REVISTA DE ECONOMÍA DEL ROSARIO, 4 (JUNIO 2001) 17-40

HISTORY DEPENDENCE IN THE ECONOMY. A REVIEW OF THE LITERATURE

Alejandro Gaviria *
Associated Researcher
Fedesarrollo
agaviria@fedesarrollo.org.co

Why did everything have to turn out the way it did?
Truman Capote

ABSTRACT

This paper presents a review on the concept of history dependence and time irreversibility in the economic literature. It presents an inventory of artificial examples where random events affect final outcomes and history cannot be reduced to the deliverer of the inevitable. It focuses mainly on recent examples and applications that, although not necessarily better than their predecessors, allow a better understanding of the mechanisms whereby history leaves its imprint. The paper reviews deterministic and stochastic models, presents some empirical illustrations and draws a few general conclusions.

Keywords: History dependence. Economic modelling.
JEL Classification: C59, N01

I. INTRODUCTION

Why did everything have to turn out the way it did? Every time one is confronted with this question a well-known dichotomy comes to mind: is everything inevitable? Or is it merely random? That is, is history merely playing out a well-defined script written by the "iron" laws of nature? Or is history driven by small chance events that are inherently unpredictable? In this paper, I attempt to shed some light on these seminal questions by reviewing the growing economic literature on history dependence and time irreversibility.

* The Central Bank of Colombia and the Alfred P. Sloan Foundation provided generous financial support. Vince Crawford, Ross Star and Cristiano Antonelli provided helpful comments and suggestions.

This paper is not exercise of advocacy. My ultimate goal is not to convince the unconvinced and reassure the skeptical about the importance of history in human affairs in general and economic phenomena in particular. Rather, my goal is to present an inventory of artificial examples where random events affect final outcomes and history cannot be reduced to the deliverer of the inevitable. I shall not attempt here to judge the plausibility of these examples; that is the reader's job. I shall present, however, some real-world illustrations that suggest, at the very least, that the ideas explored in this paper have some bearing with reality.

My review is not comprehensive. I focus mainly on recent examples and applications that, although not necessarily better than their predecessors, allow a better understanding of the mechanisms whereby history leaves its imprint. Although many past economists did recognize the importance of contingency and history, they did not possess the techniques necessary to generalize their theories and test them. Only recently, these techniques have become available, and so only recently economists (and other social scientists) have been able to study the workings of history using full-fledged models. Here I review these models, with a view not in their technical aspects but in their basic intuition. Readers interested in the details of the models can always consult the original sources. As with any review, breadth rather than depth is the guiding principal of this one.

The organization of this review is as follows. Section two focuses on deterministic models, section three focuses on stochastic models, section four presents some empirical illustrations and section five draws a few general conclusions.

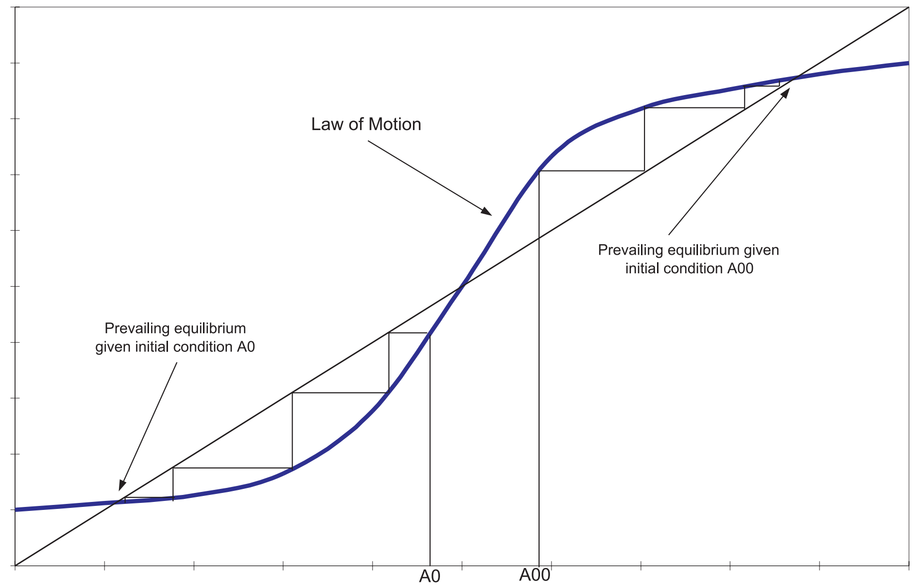
II. DETERMINISTIC MODELS WITH POSITIVE FEEDBACKS

In this section, we study a type of economic models that are often used to study discrete choices in the presence of externalities. These model share two basic characterictitics: first, they exhibit positive feedbacks in that the higher the participation in the activity in question (e.g., voting for a candidate, following a convention, committing a crime, migrating to a city), the higher the incentives for others to join in. And second, they explicitly specify a *deterministic* law of motion showing how past participation relates to current participation.

These models are often referred to as *critical mass* models as a way to describe what is perhaps their most notable property: the fact that some activities become self-sustaining once they surpass a certain threshold. The main message of these models is summarized in Figure 1. The story is well-known: there are multiple stable equilibria and the starting point —which is exogenous to the analysis— determines which equilibrium prevails. Of course, the bold curve doesn't have to be nice and regular, and it may cross the diagonal more than three times; all depending on the peculiarities of each application. These details notwithstanding, the main message follows through: history —in the form of the initial conditions— plays an important role in determining the final outcome.

FIGURE 1

Dependence on inicial conditions deterministic daw of motion



A. Threshold models with heterogeneous agents

Some *critical-mass* models assume that people are heterogeneous in that they have distinct participation thresholds—the minimum fraction of participants that compels them to participate in the activity under analysis. Technically, these models postulate a non-degenerate distribution of thresholds across individuals. Thomas Schelling (1978, p. 96) has made this point vividly, "you may dress formally if enough people do to keep you from being conspicuous, but I may dress formally only if so many do it that I would be conspicuous not to do it. You may be willing to enroll in a school in which the opposite sex outnumbers you no more than 3 or 4 to 1, but I may be unwilling to enroll in a school unless it is largely my own sex."

Threshold models with heterogeneous agents have been widely used by economists and sociologists. Schelling (1978) opened the door by suggesting many applications ranging from the dynamics of racial compositions of schools and neighborhoods to the evolution of the divorce rate. Many new applications have been proposed since then. A first notable example is Kuran's (1995) analysis of public opinion. In Kuran's model, one's expressive utility (the need to voice one's convictions) and one's reputational utility (the need to conform to the majority) determines one's political threshold (the lowest participation that impels one's support to the issue in question). If political thresholds vary across individuals, the situation depicted by Figure 1 applies directly and the evolution of public opinion will be greatly influenced by initial conditions. Andvig and Moene (1990) study the evolution of corruption using a similar framework. In their model, government officials are assumed to be heterogeneous with respect to the costs of supplying corrupt services, which implies a distribution of corruption thresholds—the minimal fraction of corrupted officials that prompts one's corruption. Here, as before, the profitability of corruption depends upon its established frequency and history does play a role by selecting the final outcome among the stable equilibria. Carrington et al. (1996) study a model of migration with moving costs that uses the same framework. In this model, moving costs decrease with the number of migrants already settled in the final destination and migrants are heterogeneous with respect to this costs. Finally, Granovetter (1978) employs a similar setting to analyze various situations involving social customs and conventions.

