SIEPR Discussion Paper No. 06-08

PATH DEPENDENCE – A FOUNDATIONAL CONCEPT FOR HISTORICAL SOCIAL SCIENCE

By
Paul A. David
Stanford University & University of Oxford

Revised January 2007

Stanford Institute for Economic Policy Research
Stanford University
Stanford, CA 94305
(650) 725-1874

PATH DEPENDENCE – A FOUNDATIONAL CONCEPT FOR HISTORICAL SOCIAL SCIENCE

By

Paul A. David
Stanford University and the University of Oxford

First draft: 19 January 2003
Second draft: 19 March 2005
Third draft: 23 May 2005
This version: 26 January 2006

Forthcoming in *Cliometrica — The Journal of Historical Economics and Econometric History,* v.1, no.2, Summer 2007.

ACKNOWLEDGEMENTS

The general shape of this paper took form as the Invited Lecture presented to the Symposium on *Twenty years 'QWERTY-effects' and path-dependence studies,* which was convened at The State University – Higher School of Economics, on 13th May 2005 in Moscow, Russia. I am grateful to Nureev Rustem (SU-HSE) for that invitation, and for his role with Leonid Borodkin (Moscow State University) and other Symposium participants, including V. Polterovich (Russian Academy of Sciences), in making the entire occasion both memorably enlightening and enjoyable. [An account of the symposium is available at: http://www.hse.ru/temp/2005/05_13_simpo.shtml, for those who read Cyrillic.] A subsequent presentation the paper to the First BETA-Workshop on Historical Economics, held on 20-21 May, 2005 at the Université Louis Pasteur de Strasbourg, provided opportunities for memorable discussions with Kristine Bruland, Claude Diebolt (the Workshop's organizer), James Foreman-Peck, Patrick Llerena, and Steve Redding, which contributed further to improving the exposition in the present draft.

Contact author: P. A. David, Department of Economics, Stanford University
Stanford, CA 94305-6072. Fax: 001+650+725-5702
pad@stanford.edu and paul.david@all-souls.ox.ac.uk

ABSTRACT

This introduction to the concept of path dependence, its pertinence for the development of historical social science, and its application in economic analysis and economic history proceeds from intuitive general ideas about history and historicity in narratives. It provides precise definitions of what is meant by describing a dynamical process as being "historical." Deterministic and stochastic formalizations of such dynamical systems are distinguished. The characterization of stochastic path-dependent processes as "non-ergodic" is explained in non-mathematical language by reference to concepts in probability theory, and a variety of representations of such processes in formal models is surveyed (including the Polya urn-process, certain kinds of Markov chain models, branching processes, and reversible spin systems) to show that while all display path dependence in other respects their properties are quite different. The diverse set of structural, micro-level conditions that can give rise to path-dependence is examined, and a further distinction is drawn between the property of path dependence and the existence of so-called 'QWERTY-effects' -- characterized by decentralized competitive market failures and consequent "lock-in" to Pareto-inefficient equilibria. Concluding sections consider the implications of the existence of non-ergodic dynamics for the methods of economic policy analysis, and the nature of the guidance that can be obtained in regard to public policy affecting endogenous technological change and institutional evolution.

Path dependence: when “history matters”

“Path dependence” is an important concept for social scientists engaged in studying processes of change, as it is for students of dynamic phenomena in nature. A dynamic process whose evolution is governed by its own history is “path dependent.” The concept, thus, is very general in its scope, referring equally to developmental sequences (whether in evolutionary biology or physics) and social dynamics (involving social interactions among economic or political agents) that are characterized by positive feedbacks and self-reinforcing dynamics.

Although the assertion that "history matters" has come to be coupled frequently with references to the concept of path dependence, the precise meaning of the latter term—and hence the significance of the former expression—more often than not remains rather cloudy. This is unnecessary as well as unfortunate. The fundamental idea is straightforward enough to be intuitively grasped without any instruction in economics; indeed, a thorough training in modern economics actually might interfere with human intuitions about history, and especially about processes involving historically contingent evolution. Even the formalizations of the concept of path dependence (to be introduced subsequently) are far from forbidding and readily will repay the effort spent in absorbing them. They will be seen to lend a useful measure of precision to descriptions of the special class of dynamical systems that are neither completely deterministic nor purely random in their workings, and in which the specific details of history govern the unfolding course of development.

A clearer grasp of what the term "path dependence" is about therefore ought to be part of the historical social scientist’s tool-kit. One would expect it to figure especially prominently in the analytical consciousness of economic historians, and to be equally familiar to all who are concerned to study the evolution of technologies, institutions, firms’ strategies and industry structures.

Empirical relevance

The foregoing prescription is not a full measure of the extent of the social and behavioral science domains that would be further illuminated by a working knowledge of the meaning and implications of this strong formulation of the idea that “history matters.” As the term is employed here, “path dependence” is merely the label for a particular class of dynamic phenomena, not a theory to account for the way that such systems behave. One may develop theories and frame hypotheses regarding the economic circumstances in which path dependent dynamics are likely to appear, and where the consequences of past events can be expected to exert greatest and most enduring influence. Whether a particular

dynamic process in a given economic setting is or is not path dependent remains an empirical question, and one that frequently is not so simple to resolve.

Much of the corpus of economic analysis upon which modern economic historians draw for their theoretical insights still envisages a form of dynamics that is path *independent*, in the sense that the unfolding sequence of events is not shaped in any significant and persisting way by particular events that occurred in the past. Such theories typically conceptualize change as continuous movement towards a prescribed outcome—upon which the details of the flow of events exercise no real causal influence. In its most elementary formulations, neoclassical theory encouraged economists to explain whatever patterns of resource allocation were observed by supposing each had emerged as a unique resultant of a particular set of given conditions. The (exogenous) "givens" were those affecting tastes (consumption "preferences," including preference between present and future satisfactions), factor endowments (population, natural resources), technological knowledge, and some legal and other institutions that established rules governing market transactions. On this view, it was useful to study economic history because it was a source of observations on the basis of which certain invariant "laws of economic motion" could be established; and also because history was seen—as Descartes saw it—to provide a fund of concrete examples that would illustrate the workings of those "time-less laws." Certainly there is ample space in which it is plausible to search for parsimonious theories of that kind, since, in some circumstances and over some intervals of time their simplified premises may correspond reasonably closely to the empirical realities. But, that will not be the case universally.

Indeed, there are good grounds for supposing that the opposite is true more often where one is concerned with the interplay of endogenous long-term changes in demography, in the state of technological knowledge, in institutional infrastructures, and in culturally formed tastes and the experientially conditioned *mores* and habits of individuals. These, of course, are the very stuff of economic and social history and the forces shaping the development and secular growth of economies. The elaboration of theories around the core concept of path dependent dynamics proves to be especially useful in those contexts. It encourages and enables the analytical historian and the economist alike to entertain the possibility that, in place of a unique-equilibrium-seeking dynamic, one should envisage a process that is seeking an evolving and historically contingent equilibrium. This provides a rationale for the methodology of *historical* social science, by indicating why it is that in order to uncover the logic (or illogic) of the world around us, we sometimes must first understand the details of how it got that way. A recognition of the existence of path dependence in economic processes—not simply as a possibility, but in some contexts as the more likely condition—thus accords a broader and more influential role for history, without sacrificing the rigor and precision to which modern economic analysis aspires.

