Path dependence, its critics and the quest for 'historical economics'

By

Paul A. David
All Souls College, Oxford & Stanford University

First draft: November 1998
Second draft: June 1999
This version: June 2000

This paper has evolved from the author's 'Keynote Address' to the European Association for Evolutionary Political Economy, at their Meetings held in Athens, 7-9 November 1997. It has been revised for publication in *Evolution and Path Dependence in Economic Ideas: Past and Present*, edited by P. Garrouste and S. Ioannides, and forthcoming in 2000 from Edward. Elgar Publishing, Cheltenham, England.

ABSTRACT

The concept of path dependence refers to a property of contingent, non-reversible dynamical processes, including a wide array of biological and social processes that can properly be described as 'evolutionary'. To dispel existing confusions in the literature, and clarify the meaning and significance of path dependence for economists, the paper formulates definitions that relate the phenomenon to the property of non-ergodicity in stochastic processes; it examines the nature of the relationship between between path dependence and 'market failure', and discusses the meaning of 'lock-in'. Unlike tests for the presence of non-ergodicity, assessments of the economic significance of path dependence are shown to involve difficult issues of counterfactual specification, and the welfare evaluation of alternative dynamic paths rather than terminal states. The policy implications of the existence of path dependence are shown to be more subtle and, as a rule, quite different from those which have been presumed by critics of the concept. A concluding section applies the notion of 'lock-in' reflexively to the evolution of economic analysis, suggesting that resistence to historical economics is a manifestation of 'sunk cost hysteresis' in the sphere of human cognitive development.

Keywords: path dependence, non-ergodicity, irreversibility, lock-in, counterfactual analysis
JEL Classification: A1,B0, C4, D9, N0, O3

Contact author: Prof. P. A. David, All Souls College, Oxford OX1 4AL, UK
Tel. 44+1865+279313; Fax: 44+1865+279299;
E-mail:paul.david@economics.ox.ac.uk or <pad@leland.stanford.edu>

Path dependence, its critics and the quest for 'historical economics'[1]

1. Introduction

Contemporary research and writing being undertaken in the genre of evolutionary economics can be viewed as part of a broader, more catholic intellectual movement, one that I would characterize as a quest for *historical social science.* Yet, a decade after it began to be trendy among economists to say that 'history matters', some things remain less than entirely clear about the possible meanings attached to that phrase, if indeed it is taken to carry any substantive content at all. For me, at least, the expression 'history matters' does carry a quite precise set of connotations, namely those associated closely with the concept of *path dependence*. The latter refers to a property of contingent, non-reversible dynamic processes, including a wide array of processes that can properly be described as 'evolutionary'. The set of ideas associated with path dependence consquently must occupy a central place in the future, historical social science that economics should become.

However, by now you may well have begun to wonder whether the matter of history mattering really has been greatly clarified by my tying it to a second catchy expression that, unfortunately like the first, has come to be invoked more frequently than it is defined. What *is* 'path dependence' anyway? Does it have a meaning more precise than the slogan: 'history matters'? Is it about 'the Economics of QWERTY' or about something more general? If we were to conduct a systematic survey, even one confined to the academic economics profession, it probably would confirm my casual impression that the rising popularity of the term 'path dependence' has spawned a variety of usages, a perceptible measure of confusion, and even some outright misinformation.[2] If there are few who are prepared to dissent from the assertion that 'history matters', there are more who wonder whether history matters in ways that are important for economists to think about, and, there are many more who hold diverse and sometimes contradictory notions of how it comes about that history matters.

My immediate task on this occasion, therefore, is to try to clarify the meaning and amplify the economic significance of 'path dependence'. My hope is that the results of such an undertaking will enable others better to appreciate some of the salient implications for our discipline of recovering a conceptualization of change as a process that is *historical*, including implications for the way economic policy analysis is approached. A task so simple to describe, however, is not necessarily so easy to perform. For one thing, much of the training of the modern economist tends to weaken the receipients' natural, intuitive understanding of historical causation. Consequently, some remedial work is required in addressing an audience of academic economists, many of whose members' advanced education will have left them severely incapacitated in this particular regard.