B. Threshold models with homogenous agents

The previous examples are just a small subset of a wider class of models involving positive feedbacks and deterministic adjustments—the theme of this section. Other examples that don't postulate threshold distributions but that still adhere to the characterization given earlier are reviewed in what follows. One first example is Krugman's (1991) analysis of industry location. Here, one industry's workers are other industries' customers, which causes the emergence of "agglomeration" economies. The key point here is that given sufficiently strong economies of scale and sufficiently low transportation costs, multiple equilibria arise and initial advantages end up dictating the pattern of location. Learning dynamics as applied to coordination games are also good examples of the class of models examined in this section. Coordination games provide in turn an excellent framework to study the emergence of conventions, "rules that have never been consciously designed and that it is in everyone's interest to keep" (Sudgen 1986, p. 54). Three notable examples in this respect are Sudgen's (1986) analysis of conventions, Chu's (1993) model of traffic order evolution, and Aiyagari's (1988) model of coordination failure in the macroeconomy.

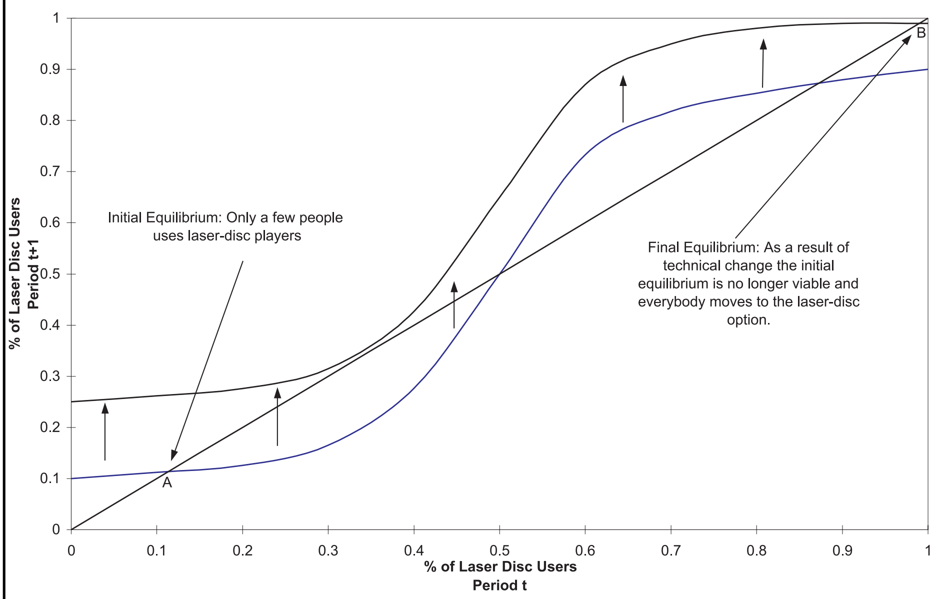
Multiple equilibria and history dependence are common properties of all the models presented in this section. Potential inefficiencies are also a property of these models: nothing in them guarantees that the best equilibrium —provided that there is one— will be selected. In fact, as a result of a bad start, things may go awry. Or to use the previous examples, people may end up following a costly convention or companies may end up locating in the wrong place. Moreover, once equilibrium is reached, it is difficult to exit from. This excess of inertia is often referred to as *lock-in*.

C. Exit from lock-in

A common topic in this literature deals with the issue of exit from lock-in (or how to overcome the legacy of history). The problem is as follows: each person (or firm) would benefit from switching to a better "option" as long as others follow suit, but individually none dare change in case others do not follow. Three different mechanisms of exit from lock-in in the absence of a coordinating authority have been examined in the literature: change in fundamentals, self-fulfilling prophecies and random mutations. A quick overview is presented below.

(i) Changes in fundamentals. Here changes in technology (or preferences) lead to changes in behavior that usually involve sudden shifts. This is illustrated in Figure 2: a stable equilibrium, point A, becomes unsustainable as a result of an exogenous shift in the *critical-mass* curve, which in turn prompts a large-scale movement toward B. An example, based on Krugman (1996 p. 56-60), may help clarify this point. There are two video-playing technologies: ones uses tapes and the other laser discs. At present time, people prefer the tape technology because stores stock a greater variety of cassettes. Likewise, stores carry mostly cassettes because most people have VCRs to play them and only a few have DVD players. Thus, the tape technology, arguably the inferior one, is right now the prevalent equilibrium. Yet the laser technology is getting better, which gradually increases the advantage of laser discs and causes the critical-mass curve to shift. As shown in Figure 2, eventually, everybody will adopt the laser technology and the *lock-in* to the tape technology will come to an end.

FIGURE 2
Exit from lock-in and technical change



(ii) Self-fulfilling prophecies. The argument here is as follows: if most people expect everybody to change, most people will change and their expectations will prove right. So if most individuals come to believe that everybody is for a change, there will be a mass exit from the current equilibrium without any change in fundamentals. Under what circumstances may expectations override the legacy of history? According to Krugman (1991), the role of expectations is greater, the higher people weight future income streams, the higher the extent of the external economies involved, and the less costly it is to make the change (e.g., to relocate an industry is arguably more costly than to change a convention).

(iii) Random mutations. The standard assumption here is that individuals experience independent random *mutations,* which is to say that every once in a while they switch to alternative activities without apparent reason (maybe just to experiment or maybe due to involuntary errors). Now, if enough individuals mutate simultaneously, the current equilibrium may be subverted. Two qualifications are in order: first, this argument applies only to small populations because when populations are large the expected time to upset the current equilibrium is not reasonable in any economic context. And second, lock-in is no longer an issue when *mutations* are present because the system will spend some time in each equilibrium, and the longest time at the equilibrium with the widest *basin of attraction* (Kandori, Mailath and Rob, 1993 and Kimur, 1993).

All three mechanisms of exit from lock-in described above stress the same point: if the transition from one equilibrium to another ever gets started, it will happen quickly. This is sometimes denoted as punctuated equilibrium—an expression first used by evolutionary biologists to describe a pattern of evolution in which long periods of quiescence are punctuated by brief interludes of dramatic change. So although history casts a long shadow when positive feedbacks are present, revolutions — in the usual sense of sudden subversions of the status quo— shouldn't be ruled out.

III. STOCHASTIC MODELS OF ALLOCATION WITH POSITIVE FEEDBACKS

As in the previous section, positive feedbacks are at the heart of the models considered here. Unlike the previous section, the laws of motion studied here contain stochastic elements. This not only introduces some mathematical complications, but also gives us a better look at the workings of history. In this section, we study seven different types of processes that involve both positive feedbacks and stochastic dynamics.

A. Allocation processes

Allocation processes are thoroughly described by Arthur (1994). A quick overview and an illustrative example are presented below. Here an allocation is made each period to one of k categories according to some vector of probabilities $\mathbf{p}$. The process start at time 1 with an initial vector of allocations, $\mathbf{y}_1 = (y^1, y^2, \ldots, y^k)$, where and y^i represents the number of initial units in category i. This in turn defines the initial vector of proportions, $\mathbf{x}_1 = (y^1/w, y^2/w, \ldots, y^k/w)$, where $w = \sum y^i$. The key assumption here is that proportions ($\mathbf{x}$) get translated into probabilities ($\mathbf{p}$) by means of a monotonically increasing function $p: S^k \rightarrow S^k$. That is, $\mathbf{p} = (p^1(\mathbf{x}), p^2(\mathbf{x}), \ldots, p^k(\mathbf{x}))$[1]. This implies that the new "unit" arriving at period 2 will go to category i with probability $p^i(\mathbf{x}_1)$, the one arriving at period 3 will go to the same category with probability $p^i(\mathbf{x}_2)$, and so on for the next units.