Path dependence in analytical narrative history

Although the term can be defined precisely, and the concept of path dependence is amenable to further elaboration with the aid of tools from probability theory, it is possible also to grasp the core idea intuitively, by exercising common human understanding of the contingent character of seemingly inconsequential events that nevertheless set in motion a sequence of changes which make for an arresting outcome, and hence an interesting historical narrative. Any deterministic system's evolution will be governed by its initial conditions and its "law of motion" (the dynamic relations). Thus, the most primitive "plot structure" is presented by the familiar illustration of a simple physical system whose dynamics are path dependent: a drop of rain falling to one side of the continental divide eventually will reach the sea on that side; but had a gust of wind directed the falling droplet ever so slightly across the divide, it eventually would have joined the sea on the opposite edge of the continent. As stories go, however, this one has but a single moment of drama in it—the gust of wind. So we may turn to the no less familiar yet rather more complicated parable related by Ben Franklin's *Poor Richard's Almanac* (1733) to illustrate to the homily that "A little neglect may breed great mischief." This is the "tale of the horse-shoe nail", for want of which a shoe was lost, then a horse was lost, and for want of a horse a rider was lost. The narrative need not be extended further--byallusion to the loss of a battle, and even a kingdom—in order to appeal successfully to the common intuition that it is quite possible for eventual outcomes to be shaped by the transmitted effects of remote and seemingly trivial occurrences, and for the magnitude of those effects to become magnified as each event leads to the next.

Reference to elementary historical narratives such as these, or, indeed, to the chain of Old Testament "begats," are a helpful first step in conveying the core idea that lately has been more formidably labeled as "path dependent social dynamics" (Blume and Durlauf, 2001; Durlauf and Young, 2001). This phrase refers to systems of human social interaction, whose motions remain under the influence of conditions that are themselves the contingent legacies of events and actions played out within each system's history. In the realities of a social system's evolution, the events in question sometimes will be recent or obtrusive enough to be salient in the collective cultural identity of the society. That condition, however, is not essential; in other cases the critical events will have taken place in a past so remote that it has slipped from the conscious memories of the actors. Although economic agents' behaviors may be the deliberate resultants of rational calculation conditioned on currently prevailing circumstances, the particularities of that decision context itself may well be consequences of the contingent chain of events to which those remote critical events had given rise.

In order to delineate more precisely the dynamic properties that are referred to by the concept of path dependence, and to arrive at an analytical grasp of the main mechanisms that cause some economic and social processes

to exhibit dynamic properties of that type, it is necessary now to go beyond the preceding rather loose evocations of the idea of path dependent change. The drama of an historical narrative becomes more acute when the presence of just such a critical juncture, “a forking of the road” has been identified in the flow of events. In the tragic form of narrative, the action of the player or players when they reach that point are foreordained, and so it really does not matter whether or not they can be said to have seen what would befall them on each of the paths that lie ahead. This is not so in the stories that economic historians typically wish to tell. Economic analysis is essentially about human choices among what the philosopher William James termed “real options,” and the question of what the actors knew, and when they knew it, sought to know it, or at least formed some expectations of it are of interest for the analytical narrative.

Consequently, the quintessential and most persuasive form for an “historical” economic history narrative rightly involves those circumstances in which alternative local states of the world are plausible as well as merely conceivable. Articulating an entire counterfactual world is not what is called for here, but it nevertheless seems important for the narrator to show that when the traveler in the tale is paused at the putative “critical” fork in the road, there was an open path which would have led to events quite different from those that eventually transpired. Furthermore, for a tale to emphasize the highly contingent nature of the eventual outcome, the actors who are being followed until their arrival at that critical branch-point, or perhaps others who might have been in their place, ought to have been capable of choosing either of the paths forward. Then, as a poem on that subject by Robert Frost relates, the actual choice can be seen to have “made all the difference.”

Thus, among the most rhetorically satisfying narratives constructed by historians are those that identify and elucidate the role of critical human actions (or failures to act) that are shaped by transient and incidental circumstances—conditions that were not obviously pertinent to the principal issues of interest in the drama, yet from whose influence a succession of unanticipated and ultimately unwanted results unfolded. The masterful account of the debacle of the Spanish Armada by the English historian Garrett Mattingly (1959) is paradigmatic, in its reconstruction of the intricate chain of events that ensued from the Duke of Medina Sedona’s being temporarily distracted from overseeing the proper provisioning of many ships in his fleet with adequate stores of fresh drinking water. This tale gains interest not only from the exquisite detail of Mattingly’s chronicle and his sympathetic treatment of the Duke, but surely also from the implied historical counterfactual: were it not for the spoilt water in those cheaply made, rotting casks, the reign of Elizabeth I and the destiny of England among the European powers well might have turned out to be far less “Glorious.” (Again, this is a story in which, when told from the Spanish side, “a little neglect” is seen to breed “great mischief.”)

Another, more explicit and possibly more compelling presentation of the way that remote and transient circumstances left a profound and persisting impress upon national development emerges in the story of the separation of The

Netherlands from Belgium, as told by the great Dutch historian, Pieter Geyl (1936, 1961-1964). Geyl begins with the observation that the northern and southern Netherlands originally were neither politically nor religiously differentiated: they were a set of fiefs held jointly under Spanish dominion, and within each territory there was a similar division between Catholics and Protestants. The regional topographies were different, however. This turned out to matter when the Spanish sought to put down the revolt in the south, because Protestant refugees fled northward toward the safer havens afforded by that region's convoluted coastline. That shifting of the religious balance transformed the north into the center of Protestantism and successful resistance against Spanish sovereignty. When Spain ceded independence to the north, the divergence of religious and cultural orientations that had emerged between the regions' populations was further reinforced. Eventually, two centuries later, when, following the defeat of Napoleon, the Congress of Vienna decided to restore the unity of The Netherlands, the reunified state proved impossible to maintain in the face of the persisting divergence of religious and national sentiments.

One might see in Geyl's narrative a parallel with the biological process in which the separation of a land mass due to geological or climatological change allows genetic drift to create new traits which become fixed in a small, isolated segment of the animal species that originally was distributed throughout the entire territory. Had the carriers of those mutations been able to inter-mix and breed with the main population, it is likely that the recessive traits would have disappeared. But, the mutant strain could become fixed in an isolated sub-population and therefore could attain evolutionary stability: at a future date when the land mass became unified again (say, by the receding of waters and the emerging of land bridges), speciation would have proceeded to the point that interbreeding became biologically impossible. In a similar vein, the deep structure of Mattingly's story of the Armada could be held to bear a resemblance to the "tale of the horse-shoe nail," even to the appended moral about the adverse consequently that may come from "a little neglect."

To be sure, the fixing of a dominant religious orientation and the development of regional (and hence, ultimately national sentiment) is not a matter of biological replication. Likewise, in the simple physical model of the rain-drop paused above the continental divide, the action of the random zephyr on the water-droplet is the only bit of "history" that matters, and that story is at once too contrived and mechanical to provide interesting insights into the determinants of the course of a major battle at sea, let alone the complex forces that shaped the development of the Tudor State.

Even with this important *caveat*, however, the foregoing brief narratives serve to impart a positive message. The possible existence of a multiplicity of locally stable equilibria (dynamic attractors of the system) is seen to pose two generic explanatory challenges for the historian. First, what reason is there to suppose that there is more than a single equilibrium, especially when the apparent stability and continuity observed around us would suggest that there is only one? This is the part of the challenge that requires analysis of the existing

structure to demonstrate the plausibility of a counterfactual state. Second, if another state might have obtained, how did we arrive in this one? This is the problem of identifying the mechanism(s) of "selection"—which, in the conceptualization of path dependence as a branching process, corresponds to looking for critical bifurcations in the sequence of development, and for the factors that conditioned the actions taken at those historical junctures.

One must emphasize that we are concerned here only with “possibilities” and not with universal empirical certainties. The case for a more general appreciation of historicity in social and economic processes does not rest on the presumption that path dependence is ubiquitously present and of equal importance everywhere and in every historical epoch. Quite the opposite is true. Many of the problems that properly engage the attention of economists can be tackled perfectly well with the familiar tools for treating the familiar static equilibrium of supply and demand analysis, or with the analogous apparatus of “steady state” dynamics—because the situations in question can reasonably be characterized as possessing a unique globally stable equilibrium that is unique. But, on the other hand, the latter condition cannot be held to be a reasonable simplification everywhere, and at all times.

