aligned} & \lim_{s\to\infty} E\left(\log \tfrac{\hat{k}_{t+s+1}}{\hat{k}_{t+s}}\,|A_t\right) = \\ & = \lim_{s\to\infty} E\left(\log(\hat{k}_{t+s+1})\,|A_t\right) - \lim_{s\to\infty} E\left(\log(\hat{k}_{t+s})\,|A_t\right) = 0. \end{aligned} \tag{53}$$

Under the same conditions, by definition, for i=0 or i=1

$$\begin{aligned} & \lim_{s\to\infty} E\left(\log \tfrac{\Delta k_{t+s}}{k_{t+s}}\,|A_t\right) = \\ & = \lim_{s\to\infty} E\left(\log \frac{e^{\frac{\nu_{i^*t+s}}{1-\alpha}}\hat{k}_{t+s+1}}{e^{\frac{\nu_{i^*t+s-1}}{1-\alpha}}\hat{k}_{t+s}}\,|A_t\right) = \\ & = \tfrac{1}{1-\alpha} E(\xi_{it} + \psi_{it}). \end{aligned} \tag{54}$$

As a consequence:

$$\lim_{s\to\infty} E\left(log \frac{\Delta y_{t+s}}{y_{t+s}}\,|A_t\right) = \frac{1}{1-\alpha} E(\xi_{it} + \psi_{it}). \tag{55}$$

Proposition 3.4 *If* $E(\xi_{1t} - \xi_{0t}) > E(\psi_{0t})$ *then*

$$\mu\left(\lim_{t\to\infty} i_t^* = 1\right) > 0$$

and

$$\mu\left(\lim_{t\to\infty} i_t^* = 2\right) > 0.$$

Moreover $\mu\left(S\left(\lim_{t\to\infty} i_t^*\right) = \{1,2\}\right) = 1$, *where* $S(Z)$ *denotes the support of the random variable* Z.

The proof is based on the following lemma:

Lemma B.4 *If* $E(\xi_{1t} - \xi_{0t}) > E(\psi_{0t})$ *then the process* X_t *converges almost surely to the set* $B = \{\vec{x}_1, I_{1,2}, \vec{x}_2\}$ *where* $\vec{x}_1$ *is such that* $x_{12} > 1 - 2x_{11}$ *and* $x_{21} < x_{11}$, $\vec{x}_2$ *is such that* $x_{22} > 1/2 - 1/2x_{21}$ *and* $x_{21} > x_{11}$, *and* $I_{1,2} \subset \{(x,y) : x = y, x > 1/3, x < 1/2\}$; *moreover* $\mu\left(\lim_{t\to\infty} X_t = x_1\right) > 0$ *and* $\mu\left(\lim_{t\to\infty} X_t = x_2\right) > 0$.

Proof The proof is analogous to that of lemma B.2, replacing theorem A.4 with theorem A.5. However to apply theorem A.5 we need an auxiliary function F which satisfies conditions (iii) and (iv). Consider functions of this type:

$$F(x,y) = \sum_i 1_{A_i}(x,y)\left[\int_0^x a_i(u)du + \int_0^y b_i(v)dv + C_i\right], \tag{56}$$

where $\{A_i\}$ is a partition of T_2, $1_A(.,.)$ is a simple function which takes value 1 on A and 0 otherwise, C_i are constants, $a_i(.)$, $b_i(.)$ are constant functions non-necessarily different from 0. With an appropriate choice of A_i, C_i, $a_i(.)$ and $b_i(.)$, F satisfies conditions (iii) and (iv) of theorem A.5.

Proof of Proposition 3.4 Notice that there is an open neighborhhood of x_1 such that $i_t^* = 1$ and there is an open neighborhood of x_2 such that $i_t^* = 2$. Then apply lemma B.4.

Proposition 4.1 *If* $-\frac{1}{2}\frac{E(\xi_{1t})-E(\psi_{0t})-E(\xi_{0t})}{E(\xi_{1t})+E(\psi_{0t})+E(\xi_{0t})} < a < \frac{1}{2}\frac{E(\xi_{1t})+E(\psi_{0t})-E(\xi_{0t})}{E(\xi_{1t})+E(\psi_{0t})+E(\xi_{0t})}$ *then*

$$\mu\left(\lim_{t\to\infty} i_t^* = 1\right) = 1.$$

Proof We need to show that $h(.)$ has no zeroes for $x < 1/2 + a$, then by remark A.1 the statement of the theorem follows; notice in fact that condition (i) of remark A.1 holds because

$$\mu\left(F_a(x)E(\psi_{1t}) < 1/2, \quad (1 - F_a(x))E(\psi_{0t}) < 1/2\right) \leq L < 1/2.$$