Three remarks are in order. First, higher proportions imply higher probabilities, and thereby the presence of positive feedback. Second, the function **p** depends on the distribution of preferences and possibilities of the problem in question. And third, allocation processes are clearly stochastic in that the "state" of the system (**x**) doesn't determine where the new "unit" goes, but the probabilities over all possible allocations.

When dealing with allocation process, we are interested in studying how the proportions in each category build up. It's not difficult to show (see Arthur 1994, p.54) that the expected motion of the proportions obeys the following equation:

$$E(\mathbf{x}_{t+1}/\mathbf{x}_t) - \mathbf{x}_t = \frac{1}{t+w}(p(\mathbf{x}_t) - \mathbf{x}_t) \quad (1)$$

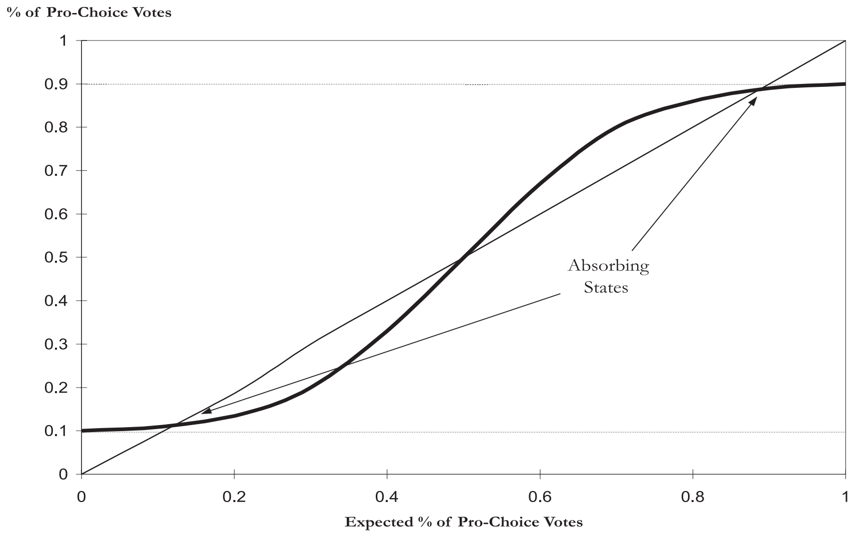
where $\mathbf{x}_t$ is the vector of proportions at time t. Clearly, future proportions will increase —at least on an expected basis— as long as probabilities exceed current proportions. Moreover, depending upon the function p, there may be several proportions at which deterministic motion is zero: p(**x**) = **x**. Some of these *fixed-points* will be stable (attractors) and some will be unstable (repellers), but the latter are of little importance. Indeed, Arthur et al. (1983) show that any allocation system converges, with probability one, to one of the stable fixed points of p. Which attractor gets selected depends on small historical accidents: who arrives first, how people (or firms) bunch together in the "arrival line", who decides to wait and see and who decides to enter right away, and so on. That is why allocation processes are said to be path-dependent: early history determines the final outcome in a way that is unpredictable at the outset.[2]

The following example may help get across some of the points raised earlier. Let's assume that we are in the middle of the convention of the US Republican Party, where one of the main events is taking place. The delegates are voting to decide whether to include an anti-abortion declaration in the party platform. The referendum is organized in a sequential manner: delegates arrive one at time, observe the current distribution of votes, and decide whether to go pro-life or pro-choice. The preferences —represented here by the distribution of thresholds— are summarized in Figure 3. As shown, 10 percent of the voters will vote pro-choice regardless of the expected outcome, a similar percentage will cast a pro-life ballot no matter what, but most people will base their votes upon the expected outcome. The voters here are myopic

1 Models having this property are often referred to as generalized urn process. See Dosi and Kaniovski (1994) for a comprehensive discussion of these processes, including several generalizations.

2 In contrast to the models studied in the previous section, the final outcome in allocation processes isn't entirely dictated by initial conditions: "fluctuations dominate motions at the outset; hence; they make limit points reachable from any initial conditions" (Arthur 1987, p. 20).

FIGURE 3
Distribution of thresholds abortion referendum

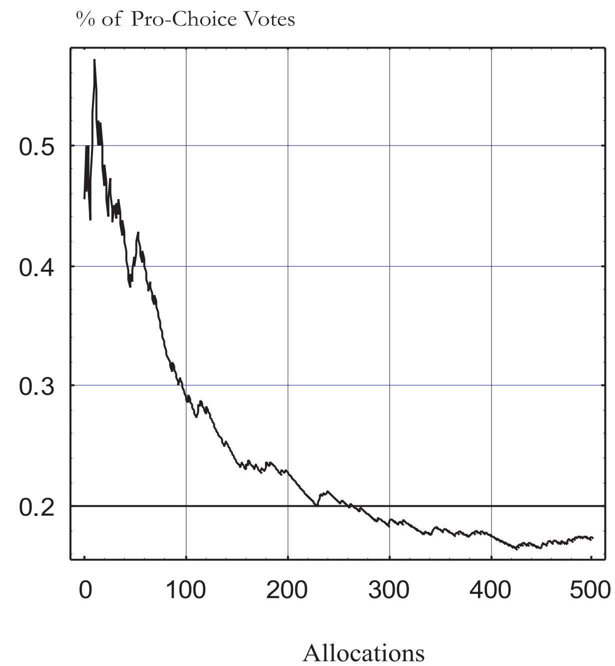


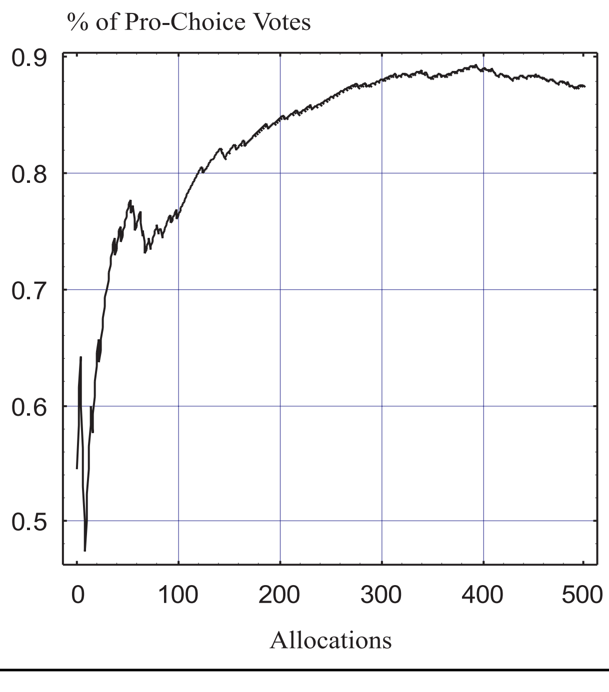
in that they all regard the outcome they observe when casting their votes as the best possible forecast of the final outcome.