To be able to answer the question of where the concept of path dependence properly comes into it own—and conversely where, in that sense, it is reasonable to proceed under the simplification that “history does not really matter”—ought to be of interest both to economists and to historians seeking the right narrative modes to employ when relating their understandings of the past. In an important respect, identification of the respective properties of the many dynamic processes taking place in the world around us can be viewed at the very first thing one should try to establish, especially before undertaking actions intended to make a change for the better. Obviously, an elementary prerequisite for deciding such matters is the simple recognition of the question’s existence and pertinence, which is to say, the possibility that some economic processes are characterized by path dependence whereas other are not.

Precise definitions

Path dependence may be defined with regard to the relationship between a system’s process dynamics and the limiting configuration(s) that it ultimately takes on. This is one of two mutually compatible conceptualizations, and it is most straightforwardly employed when a deterministic process is said to have “led to a path dependent outcome”, i.e., a particular equilibrium among a number of potentially attainable limiting states. Alternatively, the property of path dependence may be defined for the class of stochastic processes, by referring not to “an outcome,” but instead to the nature of the limiting probability distribution that eventually will govern the movements of the system under examination.

One route to a precise definition of the term starts by distinguishing between path dependent dynamics, and all the rest. The latter appropriately are labeled “path independent”, because their dynamics guarantees that they converge to a unique globally stable equilibrium configuration regardless of where they started, or how they approached that eventual outcome. Models of deterministic convergence to a limiting position are very familiar in the literature of economic dynamics, much of which has been patterned on classical mechanics. In the case of a stochastic system, however, the eventual limit of the convergent process will be an invariant (stationary) asymptotic probability distribution that is continuous over the entire feasible space of outcomes. Putting this somewhat differently, in the limit a unique positive probability is assigned to every one of the configurations it could assume, or to every imaginable position the system can visit within some well-defined space. Thus, the probability of the non-recurrence of any one among the feasible states during the path independent system’s history is zero. No matter where it starts, a system of this kind will visit all of its states eventually—some of them more frequently than others; quite evidently it is able in due time to shake free from the influence of the initial conditions or of any among the past state(s) that it attained. Were that not the case, the system’s limiting behavior would not remain invariant throughout the course of its movements.

The technical term for the conditions just detailed is “ergodicity.” In physics, ergodic systems are said to be "connected," precisely because it is possible for them to transit directly or indirectly between any arbitrarily chosen pair of states, and hence for the process eventually to re-visit all the states that are compatible with the energy of the system. With that model as a reference point, path dependence may now be defined in negative terms as the property of systems whose dynamics are “non-ergodic.” The latter constitute the exceptions to the foregoing class of processes—where details of the system’s past motions were seen not to matter in the long run. Where the condition of non-ergodicity rules, however, random movements and other perturbations that disturb the system will not necessarily be “averaged away” with the passage of time. The Central Limit Theorem of statistics ceases to hold in this world, and random, transient perturbations can exert a lasting influence upon the movements of the system. This is the domain where the time series tracing the movements in, say levels of economic activity would be found to possess “unit roots” (Durlauf 1993).

Framing matters purely in these negative terms might suggest that path dependence is something of an “aberrant property” that contravenes the (ergodic) norm by leaving the system unable to shake free from the influence of events in its past. To avoid imparting such an anti-historical bias, it is helpful also to supply a positive definition, one stating what a path dependent process is, instead of what it is not. This is simple enough to do: a path dependent stochastic system is one possessing an asymptotic distribution that evolves as a consequence (function) of the process's own history. Historians should be reassured to notice (from the contrasting definition of a stochastic ergodic process) that the conceptualization of a contingent, path dependent dynamics as

"non-ergodic" is one that accommodates their traditional doubts regarding the likelihood of "history repeating itself" in any strict sense.

Formal representations in stochastic systems

The preceding definitions of path dependence subsume properties that are shared by a variety of deterministic and stochastic processes that in other respects possess quite distinctive features. Among the simplest representations are those stochastic systems that are formulated mathematically by an invariant set of finite (positive) "transition probabilities" that are *state-dependent*. The latter probabilities express the likelihood of being in any state (or category) *i* at time *t*+1, conditional on having been in any state *j* (including *j* = *i*) at time t. Such probabilities can be concisely arranged in matrix form by listing horizontally all the states where the system could be observed at time *t*, and listing vertically the (same) set of states to which it could have moved by time *t*+1. Transition probability matrices of this kind define what are referred to as "first order" or homogeneous *Markov chains*. (Their "first order" character signifies that it is only the current state occupied by the system, and no anterior states in its history, that affects the probability of where it will go next.) Homogeneous Markov chains are familiar constructs in economic models of the evolving distribution of workers among employment states, firms among size categories, family lineages among wealth-classes or socio-economic (occupational) strata, and the rankings of whole economies among in the international distribution of per capita income levels.

In such applications, it is quite common for the transition matrix formed by those probabilities to be specified in such a way as to ensure that the dynamic process is ergodic. The states of a Markov chain are said to be ergodic when they are *persistent*, *aperiodic* and *not null*. For a state that is ergodic, the probability of remaining in that state converges at the limiting, infinite step in the process to a positive constant—which is the inverse of the state's "mean recurrence time." From this definition of ergodicity, it follows that the distribution of the system's position among its possible states will be governed by the repeated iteration of the transition probability matrix, and will therefore converge eventually to an invariant asymptotic probability distribution in which positive probabilities are assigned to all the possible states.

When there is a single "absorbing state" or subset of connected states from which there is zero probability of escape to the subset of transient states, the Markov chain will converge weakly to that particular "attractor." Nonetheless, it remains ergodic—in the sense that regardless of where it may have started, the motion of the system eventually shakes off its initial conditions and goes to that unique absorbing state. On the other hand, when there are two or more absorbing subsets (distinct states, or regions of equilibria that are locally stable), the homogeneous Markov chain is *non-ergodic*: both the process represented by

the chain of transitions, and the eventual outcomes can be said to be path dependent.

One may see readily enough that in a system governed by a homogenous Markov process, if there exists at least one transient (non-absorbing) state from which (whether directly or indirectly) any of the multiple absorbing states can be reached, the realization of the random process of the system when it has arrived at such a point in its history—i.e., at the critical “forking of its path"—will “select” probabilistically among the alternative locally stable attractors. Thus, even within the class of probabilistic (Markov) models that has found wide application in the hands of conventional (“mainstream”) social scientists, it is necessary to acknowledge the empirical possibility of limiting outcomes being contingent upon improbable transitions that occur at some historically remote moment on the developmental path.

The foregoing definitions of path dependence also subsume other stochastic processes that possess a multiplicity of limiting distributions -- which generally is the case for “branching processes” or “branching systems.” In the latter class of non-Markovian probabilistic systems, the prevailing probabilities of transitions among states are functions of the sequence of past transient states that the system has visited. When a branching process is subject to local irreversibilities, it takes on the property of non-ergodicity. That is the case most familiarly in biological evolution, because speciation is a non-reversible action. Of course, as the previous metaphorical allusion to the “fork in the road” implies, there are other, analogous physical systems exhibit the same property. Indeed, observing certain physical systems, we are aware of the existence of critical conditions (such as externally controlled temperature levels) that can cause an irreversible “phase change” in their macro states—as in the phenomenon of liquefaction of gases, and the sudden alignment of the electrons in an iron rod that makes it a ferromagnet.