Recall that $h(.)$ is

$$h(x) = \begin{cases} E(\psi_{1t}) + E(\xi_{1t}) - xE(\psi_{1t} + \xi_{1t} + \xi_{0t}) & \text{if } x \geq 1/2 + a \\ \\ \begin{array}{c} F_a(x)E(\psi_{1t}) + E(\xi_{1t}) + \\ -xF_a(x)[E(\psi_{1t}) - E(\psi_{0t})] + \\ -xE(\xi_{1t} + \psi_{0t} + \xi_{0t}) \end{array} & \text{if } 1/2 - a < x < 1/2 + a \\ \\ E(\xi_{1t}) - xE(\xi_{1t} + \psi_{0t} + \xi_{0t}) & \text{otherwise} \end{cases}$$

where $F_a(x) = \frac{1}{2a}\left(x - \frac{1}{2} + a\right)$ for $1/2 - a < x < 1/2 + a$. Notice that $h(.)$ is continuous. Consider first $x \leq 1/2 - a$; in this interval $h(.)$ has no zeroes if $E(\xi_{1t}) - \left(\frac{1}{2} - a\right)E(\psi_{0t} +$

$\xi_{1t} + \xi_{0t}) > 0$, which leads to $a > -\frac{1}{2}\frac{E(\xi_{1t})-E(\psi_{0t})-E(\xi_{0t})}{E(\xi_{1t})+E(\psi_{0t})+E(\xi_{0t})}$. Now, consider $1/2 - a < x$ and take first $E(\psi_{0t}) = E(\psi_{1t})$; notice that $h(.)$ is linear for $1/2 - a < x < 1/2 + a$; if $E(\psi_{1t}) + E(\xi_{1t}) - \left(\frac{1}{2} + a\right) E(\psi_{1t} + \xi_{1t} + \xi_{0t}) > 0$, which leads to $a < \frac{1}{2}\frac{E(\xi_{1t})+E(\psi_{0t})-E(\xi_{0t})}{E(\xi_{1t})+E(\psi_{0t})+E(\xi_{0t})}$, then $h(1/2 + a) > 0$, therefore if

$$-\frac{1}{2}\frac{E(\xi_{1t}) - E(\psi_{0t}) - E(\xi_{0t})}{E(\xi_{1t}) + E(\psi_{0t}) + E(\xi_{0t})} < a < \frac{1}{2}\frac{E(\xi_{1t}) + E(\psi_{0t}) - E(\xi_{0t})}{E(\xi_{1t}) + E(\psi_{0t}) + E(\xi_{0t})} \tag{57}$$

the only possible zero of $h(.)$ is for $x < 1/2 + a$. Now notice that if $E(\psi_{0t}) < E(\psi_{1t})$ then $h(.)$ shifts upwards, therefore if (57) holds the only possible zero of $h(.)$ is again in $(1/2 + a, 1]$.

8 References

Acemoglu, D., and F.Zilibotti, Was Prometheus Unbound by Chance? Risk, Diversification and Growth, CEPR Discussion Paper No. 1426, 1996.

Aghion, P., and P.Howitt, A Model of Growth Through Creative Destruction, *Econometrica*, **60**, 323-51, 1992.

Antonelli, C., *The Economics of Localized Technological Change and Industrial Dynamics*, Dordrecht: Kluwer Ac., 1995.

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

Arthur, W.B., Competing Technologies, Increasing Returns and Lock-In by Historical Events, *Econ.J.*, **99**, 116-31 1989.

Arthur, W.B., On Designing Economic Agents That Behave Like Human Agents, *J. of Evolutionary Econ.*, **3**, 1-22, 1993.

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, 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.

Atkinson, A.B. and J.Stiglitz, A New View of Technological Change, *Econ. J.*, **79**, 573-78, 1969.

Azariadis, C., The Economics of Poverty Traps. Part One: Complete Markets, *J. Econ. Growth*, **1**, 449-86, 1996.

Bassanini, A., A Note on Brock and Mirman's Proof of the Ergodic Theorem in Stochastic Growth Models, CEIS Working Paper n.17, Università di Roma "Tor Vergata", 1996.

Bassanini, A., Modeling Path-Dependent Growth Through Technological Change: A Polya Urn Scheme Formulation, unpublished manuscript, 1997.

Bassanini, A. and G.Dosi, Competing Technologies, International Diffusion and the Rate of Convergence to a Stable Market Structure, IIASA Interim Report, forthcoming, 1997.

Becker, G.S., K.M.Murphy and R.Tamura, Human Capital, Fertility, and Economic Growth, *J. Pol. Econ.*, **98**, S12-37, 1990.

Benaim, M. and M.W.Hirsch, Learning Processes, Mixed Equilibria and Dynamical Systems Arising from Repeated Games, unpublished manuscript, 1996.

Benveniste, A., M.Métivier and P.Priouret, *Adaptive Algorithms and Stochastic Approximation*, Berlin: Springer, 1990.

Brezis, E.S., P.Krugman and D.Tsiddon, Leapfrogging in International Competition: A Theory of Cycles in National Technological Leadership, *Amer. Econ. Rev.*, **83**, 1211-19, 1993.

Brock, W.A. and L.J.Mirman, Optimal Economic Growth and Uncertainty: The Discounted Case, *J. Econ. Theory*, **4**, 479-513, 1972.