What is the final outcome of this hypothetical example? We don't know. An answer would require an infinitely detailed knowledge of the sequence of voting that, arguably, we don't have. Indeed, the only thing we can venture to say a priori is that the process will end up converging to one of the two attractors. Figure 4 shows two different runs of a computer simulation of the abortion referendum: whereas in the top-graph the pro-life ticket gets 90% of the votes, in the bottom graph the pro-choice ticket gets a similar fraction. Interestingly, despite our perfect knowledge of both the preferences and the voting procedure, any prediction will be wrong with probability one-half. In sum, history decides then —in an unpredictable and capricious manner— whether or not the an anti-abortion declaration will become part of the Republican party platform.

Allocation processes have been used to study a wealth of different phenomenon. Schelling (1978) provides some vivid examples ranging from people crossing a congested intersection to the seating pattern of auditoriums. Arthur (1994, chap. 5) presents an interesting application of allocation process to the transmission of information based on experience. In Arthur's model, risk averse buyers sample past purchasers in order to gather information about competing products. This introduces an *information feedback* in the process: buyers are likely to learn more about a commonly purchased product than one with few previous users. As a result, products that by chance win market share early on gain an informational advantage, which in turn tend to increase their market share. In the same vein, Arthur (1994, chap. 8) and Roth

FIGURE 4
Two different histories of the voting referendum





and Erev (1995) model individual human learning by using allocation processes. Here earlier successes reinforce past actions, which prevents further experimentation and thus precludes the adoption of potentially better actions. Finally, Arthur (1994, chap. 4) and Krugman (1996, p.39-46) have used allocation processes to study the emergence of urban conglomerates.

B. Market adoption processes with learning by using

This type of processes was first introduced by Arthur (1989) in order to study competition between unsponsored technologies (steam and gasoline engines at the turn of the century, and light and heavy water nuclear reactors during World War II are common examples). The crucial point here is the presence of increasing returns by adoption stemming from learning by using: "as adoption accumulates, the usage and experience that accumulate with it become incorporated into more reliable and effective variants" (Arthur 1989). It's also usually assumed here that adopters have different natural *tastes* for the available technological options, and that the variant each adopter choose is frozen in design at his time of choice, so that his payoff is not affected by future adoptions.

Market Adoption processes are different from Allocation processes in two respects. First, randomness is no longer determined by the probability of new allocations, but it stems directly from assuming that the order of arrival at the "adoption window" is unknown.[3] Second, the long-run behavior of the system depends now on whether the improvement caused by successive adoptions grows without bound. If so, monopoly by one technology must eventually occur. Otherwise, "market sharing" is a distinct possibility (Arthur 1989).

The following example –adopted from Arthur (1989)— may both clarify some points and fill in some details left unexplained in the previous exposition. Let's assume that there are two different technologies competing for the satellite launching market. Let's call the first one Apollo and the second one Atlantis. Further, let's assume that there are two types of adopters (or purchasers of satellite launching services). Let's call the first ones Americans who have a natural preference for the Apollo technology, and let's call the second ones Europeans who initially prefer the Atlantis technology. The preferences are shown in Table 1. Obviously, the preferences change with each adoption as a result of the learning-by-using effect mentioned earlier. Figure 5 shows two possible histories of the satellite launching business.[4] As shown, chance circumstances behind the order of arrival of Americans and Europeans at the *adoption window* determine which technology comes to dominate the market. Again, path-dependence, unpredictability and lock-in by historical accident are notable characteristics. Again, nothing guarantees that the best technology is chosen. Again, history may prove Dr. Pangloss wrong.

---

3 Technically, we are no longer dealing with generalized urn (or Polya) processes but with random walks with absorbing barriers.

4 In this example, it was assumed that the two type adopters are equally likely to arrive at the *adoption window* at all moments.

FIGURE 5
Market adoption process satellital technologies

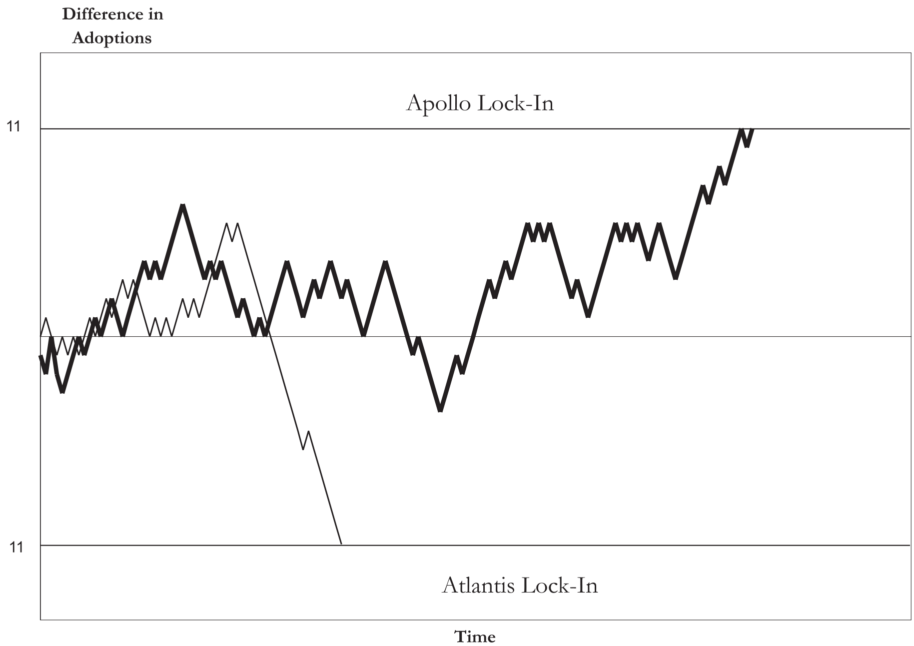


TABLE 1
Technological preferences in satellital launching

	Apollo	Atlantis
Americans	$12+N_{ap}$	$1+N_{at}$
Europeans	$1+N_{ap}$	$12+N_{at}$

N_{ap} : Number of Apollo adopters.
N_{at} : Number of Atlantis adopters.

Finally, a word about the role of expectations in market adoption models is necessary. The first thing to remember is that we assumed earlier that the payoffs aren't affected by future adoptions. Obviously, this assumption makes any consideration of forward-looking behavior irrelevant. What will happen if we allow future payoffs to be affected by later adoptions? Or, in other words, what will happen if we try to use *market adoption* processes to study, say, the formation of urban systems where what matters is not the payoff at the moment of the *adoption* but rather the payoff after all the dust settles down? Katz and Shapiro (1983, 85), and Arthur (1989) have tackled this question. Two main conclusions emerge from their analyses: first, expectations may result in faster lock-in, and second, self-fulfilling prophecies are a latent possibility.