Among the stochastic physical systems whose non-ergodic dynamics have been the subjects of extensive formal investigation, and therefore offer analogies that have caught the interest of students of social system dynamics, the “reversible spin systems” studied in particle physics pertain to the dynamics of ensembles of interacting entities or agents that are finite in number (see e.g., Liggett 1985, particularly on the so-called “voter model,”;David 1988a, 1993a, 2005a, for applications). A contrasting case is the generalized Polya process, which envisages an urn containing colored balls that are being sampled with over-replacement, so that addition balls of the same color as the one drawn at each step, are returned to the urn. Here, convergence of one or another of the system’s stable equilibria (attactors) for the distinction of colors is attained only when the size of the population also goes to the infinite limit (see Arthur 1994; building on Arthur et al. 1983, 1986).

Prediction ex ante of the limiting (equilibrium) configuration is not possible in the case of the frequently cited path dependent Polya-urn model. But, for other processes that are every bit as path dependent, as the reversible spin systems, information about the initial configuration of the system is a sufficient

basis for assigning probabilities to the elements of the sub-set of configurations that represent the system's potential equilibrium states (see e.g., David 1993a). We hardly need to enter further into such esoteric technical matters to extract the conclusion that path dependence is not a property that is tightly coupled with other, very specific attributes of dynamic systems, such as the feasibility of predicting their limiting states. As shall be seen below, there is similarly no tight coupling between the questions of whether or not the outcome of a dynamic process of resource allocation is non-ergodic and whether or not it will be economically optimal.

"Branching processes" or "branching systems" constitute still another important and well-studied class of structures whose dynamics may be path dependent. It is interesting to notice the circumstances under which that condition will obtain, because they reveal what appears to be a more general condition that underlies this property. In a branching process the prevailing probabilities of transitions among states are functions of the sequence of past transient states which the system has visited. When a branching process is subject to conditions of local irreversibility's, however, it takes on the property of "non-periodicity". What is meant here by a local irreversibility is simply conveyed by the metaphor of a wanderer arriving at a "fork in the road", from which diverging trails lead to two or more distinct regions *between which there are no other connecting routes*. Suppose that if one continues along the right-hand path it would be possible to reverse direction and return to the fork; whereas traffic on the other path can only move forwards: this is the irreversibility. If the left-hand fork is taken, therefore, it will be impossible to return to the junction and so gain access to the right-hand branch. The wanderer's state space would have become "disconnected," and her future journey will be confined thereafter to positions reachable only along the left branch. This corresponds to the event of "speciation" in biological evolution.

Underlying causes

The existence of a multiplicity of locally stable equilibrium (dynamic attractors of the system) therefore presents analysts with two generic explanatory challenges. First, what reason is there to suppose that more than a single equilibrium exists, especially when continuities and structures observed in the social world around us could suggest that there is only one? This part of the challenge requires examination of the existing structure of the system to demonstrate the plausibility of a counterfactual equilibrium state. Second, if another such state might have obtained, how did we arrive at the one we are in? This poses the problem of identifying the mechanism(s) of "selection"—which, in the conceptualization of path dependence as a branching process, corresponds to looking for critical bifurcations in the sequence of development, and for the factors that conditioned the actions taken at those historical junctures. The selection mechanisms are manifold, and bewildering in their variety, which is why

in formal modeling exercises their operation typically is represented by the stochastic part of the system.

Something more general and illuminating than that, however, can be said in regard to the nature of the structural conditions that underlie the non-ergodic property of dynamical systems. A fundamental condition, which appears to underlie all of the instances in which economic resource allocation processes exhibit path dependent dynamics, is presence of micro-level irreversibilities—as in branching systems were a path forks into several tracks that remain separated and un-retraceable irreversibility. There is a simple economic interpretation of "irreversibility" in the foregoing context: a finite, and possibly substantial cost must be incurred to undo the effects of the resource allocation decision in question. Among the most readily recognizable economic irreversibilities are those associated with investment in durable assets, the cost of which are "sunk", i.e., not recoverable in some substantial part, even when the asset is not used. Under this broad heading one can list investments in the form of physical structures and equipment, in human capabilities acquired through education and training, in knowledge assets created by expenditures for discovery and invention. Similarly, one can view the reputations of individuals, firms and organization as belonging to the "assets" created by irreversible investments. Acquiring a deterrent reputation for readiness to engage in mutually destructive economic competition (e.g., price-wars), or credibility of commitment to a stated policy, usually requires undertaking some costly present resource expenditures that can be seen by others as a predictable response to a particular, externally recognizable situation. But actions that form "bad reputations" also may share the property of being irreversible, since it is likely to prove costly to neutralize their effects.

In this connection, it is important to emphasize that the individual assets in question (whether they are productive or damaging to their possessors) need not be *infinitely* durable. Quite the contrary: a system that has been built up from complementary components all of which have finite but *overlapping* service lives, under certain conditions—may collectively display the property of indefinite self-replication (see David 1971,1975:ch.5; Arrow 2004). This point sounds complicated and rather mysterious when it is expressed in these abstract terms, but one may grasp it quickly from a well-known parable of continuity in change: the story of "the ship of Theseus." In the course of that mythical vessel's long voyage, each timber, plank, and mast had to be replaced, one-by-one, so that when it returned there was not a single piece of the original structure which remained. Nonetheless, "the ship" had survived unchanged.

Not only in branching systems, but in other deterministic and stochastic processes, the existence of micro-level irreversibilities alone is sufficient to give rise to path dependent dynamics (see, e.g., Fisher 1983; Arrow 2000, 2004; David 1997b, 2001). This establishes that there is *no necessary connection between non-ergodicity and the presence of indivisibilities* (or "lumpiness") in the goods being produced, exchanged, and consumed in the system. Inasmuch as modern economic analysis attributes the phenomenon of increasing returns to

scale in production or consumption activities to the presence of invisibilities in one or more of the required “inputs,” a further proposition is implied: no necessary connection exists between conventionally defined “increasing returns” and the phenomenon of path dependence.

Assertions to the contrary, which appear all too frequently in recent writings on the subject, simply are wrong, or stem from an unconventional usage of the term “increasing returns”—applying it in reference to phenomena that actually do not involve indivisibilities. The latter is quite clearly the case in Arthur’s (1989) use of the phrase “increasing return to technology adoption.” The latter actually referred to self-reinforcing, positive feedback mechanisms governing decisions such as the choice whether to select one or another among alternative production techniques, or consumer goods, or geographical locations for production activities.

Where history matters most, for “mainstream” economists: market failures

It has been seen that the property of the large class of dynamic phenomena to which the term path dependence refers really is neither arcane nor esoteric, even when one is quite precise about its definition. This makes it rather puzzling that the intuitively natural representation of dynamics as being “historical” should not feature in so many of the dynamical models that economics have developed and sought to apply in understanding or prescribing for the world around them. Stranger still has been the resolute resistance encountered by the idea of path dependence in some quarters of modern academic economics, where the phenomenon itself is dismissed either as being devoid of any important economic consequences. Where that not sufficiently dismissive, path dependence as a conceptualization of the way that “history matters” in the economy also has been denounced as harboring the contention that “markets fail”; it therefore is suspect of being a sort of intellectual “Trojan Horse” which harbors within it dangerous economic policy prescriptions derived from ideological doubts that the best course for society is to entrust almost all matters involving resource allocation to “the workings of free markets” (see, e.g., Liebowitz and Margolis 1990, 1994, 1995).

The objection raised against giving attention to the economic implications of path dependence on the grounds that its consequences must be trivial would seem to put the cart before the horse: consequentiality of economic impact is an empirical matter that requires assessment of the findings of an extensive program of careful quantitative historical research, not a judgment arrived at by aprioristic arguments, much less than by tautologies fashioned by conveniently redefining what is meant by the concept of path dependence (see David 1997b, 2001b).

The second style of objection (i.e., that "path dependence" falsely implies that a decentralized competitive market process can result in a socially inefficient outcome) is a blatant logical *non sequitur*, because the concept (as understood

here) itself carries no necessary implications whatsoever in regard to the existence or non-existence of “market failure”. Nevertheless, frequent repetition of the contrary allegation has created considerable confusion on this point, a muddle that is best cleared up by examining it explicitly.