Brock, W.A. and S.N.Durlauf, A Formal Model of Theory Choice in Science, SSRI Working Paper n.9707, University of Wisconsin at Madison, 1997.

Caselli, F., Technological Revolutions, unpublished manuscript, 1996.

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

Cowan, R. and P.Gunby, Sprayed to Death: Path-Dependence, Lock-In and Pest Control Strategies, *Econ.J.,* **106**, 521-42, 1996.

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., *Technical Choice, Innovation and Economic Growth*, Cambridge: Cambridge University Press, 1975.

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

David, P. and D.Foray, Dynamics of Competitive Technology Diffusion through Local Network Structures: The Case of EDI Document Standards, in L.Leydesdorff and P.van den Besselaar (eds.), *Evolutionary Economics and Chaos Theory*, London: Pinter, 1994

Desmet, K., A Perfect Foresight Model of Regional Development and Skill Specialization, unpublished manuscript, 1997.

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

Dosi, G., Sources, Procedures and Microeconomic Effect of Innovation, *J. Econ. Lit.*, **26**, 1120-71, 1988.

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.

Dosi, G., K.Pavitt and L.Soete, *The Economics of Technological Change and International Trade*, New York: NYU Press, 1990.

Duflo, M., *Algorithmes Stochastiques*, Paris: Springer, 1996.

Durlauf, S.N., Nonergodic Economic Growth, *Rev. Econ. Studies*, **60**, 349-66, 1993.

Durlauf, S.N., Path-Dependence in Aggregate Output, *Industrial and Corporate Change*, **3**, 149-71, 1994.

Galor, O., and D.Tsiddon, Technological Progress, Mobility and Economic Growth, *Amer. Econ. Rev.*, **83**, 363-82, 1997.

Grossman, G. and E.Helpman, *Innovation and Growth in the Global Economy*, Cambridge, MA: The MIT Press, 1991.

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

Jones, C.I., R&D-Based Models of Economic Growth, *J. Pol. Econ.*, **103**, 759-84, 1995.

Joshi, S., Recursive Utility and Optimal Growth Under Uncertainty, *J. Math. Econ.*, **24**, 601-17, 1995.

Jovanovic, B. and Y.Nyarko, Learning by Doing and the Choice of Technology, *Econometrica*, **64**, 1299-310, 1996.

King, M. and M.Robson, A Dynamic Model of Investment and Endogenous Growth, *Scandinavian J. of Econ.*, **95**, 445-66, 1993.

Kirsch, D., The Accidental Early Demise of the Electric Car, unpublished manuscript, 1996.

Krugman, P., History vs. Expectations, *Quart. J. Econ.*, **106**, 651-67, 1991.

Levin, R., W.Cohen and D.Mowery, R&D Appropriability, Opportunity and Market Structure, *AEA Papers and Proceedings*, **75**, 20-4, 1985.

Liebowitz, S.J. and S.E. Margolis, Path-Dependence, Lock-In and History, *Journal of Law, Economics and Organization,* **11**, 205-26, 1995.

Majumdar, M., T.Mitra and Y.Nyarko, Dynamic Optimization under Uncertainty: Non-Convex Feasible Set, in G.R.Feiwel (eds.), *Joan Robinson and Modern Economic Theory*, New York: NYU Press, 1989.

Mirman, L.J., The Steady State Behaviour of a Class of One-Sector Models with Uncertain Technology, *J. Econ. Theory*, **6**, 219-42, 1973.

Mokyr, J., *The Lever of Riches: Technological Creativity and Economic Progress*, Oxford: OUP, 1990.

Nelson, R., Research on Productivity Growth and Productivity Differences: Dead Ends and New Departures, *J. of Econ. Lit.*, **19**, 1029-64, 1981.

Parente, S., Technology Adoption, Learning-by-Doing, and Economic Growth, *J. Econ. Theory*, **63**, 346-69, 1994.

Silverberg, G., G.Dosi and L.Orsenigo, Innovation, Diversity and Diffusion: A Self-Organization Model, *Econ. J.*, **98**, 1032-54, 1988.

Smulders, C., and Th.van de Klundert, Imperfect Competition, Concentration and Growth with Firm-Specific R&D, *Eur. Econ. Rev.*, **39**, 139-60, 1995.

Stokey, N.E., and R.E.Jr.Lucas (with E. Prescott), *Recursive Methods in Economic Dynamics*, Cambridge, MA.: Harvard University Press, 1989.

Stoneman, P. (ed.), *Handbook of the Economics of Innovation and Technological Change*, Oxford: Blackwell, 1995.

Tell, F., Innovation, Abandonment and Technological Progress: Path-Dependence and Knowledge Management in the High Voltage Power Transmission Industry, unpublished manuscript, 1997.

West, L.J., Keyboard Efficiencies Revisited: Direct Measures of Speed on the Dvorak and Standard Keyboards, *The Delta Pi Epsilon Journal*, **36**, 49-61, 1994.