C. Informational cascades

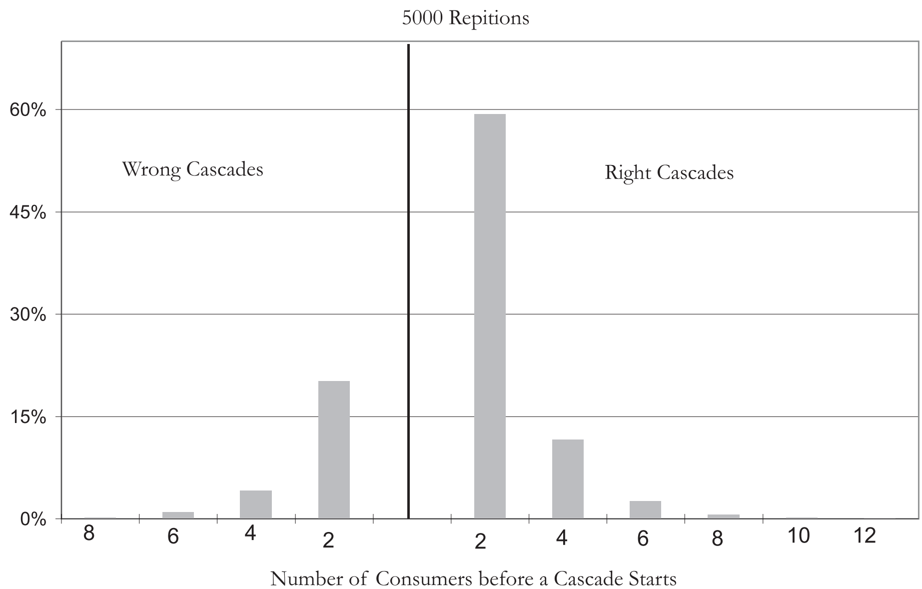
Another type of popular models of sequential choice under uncertainty focus on the transmission of information when communication is not possible. These models usually go under the heading of informational cascades. As in the previous models, people arrive sequentially at the adoption window; unlike them, the order of arrival is now known to all. In a typical example, each individual has to decide whether to adopt or reject certain entity; a new technology, a new brand, a new form of behavior. The gain to adopting is unknown but fixed in that it doesn't depend on the number of adopters. So, network externalities are ruled out by assumption. What is key here is that each individual first privately observes a signal and then makes a decision of whether or not to adopt based on two types of information: his own signal and the decisions made by those ahead of him.

The decision of the first person is trivial: he only adopts if he gets a *high* enough signal. The second person faces a tougher problem because she has to take into consideration not only her own signal, but also any inference she can make on the basis of the decision of her predecessor. Latecomers face a similar problem with the difference that the later they arrive, the more information they have. The main feature of these models is that, sooner or later, it will be optimal for somebody (and therefore for all individuals who follow him) to ignore his own signal and just mimic the decision of his predecessor. An informational cascade will then get started. A cascade can be either positive —wherein all individuals adopt— or negative —wherein all individuals reject. Further, and more importantly, an accumulation of historical accidents —a sequence of wrong signals due to mere random variation— may trigger a wrong cascade, namely, people may end up rejecting a good choice or accepting a bad one.

Consider the following binary signal example in the spirit of Bikhchandani et al. (1992). Let's assume that a group of consumers are trying to decide whether to adopt the latest version of computer software. They make their decisions in a sequential manner (e.g., they all coincide in the same computer shop). The price of the new product is set to $100. The intrinsic value of the product is the same for all individuals and is either $200 or $0, each with a prior probability ½. All consumers have had the opportunity to observe a binary signal regarding the true value of the software (e.g., they all had access to a trial version of the program). The signal probabilities as follows the probability of getting a high signal is 0.7 when the true value is $200, and 0.3 when the true value is $0.

Figure 6 shows the result of a simulation of the previous example. Five thousand repetitions were implemented, all of them under the assumption that the true value of the new software equals $200. Two things are noteworthy. First, informational cascades start rather soon: it never took more than 12 individuals to start one, and 80 percent of all cascades started right after the first two individuals. Second, wrong cascades —a rejection of the new program in this context— occurred in 25 percent of the cases. Once again, historical accidents may drive the process away from efficiency. Once again, history matters and predictability is all but possible.

FIGURE 6
How soon does a cascade start?



Informational cascades differ from both allocation and *market adoption* process in several ways. First, the sources of path dependence are different. Unlike the previous models, path dependence doesn't stem here from the presence of network externalities or any other form of coordination effects. Rather, it arises as a result of the way uncertainty is resolved when decisions are made in a sequential manner. Second, lock-in is no longer a property of these models. On the contrary, informational cascades tend to be rather fragile; namely, the arrival of a little information or the mere possibility of a value change can subvert a cascade.

Informational cascades have been proposed as a possible explanation for many instances of imitative behavior. Bikhchandani et al., for example, have argued that fads among doctors (they all adopt the same surgical procedures), investors (they all pick the same *winners*), and animals (they all settle in the same territory), can be regarded as informational cascades. However, these applications —and any others that can be proposed— are speculative in that it is usually not possible to establish whether the observed conformity is due to informational factors rather to some sort of network externalities.

D. Bandit problems

The so-called bandit problems have a long tradition in probability theory. A typical example goes like this. There is a slot machine with two or more arms. The jackpot probabilities of each of the arms are unknown but independent. Each play gives either a success (we hit the jackpot) or a failure (we get zero). The goal is to play the arms —one at a time— so as to maximize the expected discounted winnings. An optimal strategy must find the right balance between two conflicting approaches: exploitation (profit now by pulling the arm with the current higher subjective probability) and exploration (pulling a *bad* arm in order to learn more about its *true* probability). Among the applications of Bandit problems, we may mention clinical trials, job search, and technology adoption under uncertainty and employment hazards (See Berry and Fristedt 1985 for a comprehensive survey).

Although easy to state, bandit problems are deceptively difficult to solve. More importantly, lock-in by historical accident and path-dependence arises naturally in these problems. This is so regardless of whether an optimal strategy is implemented. Consider the following common example: as managers of a telecommunications company, we are trying to choose between two technologies of unknown merit. Let's call them digital and analog. The digital technology is better in reality, but we don't know it.[5] Indeed, prior to the analysis, we believe that both technologies are *worse* than they really are, and that they both are *equally* attractive. Under these circumstances, we should try out both technologies in the early rounds. Of course, as they are increasingly used, we will learn more about their true merits. However, it is possible that —out of sheer bad luck— the digital technology delivers poor results in the first tryouts, which in turn will lower our beliefs about its profitability. Thus, we erroneously will come to believe that the analog technology is better, and eventually we will stick with it forever. After repeated use, we will know how good (or bad) the analog technology is and, mistakenly, we will be convinced that it is much better that its *unlucky* rival. As pointed by Cowan (1991. P. 807), "because the superior technology is not being used, our beliefs about it do not change— it has no way to demonstrate its superiority, and we are locked-in to" the inferior technology. [6]

While the previous models can be thought of as examples of unregulated markets involving sequential choices, bandit problems can be thought of as archetypal of the problem faced by a social planner trying to intervene in the same markets. As shown, inefficiencies can occur in any case. In sum, neither the market nor the smartest of the central planners are free from backing the wrong horses.

---

5 When we say that the digital technology is better, we mean that its distribution of returns first-order-stochastically dominates that of the analog technology. Of course, this doesn't rule out that the analog technology will sometimes perform better.

6 As is well-known in the bandit literature, "the optimal policy will, with positive probability, lead to the inferior technology being adopted infinitely often and the superior technology being adopted only a finite number of times" Cowan (1991).

E. Spatial models

Spatial models are models of imitative behavior in which individuals change their behavior so as to conform to individuals in their vicinity. These models are closely related to the so-called Ising models in physics (Gaylord and Wellin 1995, chap. 6) and to voting models in probability theory (Ligget 1985).