Before tackling that task, however, it may be remarked that the erroneous conflation that depicts the novel contribution of the idea that economic processes may be path dependent as residing in its implied theory of market failure is all the more peculiar because “mainstream” economics has a quite elaborate theory of the sources of sub-optimality in the workings of decentralized, perfectly competitive markets. Indeed, that body of theory has been around for quite a long time, whereas the appearance of formal ideas about non-ergodic dynamics is a much more recent departure in the economics literature. Moreover, it was in the context of studying topics in *static* general equilibrium analysis that economists arrived at these fundamental theoretical propositions about the existence of “market failure”—namely, that the Pareto optimality of allocations yielded by via atomistic, competitive markets was not guaranteed when a stringent set of convexity conditions on production and preference sets was violated, and/or when markets for all extant and contingent commodities did not exist.

Now, one reasonably may question the usefulness for pragmatic policy purposes of defining market failure in a way that makes the fulfillment of those exacting conditions the relevant benchmarks for deciding when markets are performing efficiently. Nevertheless, it remains utterly bizarre to suggest that the key contribution in introducing formal ideas about path dependence—a property defined in reference to dynamical processes (both deterministic and stochastic)—consists in reiterating those familiar propositions concerning market failure, propositions that are usually taught to beginning students of economics in the context of purely static welfare analysis.

Turning then to the question of the connection between the two ideas that are in play here, the difference between them is apparent. One of them, the core content of the concept of path dependence as a dynamic property, envisages the possibility that historical change may take the form of an irreversible branching process. The other idea is that it is entirely possible for some “branchings” in the course of a market-guided process to be deemed “socially regrettable,” because they led to inextricable economic inefficiencies that could have been averted—in some counterfactual but equally attainable course of historical development. Since the two ideas are quite distinct, for them to become logically coupled requires an argument that the presence of phenomena to which the concept of path dependence refers necessarily, implies the existence of conditions that can be shown analytically to cause competitive market processes to “fail”—yielding outcomes that are inefficient, and therefore sub-optimal from the societal viewpoint.

Economies of scale are among those troubling sources of competitive market failure, but it has be seen that path dependence *as such* may exist in the absence of invisibilities that give rise to economies of scale. So, it follows that

path dependence does not imply that competitive markets fail. Indeed, quite the contrary proposition holds: under full conditions of convexity in preference and production sets a *non-tatonnement* general equilibrium process (in which costless re-contracting is not available to the agents) will converge in a strictly path dependent manner on one among the continuum of valid 'core' solutions. It thereby satisfies the criterion that the outcome is socially welfare-efficient in the Paretian sense. This well-known theoretical result (see Fisher 1983) should be sufficient to expose the analytical error of supposing "market failure" to be among the necessary properties of a path dependent economic process.

Market efficiency, however, is not the only aspect of market outcomes that is of interest. In the process just envisaged, it is the distribution of the "gains from trade," and therefore the particular distribution of economic welfare among the trading parties—one such, among the many Pareto-efficient allocations at which the market process could have arrived—that will have been determined by the historical sequence of market transactions. To be sure, modern economic analysis devotes special attention to questions concerning allocativel efficiency, because, in principle if not in practice, that is a goal can be pursued without having to decide on whose welfare to reduce in order to raise the welfare of others. Given the centrality of that approach in classical welfare, it does seem important to acknowledge that among the great variety of structural conditions that can give rise to the existence of multiple equilibria and path dependent dynamics, there are some that also would prevent the workings of competitive markets from arriving unerringly at allocations that are socially efficient in the sense of Pareto.

Prominent in that company are micro-level irreversibilities in the behavior of agents, due to learning by doing and the habituation of tastes, and externalities affecting non-market interactions in the spheres of consumption and production, including those which give rise to co-ordination games that end in "co-ordination failures." Consequently, while the logical relationship between path dependence and market failure is neither one of necessity nor sufficiency, there are indeed some potentially important underlying connections between the existence of path dependence and the welfare optimality of decentralized resource allocation in market economies.

The latter observation clearly raises the general policy relevance of historical research that would undertake to identify where and when economic processes exhibit the property of path dependence. This accounts for the fact that among economists at large most of the interest in path dependence has tended to fix upon possibilities that suboptimal equilibria will be 'selected'—almost to the exclusion of all other considerations. Nonetheless, as has been noticed, there is more to economic life than welfare losses due to static inefficiencies. Market structures in some areas of the economy—especially those where network externalities and bandwagon dynamics are especially important—are particularly susceptible to the shaping influence of specific historical events in the evolution of the industries concerned. Insofar as market structure itself strongly conditions entrepreneurial behavior and managerial strategies of

innovation, path dependence then must be reckoned among the conditions that govern *dynamic* efficiency, as well as the static efficiency proprieties upon which traditional welfare analysis tends to focus.

But, questions about the distribution of welfare gains from trade and technological and institutional innovation surely do matter; these are not incidental issues that a socially useful discipline can afford to ignore, or to study with inadequate conceptual tools. Indeed, in private and public policy deliberations such questions often exert an influence far more powerful than that of narrow efficiency considerations. The identities of winners and losers in market rivalries understandably are matters of intense interest to the owners and employees of the affected enterprises, and to the governments of countries and regions where those agents are situated. More generally, all manner of political and social *sequelae*, as well as issues of equity and social justice, are attached to the dynamics governing the evolution of income and wealth distributions and related processes of socio-economic stratification. The analysis of positive feedback mechanisms that impart path dependent properties to those processes can significantly enhance economists' abilities to understand and predict distributional phenomena of the kind that recurrently have occupied the centre of many economic policy debates, and most likely will continue to do so in the future.

Path-constrained amelioration and its policy implications

There is, however, another way to look at the question of the policy significance of the conditions underlying path dependence as sources of the inefficiencies associated with "market failure." It may be that the selection of Pareto-dominated equilibria in positive feedback systems never results in inefficiencies that grow big enough (in the sense of Liebowitz and Margolis 1995) to impress contemporary observers, especially not those who can imagine a variety of clever, if costly, mechanisms for organizing collective escapes from locally sub-optimal situations. This, is a cogent empirical point to be made in objecting to highly simplified stories of "lock-in," particularly when a current sub-optimal local equilibrium is depicted as having been arrived at directly from some historically remote transient circumstances because a selection among alternatives available at that time irrevocably committed the system to develop along a particular path.

One surely should acknowledge and allow for the possibility that when economic or other disadvantages arising from past actions come to be properly perceived in an open social system, the self-interest of actors may be engaged to seize the opportunities it affords them to benefit by offering a more satisfactory solution. This could be a better product design, a new and more efficient technique of production, a different organizational structure, or even a reformed set of legal regulations. Is it one thing, however, to glibly contemplate the

possibility that effective responses of that sort would have been forthcoming, and a quite another matter – indeed, a considerable challenge to articulate the specific mechanisms by which one or another such response could be induced. The latter kind of historical research, however, is needed to establish the likelihood that under the prevailing circumstances particular corrective feedback mechanisms actually could come into play, and to determine how rapidly the "repair" could be effected, if indeed it would be forthcoming at all.

The contention that the process of "market competition" eventually works to rectify the mistakes of profit-motivated agents by harnessing the interests and capabilities of other profit-seekers—who will find opportunities for gain by eliminating existing sources of inefficiency—undoubtedly warrants serious consideration in this regard (see Liebowitz and Margolis, 1990; Puffert 2002). But, that means it ought not be accepted on faith. Rather the opposite seems to be the position adopted by those who argue that it is implausible to suppose that market incentives will not operate rapidly to eliminate sources of substantial inefficiencies in the production and use of commodities, and that the burden of proof lies upon those who claim that sub-optimal outcomes of decentralized market-guided choices in the past might have saddled subsequent generations with quantitatively significant continuing economics costs.