Consider the following example in the spirit of David (1992). A group of students are taking a decisive final exam. The students are arranged in a circle, and they all face a common dilemma: to cheat or not to cheat. All students solve this dilemma in the same manner: they turn their heads either to the right or to the left, observe whether or not a neighbor is cheating, and then mimic his behavior.[7] Interestingly, several predictions about aggregate behavior can be made in this case. First, there are two absorbing states and both exhibit full conformity (i.e., eventually everybody will cheat or nobody will). Second, although it is impossible to predict at the outset which absorbing state will prevail, we know that the cheating equilibrium will occur with a probability equal to the proportion of *cheaters* at the beginning of the exam (David 1992, p. 216).

Spatial models have been used by Puffert (1986) to study historical competition between railroad gauges, by Glaeser, Sacerdote and Scheinkman (1996) to study social interactions and crime, and by Ellison (1993) to study the evolution of coordination when individuals interact only with their neighbors. Of particular interest is Ellison's finding that if individuals' choices are subject to random perturbations and players interact with small sets of neighbors, chance events become less important and evolutionary forces acquire preeminence.[8]

F. Genetic algorithms

Although genetic algorithms were initially proposed as a numerical method to solve complex optimization problems, they can also be understood as descriptive models of the evolution of human designs. Genetic algorithms, viewed as positive models, emphasize the importance of precedence in the evolution of human designs in general and technology in particular. This point has long been recognized in evolutionary biology. Gould (1982, p.383) puts it best: "The constraints of inherited form and developmental pathways may so channel any change that even though selection induces motion down permitted paths, the channel itself represents the primary determinant of evolutionary direction".[9] In other words, once a design has taken a particular path, it tends to stay in it, and so many of the end results will have no other explanation than the shadow of their past.

---

7 Formally, this problem corresponds to a class of problems usually referred to as of interacting Markov processes (see Kindermann and Snell, 1980).

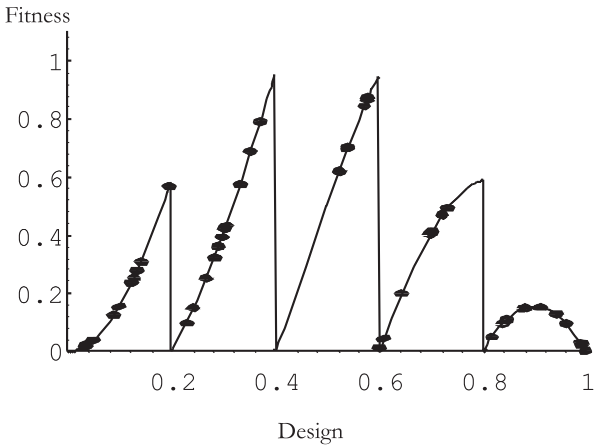
8 Young (1996) reaches a similar conclusion in his analysis of the evolution of the rules of the road.

9 Gould (1989, p.301) offer some examples that may help to get this point across. "Pandas, to eat bamboo, must build an imperfect thumb from a nubbin of a wrist bone, because carnivorous ancestors lost the requisite mobility of their first digit. Likewise, many animals of the Galapagos differ only slightly from neighbors in Ecuador, though the climate of this relative cool volcanic islands diverges profoundly from conditions on the adjacent South America mainland".

Consider the following example. Let's assume first that at a given point in time there are 50 different designs for a given artifact, say a motorcycle. Of course, some designs are better than others, which in this context means that they have a higher *fitness*. Figure 7A shows a hypothetical fitness landscape for motorcycle designs along with the 50 initial designs represented by dots. A genetic algorithm can then be used to model the evolution of the designs.

FIGURE 7
Genetic algorithms and history dependence

A. Initial Designs and Fitness Landscape



A typical algorithm consists of three parts. (i) Selection: each string is selected with a probability proportional to its fitness value. (ii) Crossover: in a pair of selected designs (strings), a position along the string is chosen, and the right and left part of each string are swapped.[10] And (iii) mutation: each gene (a particular digit in the string) is changed at random with a small probability.

Figures 7B and 7C show the resulting designs after 500 generations for two different histories of the world.[11] In both cases average fitness increases, but evolution may take different paths. As shown, the final designs cluster either to the left of 0.4 or to the left of 0.6. The story is by now a familiar one: historical accidents move the design toward an initial path, which greatly constraints future evolution and determines the general direction of the process. Although highly stylized, this example captures important features of the evolution of motorcycle design as told by Petroski (1994,

[10] Prior to the analysis we rewrite each design, a real number between zero and one, as a binary string of length ten.

[11] The analysis was done using a mathematics program adapted from Bengtsson (1994).

p. 175), "what might ultimately come down to an arbitrary choice among competing configurations, as manifested in the location of the fuel tank, for example, in time can become so strongly associated with motorcycles that, even if functionally relocated in a new (and improved) design, a vestigial tank (a *survival form*) may be retained in what has become the customary location".

FIGURE 7

B. Designs after 500 generations one history

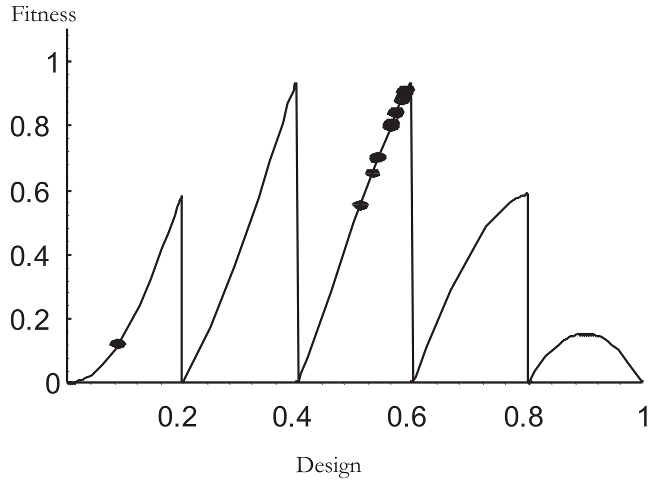
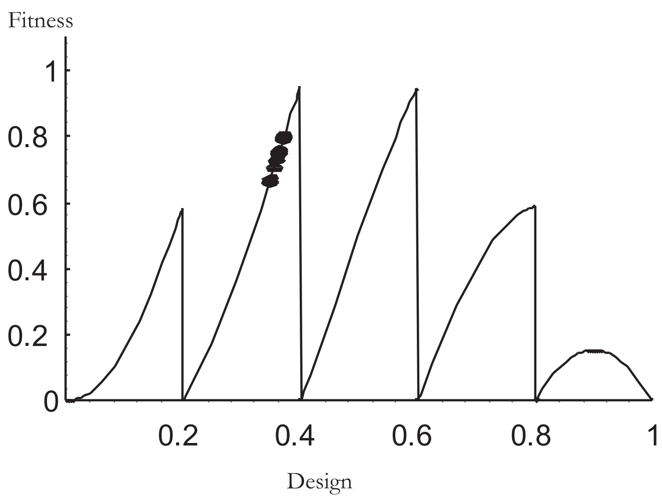


FIGURE 7

C. Designs after 500 generations a different history



The analogies of technological and biological evolution has been stressed by Basalla (1988) and Kauffman (1995), among others. In Kauffman's words, "it seems worthwhile to consider seriously the possibility that the patterns of branching radiation in biological and technological evolution are governed by the same laws". However, we are yet to define what is meant by *fitness* in technological evolution. One thing is clear, though: contingency is a powerful force behind the designs of both nature and man.

G. Switching models

These models focus on the dynamics of *markets* with the following features. There is a fixed number of *buyers* that must choose one among K available categories (e.g., brands of durable goods or candidates in an election). *Buyers* are fickle and often switch from one category to another. And buyers switch with higher probability toward those categories with a higher *market share*.

Consider the following example. There is an election involving two candidates (A and B) and a finite number of voters, some of which support candidate A and some of which support candidate B. Voters tend to switch toward popular candidates in that the probability that a supporter of A becomes a supporter of B is higher, the higher the current support for candidate B. Weidlich and Haag (1983) show that, under these conditions, the distribution of public opinion will converge to a bi-modal distribution, meaning that one candidate will eventually garner all popular support in a way that is unpredictable at the outset.

Smallwood and Conlisk (1979) use a similar framework to study the behavior of a market in which fickle customers switch toward more popular brands. In this model, customers can costlessly dump their current brand of a durable good and choose among a fixed number of alternatives, which they do with some fixed probability. Costumers tend to switch toward more popular brands with some propensity that can be manipulated by the modeler. Smallwood and Conslik show that the higher this propensity, the higher the probability that the market will be dominated by an inferior brand.

IV. EMPIRICAL EVIDENCE

The models presented in the previous sections make it clear that historical dependence and lock-in by historical accident can occur in artificial worlds built upon some convenient assumptions. But do they occur in reality? The answer to this question seems to be yes. The most famous illustration is the keyboard design of most typewriters. As is well-known, the top row of alphabet keys of the standard typewriter reads QWERTY. Why? David (1985) spells out the details of what has become a popular story. The QWERTY arrangement was introduced as solution to a problem in the early days of the typewriter: "the tendency of the typebars to clash and jam if struck in rapid succession". The idea was to minimize the collision problem by separating those letters that usually come together in the English language. Although this problem lost relevance long ago, the QWERTY layout is still the standard and will likely be the standard for the years to come despite the presence of what appear to be more efficient arrangements. The reason is simple. Most of us are QWERTY-programmed typists unwilling to bear the cost of a change. As a result, manufacturers of computers and typewriters will continue producing QWERTY keyboards, which will in turn cause the new generations to become QWERTY-programmed typists, and so on *ad infinitum*.

Although the QWERTY tale has become a *cliché*, it is not without its detractors. Among the more forceful are Liebowitz and Margolis (1990), who claim that the basic elements of the story are wrong. They claim, first, that the QWERTY layout was not purposively designed to slow down typist and, second, that this layout is not clearly inferior to some well-publicized alternatives. In general, Liebowitz and Margolis (1994) strongly suggest that compelling examples of lock-in by historical accident to inferior standards are yet to be produced. One can argue, however, that the FORTAN programming language in the 1950s, the DOS computer operating system in the 1980s, and the TCP-IP protocols that rule the internet today are clear examples of inferior standards that became dominant as a result of network externalities.

Cases of lock-in by historical accident to inferior technologies have also been documented. These stories rely on the following counterfactual: if technology B (the loser) had been adopted instead of A (the winner), that is, if the research and development invested in A had been instead invested in B, the hypothetical current form of technology B would be superior to A. Although counterfactuals can always be regarded as mere speculations, some plausible stories have been told. Cowan (1990) has argued that a series of trivial circumstances locked virtually the entire US nuclear industry into light-water technologies, leaving behind the high-temperature gas-cooled reactors that, arguably, would have been better.[12] In the same vein, Arthur (1984) relates the following story of how gasoline engines becomes the standard way to power cars:

> *In 1890 there were three ways to power automobiles – steam, gasoline, and electricity—and of these one was patently inferior to the other two: gasoline. …. [A turning point for gasoline was] an 1895 horseless carriage competition sponsored by the* Chicago Times-Herald. *This was won by a gasoline powered Duryea – only one of two cars to finish out of six starters… Steam continued viable as an automotive power source until 1914, when there was an outbreak of hoof-and-mouth disease in North America. This led to the withdrawal of horse troughs – which is where steam cars could fill with water.*

What would have happened had the steam technology been adopted instead? We don't know for sure, but some engineers strongly believed that steam was the better bet (Balassa, 1970). Finally, Cowan and Gunby (1996) argue that positive feedbacks in technology adoption can explain why chemical control of agricultural pests remains

[12] Basalla (1988, p. 166) argues that "light water reactors are one of the least efficient consumers of uranium fuel." Its dominance seems to have been prompted by the successful implementation of light water heat exchangers in the first nuclear submarine, the USS nautilus.

the dominant technology in spite of the presence of what seems a superior alternative, integrated pest management (IPM).[13]

Location stories are also good illustrations of the power of historical accident. Krugman (1991), for instance, relates how Dalton, a small town in Georgia, became the carpet capital of the United States. It all comes down to the discovery of a trick to link tufts to carpets by a Georgia teenager in 1895. The emergence of Santa Clara county, California (Silicon Valley) as the center of the computer industry and the emergence of the so-called edge cities (Garreau, 1992) can also be explained along the same lines: small accidents start a cumulative process in which the presence of a significant number of firms and workers at a particular location draws even more firms and workers, and so on *ad infinitum*.

On a more general level, path dependence is a key element in North's (1990) theory of institutional change. According to North (1990, p. 7), "[The] path of institutional change is shaped by the lock-in that comes from the symbiotic relationship between institutions and the organizations that have evolved as a consequence of the incentive structure provided by those institutions." If we accept North's argument to the effect that institutions "are the underlying determinant of the long run performance of economies" (107), a simple logical reasoning will imply that path dependence provides an important clue to what is perhaps the ultimate economic question: What accounts for the difference of economic performance among nations?

On an even more general level, historians and the like have long stressed the fact that small chance events may have large downstream consequences.[14] It all started, of course, with Blaise Pascal and Cleopatra's Nose. Many have followed suit. Fogel (1989), for example, asserts that American politics in the midst of the last century was in so delicate a balance that small events could have caused a completely different outcome. In Fogel's words, "The overarching role of contingent circumstances in [the] ultimate victory [of the antislavery movement] needs to be emphasized. There never was a moment between 1854 and 1860 in which the triumph of the antislavery coalition was assured" (322). McPherson, cited by McCloskey (1990, p. 86), offers a similar picture of the same period, "Northern victory and Southern defeat in the war cannot be understood apart from the contingency that hung over every campaign, every battle, every election, every decision during the war". Other examples spanning different times and countries can surely be given by somebody with at least a passing knowledge of history.

Introspection can be also a rich source of evidence. Path dependence and lock-in by historical accident contribute greatly to shape the course of our lives. As we all

---

[13] IPM emphasizes biological controls as predators or sterile insects, crop rotation and localized pesticide application.

[14] Of course, not all historians are willing to concede that chance events play a crucial role. There has been a long tradition in history, from Marx to Braudel, that stresses a certain inevitability of the future development of human society along foreseeable lines. This view has come to be known as historicism. Karl Popper (1957), among others, has strongly opposed historicism. Popper is adamant about this. In his opinion, "for strictly logical reasons, it is impossible for us to predict the course of history".

know from personal experience, once we start down a given path, powerful forces tend to keep us on it. On the one hand, we all tend to get better in what we do, and on the other, we grow comfortable with, perhaps even to love, what we do.[15] Thus, we all have reasons to keep going in the same direction no matter how apparently innocuous the motives that led us to take it in the first place. As an adult needs barely to be told, inertia is a fact of life.