The static framework of welfare analysis within this sort of argument is conducted, and in which too many economists are still being taught to think, carries an ahistorical if not an anti-historical bias. It tends to suppress the natural human disposition to conceptualize the whole flow of current economic life as contingent upon the results of antecedent actions. Seen in truly historical perspective, a great deal of human ingenuity, especially both of the sort that is said to be "mothered by necessity" if not by opportunity, is devoted to trying to cope with the consequences of past mistakes whose economic costs are threatening to become "serious"; to assuring, somehow, that the more pernicious effects will be moderated, if not entirely abated. Of necessity (because we live without perfect foresight) much of this effort is remedial. It is undertaken *ex post*, by contriving technological "fixes" and "patches"; by commandeering temporary task forces to handle emergencies that established organizational structures and routines are discovered to be handling badly, by sustained efforts at "reforming" (not reinventing) long-standing institutions; and, yes, by concerted educational campaigns to "un-train" people who have through historical circumstance learned habits of mind, or unthinking reactions that are proving in one way or another to be personally and socially dysfunctional.

From that vantage point it seems only reasonable to acknowledge that under conditions that create common sources of path dependence and potentials for market failure, the unfolding developmental process is likely to involve more-or-less continuous processes of *path-constrained amelioration.* The latter is just a shorthand way of referring to incremental remedial resource allocation by intelligent, observant agents who are trying to cope with, and possibly gain something by removing unwanted features of the world left to them by the actions and inactions of their historical predecessors. This dynamic conceptualization

certainly seems preferable to entertaining the fantasy that history somehow unfolds without anybody noticing what is happening, and that after some considerable passage of time, when an equilibrium appears to have been reached (how can one tell?), people gather round, assess its social efficiency and consider whether there is something in the system that is worth fixing.

Still, there is more than a whiff of heterodoxy about the foregoing line of considerations. A long and deeply rooted tradition in the West, one that is intimately connected with the eighteenth century idea of 'Progress," has pre-disposed us to refer to all of that corrective investment of resources as contributing in greater or lesser degree to "economic welfare improvement." Rightly so, because in a *historically localized* sense, that is just what it is—amelioration. The ameliorative options, however, are more often than not inherited from past decisions and are quite tightly bounded by the existing critical situation. It is was the limited amount of lower memory on the mother-board of the first, hastily designed IBM PC's that created a need for devising clever software solutions to provide adequate temporary memory registers to execute the larger programs and larger data processing tasks that soon presented themselves. Another, more forward-thinking hardware specification with more lower memory surely was conceivable for the IBM engineers, and would for a while have obviated the recourse to compensating software creativity. Similarly, it was the existing software code that threatened to malfunction badly when the year 2000 dawned, not some other clock programs and data formats that had not been implemented, although those might well have been far more trivial to modify.

The resources spent in such perceived loss-avoidance activities are part of what we casually classify as productive investments that add to the net social product, and so should be included in the conventionally measured GDP. Yet, some part at least could just as well be thought of as the deferred costs of regrettable decisions made in haste to be remedied at leisure—albeit sometimes for great private profit. They might equally be labeled "regrettable economic opportunities" (see David 1999). To assess these imposed social costs it would be necessary to examine not only some eventual terminal situation, but the magnitudes of both the losses arising from the persisting inefficiencies, and the expenditures of resources in efforts to reduce them—whether or not such undertakings proved successful. Furthermore, it is important in considering situations where investment decisions involved choices among alternative technical standards or organizational rule-structures, that the persistence of un-removed diversity, and the consequent lack of interoperability in situations where there are potentially strong positive network externalities, can be a quantitatively greater burden than the selection of a technically sub-optimal standard (see Puffert 2002, 2004). Thus, the task of implementing a framework of analysis that takes account of the dynamics of path constrained amelioration is possible, but far from easy even when the historical case has been selected to render it less arduous.

The problem is that the selection of the more feasible cases is likely to bias the findings towards underestimation of the costs of path-constrained amelioration. Most of the situations in which the economic losses incurred by remaining stuck in a “bad” coordination equilibrium really could be very large are those in which the particular institution, or technology, or complex of social norms that is being examined has become highly elaborated and deeply embedded in numerous activities throughout the economy and society. In a sense, it is that condition which makes the question attractive. But, to fully assess the impact of the course of development that was involved in such cases -- for example, whether the outcome of the “Battle of the Systems” between Edison’s direct current and Westinghouse’s alternating current was socially optimal (see David 1992c), one would have to imagine a counter-factual world in which the whole general equilibrium course of evolution could have been very different.

Thinking about the implications of general purpose technologies, and the effects of major social institutions is one of the ways in which economists today, and economic historians particularly, are seeking to come to grips with this sort of challenge to systems analysis. Description of what happened generally remains more readily within reach than measurement of causal effects, precisely because of the difficulty of specifying what a plausible alternative world would have looked like. Little wonder, then that careful analytical and quantitative economic historians continue to draw back from attempting a full-scale counterfactual analysis when they approach such tempting questions as “What was the contribution to economic growth of the steam engine, or of the railroad, of the electric dynamo, and what will be the economic impact of the modern digital computer?” And the same holds, when they try to think about how to assess the effects of the eighteenth century codification of English common law, or the contribution of the corporate form of business organization to modern economic development.

“Accidents of history”?

In considering the nature of the policy lessons that might be drawn from the foregoing view of the incremental evolutionary development of complex technological systems, some remarks on the putative role played by ”historical accidents” in path dependent processes are now very much in order. Unfortunately, the use of that phrase itself is prone to cause misunderstandings. It is quite misleading to take it to suggest that some original economic irrationality or implementation error (accident) must be implicated whenever we find that positive network externalities have given rise to a sequence that turned out to be other than a globally optimal path. Indeed, only those who are hostile to the very idea of path dependence would repeatedly insist upon a literal interpretation of the phrase ”accidents of history.“ Doing so suggests that the essential feature of such processes is that the original actors in the drama—whether as contributors to the design of a technical system, or an institutional rule structure, or a

particular form of business organization, or as the initial adopters of such innovations—had to have been acting arbitrarily, or irrationally in the context of their economic circumstances. But, such an interpretation is not only logically unwarranted, it obfuscates an important and widely overlooked feature common to the histories of many network technologies, one that also has some bearing upon the way public policy might be approached in that and related area.

The facts of all the technological instances recently under re-examination—QWERTY, 64K lower memory in the IMB PC, AC vs. DC electrical current, light-water reactors, and VCR formats (see David 2001 for discussion and references) are quite consistent with the view that the behavior of the initiating actors of the drama, generally, was quite deliberate and not at all random in the sense of remaining inexplicable to the historian. Furthermore, it was reasonably conformable to the urgings of the profit motive. Yet, generally, their actions were also bounded by a parochial and myopic conception of the process in which they were engaging—in the sense that these decision agents were not primarily concerned with whether the larger system that might be (and was eventually) built around what they were doing would be optimized by their choice. In most cases they can be held to have failed entirely to foresee the complementary innovations and investments that would be influenced by their initial commitment to one rather than another course of action. In other words, their “failure of imagination” took the form of not thinking *systemically* about the technological and industrial structures that they were engaged in developing. Thomas Edison, of course, being a systems inventor *par excellence*, was an exception in that particular regard. Yet (as has been shown by David 1991, 1992c), Edison’s business strategy in the context of the “Battle of the Systems”—including his sudden decision to withdraw altogether from the flourishing electrical supply systems industry -- appears to have been driven by quite different, rather myopic, but nonetheless rational economic considerations.