V. CONCLUSIONS

The former discussion makes it clear that history dependence is a distinct possibility not only in some abstract simple models but also in the real world. Now it is time for the definitive question: so what? In my opinion at least three different answers can be offered.

First, there is a methodological point to be learned. David (1985) puts it perfectly, "it is sometimes not possible to uncover the logic (or illogic) of the world around us except by understanding how it got that way". Thus, deductions from first principles aren't always enough to explain economic phenomena. This point is worth emphasizing given the willingness of economists to view everything as stemming from a few basic postulates. Perhaps we should all be more skeptical about deductions and more aware of history. Perhaps we should all be more prone to tell stories and less willing to pull out a model at every occasion.

Second, this paper makes it clear that long term predictions of the evolution of technology, institutions and public opinion —to cite just a few cases— are doomed to fail. In a world where chance events can have huge unforeseen consequences, long-run predictions are a joke. So the next time Bill Gates, Lester Thurow or any another professional (or amateur) futurist offer us their vision of the next century, we will have new reasons to be even more incredulous than before.

Last, this paper suggests that the common dilemma of *whether or not* to intervene might not always be the appropriate one. Oftentimes we should concern ourselves with a different question; namely, how to act in a manner that has a reasonable chance of leading to a desirable outcome. However, if path dependence and lock-in by historical accident are distinct possibilities, it becomes extremely difficult to be certain of *where and when* to intervene. Timing is paramount here. In the words of Paul David, there are only *narrow* windows in which policy can be effective.

---

15 Learning by doing and habit formation in economic jargon.

REFERENCES

Aiyagari, R., "Economic Fluctuations without Shock to Fundamentals; or, Does the Stock Market Dance to its Own Music", *Federal Reserve Bank of Minneapolis Quarterly Review*, 1988.

Andvig, J. and K. O. Moene, "How Corruption May Corrupt," *Journal of Economic Behavior and Organization*, 13:63-76, 1990.

Arthur, W. B., "Competing Technologies and Economic Prediction," *Options*, International Institute for Applied System Analysis, Austria, 10-13, 1984.

Arthur, W. B., "Competing Technologies, Increasing Returns, and Lock-In by Historical Events," *Economic Journal*, 99:116-31, 1989.

Arthur, W. B., *Increasing Returns and Path Dependence in the Economy*, Ann Arbor, Michigan University Press, 1994.

Arthur, W. B., Y. M. Ermoliev and Y. M. Kaniovski, "On Generalized Urn Schemes of the Polya Kind", *Kibernetica* 19:49-56. English Trans. *Cybernetics* 19:31-71, 1983.

Basalla, G., *The Evolution of Technology,* Cambridge, Cambridge University Press, 1988.

Berry, D. A. and B. Fristedt, *Bandit Problems,* New York, Chapman and Hall, 1985.

Bikhchandani, S., D. Hirshleifer and I. Welch, "A Theory of Fads, Fashion, Custom, Cultural Changes as Informational Cascades", *Journal of Political Economy*, 100:992-1.026, 1992.

Carrington, W., E. Detragiache and T. Vishwanath, "Migration with Endogenous Moving Costs", *American Economic Review*, 86:909-30, 1996.

Chu, C. Y., "Oscillatory vs. Stationary Enforcement of Law," *International Review of Law and Economics,* 13:303-15, 1993.

Cowan, R., "Nuclear Power Reactors: a Study in Technological Lock-In", *Journal of Economic History*, 50:541-67, 1990.

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

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

David, P. A., "CLIO and the Economics of Qwerty", *American Economic Review*, 75:332-37, 1985.

David, P. A., "Path Dependence and the Predictability in Dynamic Systems with Local Network Externalities: A Paradigm for Historical Economics", C. Freeman and D. Foray (eds.), *Technology and the Wealth of Nations,* London, Printer Publishers, 1992.

Dosi, G. and Y. Kaniovski, "On Badly Behaved Dynamics. Some applications of Generalized Urn Schemes to Technical and Economic Change", *Journal of Evolutionary Economics,* 4:93-123, 1994.

Ellison, G., "Learning, Local Interaction and Coordination", *Econometrica*, 61:1.047-1.071, 1993.

Fogel, R., *Without Consent or Contract,* New York, Norton, 1989.

Gaylord, R. J. and P. R. Wellin, *Computer Simulations with Mathematica*, Santa Clara (CA), Telos, 1995.

Glaeser, E., B. Sacerdote and J. Scheinkman, "Crime and Social Interactions", *Quarterly Journal of Economics*, 111:2, 507-48, 1996.

Gould, S. J., *Wonderful Life: The Burgess Shale and the Nature of History,* New York, Norton, 1989.

Granovetter, M., "Threshold Models of Collective Behavior", *American Journal of Sociology,* 83:1420-43, 1978.

Kandori, M., G. Mailath and R. Rob, "Learning, Mutation and Long-Run Equilibrium in Games", *Econometrica*, 61: 29-53, 1993.

Katz, M. and C. Shapiro, "Network Externalities, Competition, and Compatibility", *American Economic Review*, 75:424-40, 1985.

Katz, M. and C. Shapiro, "Technology Adoption in the Presence of Network Externalities", *Journal of Political Economy*, 94:822-41, 1986.

Kauffman, S., *At Home in the Universe,* Oxford, Oxford University Press, 1995.

Kindermann R. P. and J. L. Snell, "On the Relation between Markov Random Fields and Social Networks", *Journal of Mathematical Sociology,* 7:1-13, 1980.

Krugman, P., *Geography and Trade*, Cambridge (MA), MIT Press, 1991.

Krugman, P., *The Self-Organizing Economy*, Cambridge (MA), Blackwell, 1996.

Kuran, T., *Private Truths, Public Lies,* Cambridge (MA), Harvard University Press, 1995.

Liebowitz, S. J. and S. E. Margolis, "The Fable of the Keys", *Journal of Law and Economics*, 33: 1-26, 1990.

Liebowitz, S. J. and S. E. Margolis, "Network Externality: An Uncommon Tragedy", *Journal of Economics Perspectives*, 8:133-150, 1994.

Ligget, T. M., *Interacting Particle Systems,* New York, Springer Verlag, 1985.

McCloskey, D. N., *If You're so Smart*, Chicago, The University of Chicago Press, 1990.

North, D. C., *Institutions, Institutional Change and Economic Performance,* Cambridge, Cambridge University Press, 1990.

Petroski, H., *The Evolution of Useful Things,* New York, Vintage Books, 1992.

Popper, K., *The Poverty of Historicism,* London, Routledge & Kegan, 1957.

Puffert, D., *Network Externalities and Technological Preference in the Selection of Railway Gauges,* Ph.D. Dissertation, Stanford University, 1988.

Schelling, T., *Micromotives and Macrobehavior,* New York, Norton, 1978.

Smallwood, D. and J. Conlisk, "Product Quality in Markets where Consumers are Imperfectly Informed", *Quarterly Journal of Economics*, 43:1-23, 1979.

Sudgen, T., *The Economics of Rights, Cooperation, and Welfare,* Oxford, Blackwell, 1986.

Roth, A. and I. Erev, "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term", *Games and Economic Behavior*, 8:164-212, 1995.

Weidlich, W. and G. Haag, *Concepts and Models of Quantitative Sociology*, Berlin, Springer-Berlang, 1983.

Young, H. P., "The Economics of Convention," *Journal of Economic Perspectives*, 10:105-22, 1996.