In general, what was difficult for the pioneers in any area to foresee were the complementaries that would emerge subsequently, and in so doing open the possibilities of developing a more complex, distributed system whose components were not produced or purchased in an integral system. The Remington Co. engineers who put the finishing touches on the first commercially successful typewriters to carry QWERTY into the world did not dream of the possibility of touch-typing manuals (David 1985, 1986); nor did Edison (perhaps because he was a designer of closed systems) anticipate that anyone would devise an efficient and economical converter to link DC electrical supply facilities with distant users by way of polyphase AC networks. Similarly, in more modern times, the IBM engineers in Texas, as they rushed to create a readily producible personal computer (the famous PC), were not concerned with the amount of random access memory that would be needed (remarkably soon thereafter) to load a word-processing program like WordPerfect whilst keeping an Excel spreadsheet and a LAN-modem connection open and running in the background. Likewise, neither of the rival vendor groups behind the Sony Betamax and VHS cassette formats in the early VCR market had anticipated the commercial importance of prerecorded movies and video rental stores. (For the latter case,

one should refer to the detailed historical analyses of the VHS market in Baba and Imai 1990, Cusumano et al. 1992, and Grindley 1992, none among which are noticed by Liebowitz and Margolis 1994).

The point here is not that these inventors and entrepreneurs ought to have seen the shape of the future. Rather it is that the shape of the larger systems that evolved was built upon their work, and thus in each case preserved, and was in some respects much constrained by it—even in the way that those who came later coped with the problematic legacies of those initial decisions that had been taken quite deliberately but with quite other and considerably more evanescent considerations in mind.

Conclusions for policy

From the foregoing it may be seen that a proper understanding of path dependence, and of the possibilities of externalities leading to market failure, is not without interesting implications for economic policy. But those are not at all the sorts of glib conclusions that some critics have alleged must follow if one believes that history really matters—namely, that government should try to pick winners rather than let markets make mistakes. Quite the opposite view is closer to the policy position that I put forward more than twenty years ago, and have continued to argue in regard to technology choices among rival standards and institutional reforms that once taken were likely to develop “excess momentum” (David 1987; David and Bunn 1988; David and Greenstein 1990; David 2005b). If there is one generic course that public policy should pursue in such situations, it would be to counter-act the “excess momentum” of bandwagon movements in network product and service markets that can prematurely commit the future inextricably to a particular technical standard, before enough information has been obtained about the likely technological or organizational and legal implications of an early, precedent-setting decision.

In other words, preserving open options for a longer period than impatient market agents would wish is a major part of such general wisdom that history has to offer to public policy makers in all the applications areas where positive feedback processes are likely to be preponderant over negative feedbacks. Numerous dynamic strategies can and have been suggested as ways of implementing this approach in various, specific contexts where public sector action is readily feasible.

Still more sensible and practical approaches will be found if economists give up their exclusive obsession with traditional questions of static welfare analysis and, instead of pronouncing on the issue of where state intervention would be justified in the economy, begin to ask what kinds of public actions would be most appropriate to take at different points in the evolution of a given market process. Surely, bringing about that little shift toward “thinking in time” is the really big contribution that a properly *historical* approach grounded on the

concept of path dependent dynamics could make to the future of economics and other policy relevant social sciences.

References

Arrow KJ (2000) Increasing returns: historiographic issues and path dependence. *Eur J Hist Econ Thought* 7: 171-180

Arrow KJ (2004) Path dependence and competitive equilibrium. In: Guinnane TW, Sundstrom WA, Whatley WC (eds) *History Matters: Essays on Economic Growth, Technology, Demographic Change*, Stanford University Press, Stanford

Arthur WB (1989) Competing technologies and lock-in by historical small events. *Econ J* 99 (1): 116-131

Arthur WB (1994) *Increasing Returns and Path Dependence in the Economy*, University of Michigan Press, Ann Arbor

Arthur WB, Ermoliev YuM, Kaniovski YuM (1983) A generalized urn problem and its applications (in Russian; translated in Cybernetics 19:61-71). *Kibernetika*, 19:49-57

Arthur WB, Ermoliev YuM, Kaniovski YuM (1986) Strong laws for a class of path-dependent urn processes. In: Arkin, Shiryayev and Wets (eds) *Proceedings of the International Conference on Stochastic Optimization, Kiev 1984.* Springer, New York

Baba Y, Imai K (1990) Systemic innovation and cross-border networks: the case of the evolution of the VCR systems. Paper presented to the Schumpeter Society Conference on Entrepreneurship, Technological Innovation and Economic Growth, held at Airlie House, VA, USA, 3-5 June 1990.

Blume L, Durlauf S (2001) The interactions-based approach to socioeconomic behavior. In: Durlauf S, Young HP (eds) *Social Dynamics*. MIT Press, Cambridge, MA

Cusumano MA, Mylonadis Y, Rosenbloom RS (1992) Strategic maneuvering and mass-market dynamics: the triumph of VHS over Beta. *Bus Hist Rev*, 66(Spring):51-94

David PA: *See Bibliography (below) for works on path dependence,1969-2006.*

Durlauf S N (1993) Nonergodic economic growth. *Rev Econ Stud* 60(2):349-366

Durlauf S, Young HP (2001) The new social economics. In Durlauf S, Young HP (eds) *Social Dynamics*. MIT Press, Cambridge, MA

Fisher FM (1983) *The disequilibrium foundations of equilibrium economics*. Cambridge University Press, New York

Geyl P (1936) *The Netherlands divided (1609-1648)* [Translated by S. T. Bindoff in collaboration with the author]. Williams & Norgate, London

Geyl P (1961-1964) *The Netherlands in the seventeenth century*. 2 vols. Barnes and Noble, New York

Grindley P (1992) *Standards, business strategy and policy: a casebook,* London Business School, London

Liebowitz SJ, Margolis SE (1990) The Fable of the Keys. *J Law Econ* 33:1-25

Liebowitz SJ, Margolis SE (1994) Network externality: an uncommon tragedy. *J Econ Perspect* 8(2):133-50

Liebowitz SJ, Margolis SE (1995) Path dependence, lock-in, and history. *J Law Econ Org* 11(1):205-26

Liggett TM (1985) *Interacting particle systems (Grundlehren der mathematischen Wissenschaftern 276).* Springer, Berlin

Mattingly G (1959) *The Armada*. Houghton-Mifflin, Boson

Puffert DJ (2002) Path dependence in spatial networks: the standardization of railway track gauge. *Explor Econ Hist* 39:282-314

Puffert DJ (2004) Path dependence, network form and technological change. In: Guinnane TW, Sundstrom WA, Whatley WC (eds) *History Matters: Essays on Economic Growth, Technology, and Demographic Change*, Stanford University Press, Stanford

Bibliography: Works by P. A. David on Path Dependence, 1969-2006

A chronological listing of publications dealing explicitly with conceptual and methodological aspects of path dependence, macro-level irreversibilities and hysteresis in economic processes. (**Co-authors' names** appear in boldface).

(1969), “Transport innovation and economic growth: Professor Fogel on and off the rails”, *Economic History Review*, **22** (3), December, pp. 506-25.

(1971), “The landscape and the machine: technical interrelatedness, land tenure and the mechanization of the corn harvest in Victorian Britain.”, in D. N. McCloskey (ed.), *Essays on a Mature Economy,* London: Methuen, pp. 145-205.

(1975), *Technical Choice, Innovation and Economic Growth: Essays on American and British Experience in the Nineteenth Century*, Cambridge: Cambridge University Press.

(1985), “Clio and the economics of QWERTY,”, *American Economic Review*, 75 (2), May.

(1986), “Understanding the economics of QWERTY: The necessity of history,”, in W. N. Parker (ed.), *Economic History and the Modern Economist*, London: Basil Blackwell.

(1987), “Some new standards for the economics of standardization in the information age,”, in P. Dasgupta and P. L. Stoneman (eds.), *The Economics of Technology Policy*, London: Cambridge University Press.

(1988a), “Path dependence: putting the past into the future of economics,” *Institute for Mathematical Studies in the Social Sciences Technical Report 533*, Stanford University, November.

(1988b), “The economics of gateway technologies and network evolution: lessons from electricity supply history,”, (with **Julie A. Bunn**), *Information Economics and Policy*, Vol. 3, Winter, 165-202. (French transl. published in *Flux: Cahiers due recherché au des reseaux*, 1989.)

(1989), “When and why does history really matter?’ *A Presidential Address to the Economic History Association,* Delivered at the Smithsonian Museum of Science and Technology, Washington D.C., September. (Department of Economics Working Paper, Stanford University, October 1989.)

(1990), “The economics of compatibility standards: an introduction to recent research,” (with **S. Greenstein**), in *Economics of Innovation and New Technology*, 1(1 & 2), Fall: 3-42.

(1991), “The hero and the herd: reflections on Thomas Edison and the ‘Battle of the Systems’,” in P. Higonnet, D.S. Landes and H. Rosovsky (eds.), *Favorites of Fortune: Technology, Growth, and Economic Development Since the Industrial Revolution*, Cambridge, MA: Harvard University Press.

(1992a), “Path dependence and economics,” The 1991-1992 Marshall Lectures delivered at the University of Cambridge, April 28-29. Lecture I: ‘The invisible hand in the grip of the past’; Lecture II: ‘Models of non-ergodic economic dynamics, and their implications for policy’. (Center for Economic Policy Research Working Paper, Stanford University, August, 1992).

(1992b), “Path dependence in economic processes: implications for policy analysis in dynamical system contexts”, *Background Paper – Rosselli Foundation Workshop on Path Dependence*, Torino, Italy, 29-30 May. (Center for Economic Policy Research Working Paper, Stanford University, August, 1992; revised and published as David (2005a).

(1992c), “Heroes, herds and hysteresis in technological history,” *Journal of Industrial and Corporate Change*, 1(1): pp. 129-80.

(1993a), “Path dependence and predictability in dynamic systems with local network externalities: a paradigm for historical economics,” in *Technology and the Wealth of Nations*, D. Foray and C. Freeman, eds., London: Pinter Publishers. (French transl., in *Technologie et la richesse des nations*, Paris, 1993.)

(1993b), “Historical economics in the long run: some implications of path dependence,” in *Historical Analysis in Economics*, G.D. Snooks, ed., London: Routledge.

(1993c), “Intellectual property institutions and the panda’s thumb: patents, copyrights, and trade secrets in economic theory and history,”, in M. Wallerstein, et al. (eds.), *Global Dimensions of Intellectual Property Protection in Science and Technology*, Washington, DC: National Academy Press.

(1993d), “Percolation structures, Markov random fields and the economics of EDI standards diffusion,” (with **Dominique Foray**), in *Global Telecommunication Strategies and Technological Change*, G. Pogorel, ed., Amsterdam: Elsevier Science Publishers.

(1994a), “Dynamics of technology diffusion through local network structures,”, (with **Dominique Foray**) in *Evolutionary Economics and Chaos Theory: New Developments in Technology Studies,* L. Leydesdorff, ed., London: Pinter Publishers.

(1994b), “Positive feedbacks and research productivity in science: reopening another black box,” Ch. 8.in *Technology and Economic Change*, O. Grandstrand ed., Amsterdam: Elsevier.

(1994c), “Les standards des technologies de l’information, les normes de communication et l’état: un problème de biens publics,“ Ch. 10 in *L'analyse économique des conventions*, A. Orleans, ed., Paris: Presses Universitaires.

(1994d), “Why are institutions the ‘carriers of history’? Path dependence and the evolution of conventions, organizations and institutions”, *Structural Change and Economic Dynamics*, **5** (2): pp.205-20.

(1997a), “Making use of treacherous advice: cognitive progress,” Bayesian adaptation and the tenacity of unreliable knowledge’, (with **Warren C. Sanderson**), Ch. 12 in *Frontiers of the New Institutional Economics,* J. V. Nye and J. Drobak, eds., San Diego, CA: Academic Press, chapter 12.

(1997b), “Path dependence and the quest for historical economics: one more chorus of the ballad of QWERTY,” University of *Oxford Discussion Papers in Economic and Social History*, No. 20 (November). [Available at: <http://www.nuff.ox.ac.uk/economics/history/pap20>]

(1998a), “Marshallian externalities and the emergence and spatial stability of technological enclaves,” (with **Dominique Foray and Jean-Michel Dalle**), *Economics of Innovation and New Technologies* (Special issue on Economics of Localized Technical Change, ed. C. Antonelli), 4(2&3):147-82.

(1998b), “Communication norms and the collective cognitive performance of “invisible colleges,” in *Creation and Transfer of Knowledge: Institutions and Incentives*, G. B. Navaretti, P. Dasgupta, K.-G. Maier and D. Siniscalco, eds., Berlin-Heidelberg: Springer-Verlag.

(1998c), “From the economics of QWERTY to the millennium bug”*, Stanford University Economics Department Newsletter*, Stanford CA, Fall 1998.

(1999a), “At last, a remedy for chronic QWERTY-skepticism!*,”* Discussion Paper for the European Summer School in Industrial Dynamics (ESSID), held at l’Institute d’Etudes Scientifiques de Cargèse (Corse), France, September. [Available at: http://www-econ.stanford.edu/]

(1999b), “Krugman’s Economic Geography of Development: NEGs, POGs and Naked Models in Space,” *International Regional Science Review*, 22(2), August 1999:162-172.

(2000a), “Path dependence and varieties of learning in the evolution of technological practice,” Ch. 10 in *Technological Innovation as an Evolutionary Process*, John Ziman, ed., Cambridge, Eng.: Cambridge University Press.

(2000b), “‘Just-in-time’ for Y2K, and next,…for irreversible global warming*?”, Invited Economic Memoranda for the President of the United States*, White House Office of Communications, 3 March.

(2000c) “Path Dependent Learning, and the Evolution of Beliefs and Behaviours,” in *The Evolution of Economic Diversity*, U. Pagano and A. Nicita, eds., London: Routledge Publishers.

(2001) “Path Dependence, its Critics, and the Quest for ‘Historical Economics,” in *Evolution and Path Dependence in Economic Ideas: Past and Present*, eds. P. Garrouste and S. Ioannides. Cheltenham, Glos., UK: Edward Elgar.

(2003a) “Thinking Historically About Economic Challenges: Editors’ Introduction,” (with **M. Thomas**), in *The Economic Future in Historical Perspective*, P.A. David and M. Thomas, eds., Oxford University Press for The British Academy: pp. 1-27.

(2005a) “History Matters – QWERTY Matters,” The Tawney Lecture for the Economic History Society, presented at the EHS Conference in Glasgow, Scotland, 1 April, 2001. (Revised 15 July 2005; expanded version under preparation for publication.)

(2005b) “Path Dependence in Economic Processes: Implications for Policy Analysis in Dynamical Systems Contexts,” in *The Evolutionary Foundations of Economics*, Kurt Dopfer, ed., Cambridge, England: Cambridge University Press, Spring 2005.

(2005c) “Path Dependence and Historical Social Science: An Introductory Lecture.” An Invited Lecture to the Symposium on *Twenty years of ‘QWERTY-effects’ and path dependence studie,* held at State University-Higher School of Economics, Moscow, 13 May, 2005. SIEPR Policy Paper No. 04-022 [Available at: http://siepr.stanford.edu/papers/pdf/04-22.pdf].

(2006) "Europe's Universities and Innovation--Past, Present and Future," SIEPR Discussion Paper, Stanford University (March). [Available at: http://siepr.stanford.edu/papers/pdf/06-10.html].