To put this differently, most of us have been well-schooled in working with mathematical economic models whose dynamics admit perfect reversibility and lack any strong sense of genetic causation. It strikes me that neither those economists who casually assign to the influence of 'history' the things for which their analysis does not adequately account, nor those sceptics who say 'Sure, history matters, but not for much', are adequately responding to the challenges posed by the quite different class of dynamic processes that generate sequences of causally related *events*. One of the things about 'events' that our everyday experience of change seems to confirm is that they happen – and never 'un-happen'. By contrast with the

realities of the world around us, recognition of which forces itself implicitly and often only incompletely into the consciousness of practicing economic advisors, much of the formal teaching of economic analysis refers to a very different and special class of dynamic processes in which all motion in the long run is 'continuous locomotion'. In the context of analytical structures of that kind, which are familiar enough to students of classical mechanics, 'change' may be said to occur without there being any specific, individual 'events' that have causal significance.

To abandon the learned habits of peering at the world of economics automatically and exclusively from the peculiar vantage point afforded by a certain and now certainly antiquated branch of physics, and to be able therefore to take up another and contrary perspective, cannot be simply a matter of un-learning. Something additional, and for many, something new has to be learned. That 'something' can stand alongside neoclassical economic analysis, and so enhance one's appreciation of the special features distinguishing that paradigm from what may be called *historical economics*.

In asserting that 'history matters' I do not maintain that in economic processes history always matters in the same ways. Nor would I contend that economic processes have worked in the same way throughout history. The issue of how much 'importance' should be attached to the particular category of path dependent dynamical processes, in the sense of what proportion of the changes occurring in the economy around us can best be understood in such terms, remains for me one that can must be addressed by empirical inquiries. But, like virtually all interesting empirical questions, this one cannot be resolved in an analytical vacuum. The very nature of the evidence that would be required to address it is prescribed by reference to alternative, analytical and statistical models that admit of historical changes that are path dependent, and changes that are path independent. Data acquire meaning only in the context of economic theory: as T.S. Ashton, the British economic historian, said long ago: 'the facts do not wear their hearts on their sleeves'.

To say that is not to diminish the value of 'mere facts', nor to dilute the force of the imperative to get details of the story straight. Examination of particular cases may serve to illustrate the phenomenon of path dependence, to exemplify one or another methodology of studying historical economics, and to identify and explore unresolved problems. The writing of a piece of economic history in this way may also be good fun, and, when it is well done it typically manages both to provide entertainment and to satisfy particular points of curiosity. To do it well, however, we must begin with some grasp of the conceptual issues and the theoretical framework that endows observations with meaning and import.

Therefore, on this occasion I am not going to delve into the details of selected historical cases, whether illustrative of the evolution of technologies, or of institutions and organizational forms, or of cultural beliefs. Historical economics needs greater investment in suitable theory, and the kind of theory that is required is harder than that upon which ahistorical economics has been able to rest. So I must ask that you forego for the present the enjoyment of another excursion into economic history, and, instead, attend more closely to the conceptual foundations that serve to underpin further researches into path dependence in the economy. There will be an ancillary benefit in following this course: by anchoring our discussions firmly on these foundations with the aid of some precise definitions of path dependence (in section 2), it is quite straightforward to dispose of the misleading presentations of the concept by sceptics and critics.[3] I can then proceed (in section 3) to try clearing up the confusion that has developed in the literature over the connection between path dependence and economic

inefficiency, before turning (in section 4) to take up the meaning and economic significance of the widely used term 'lock-in'.

After this necessary clearing of obscuring 'undergrowth' it will be seen (in section 5) that once we enter an explicitly dynamic framework, the questions of static welfare 'efficiency' and the meaning of 'market failure' become more complicated and involve subtle issues that the critics of path dependence have thus far failed to take on board. Moreover, the implications of path dependence for economic policy studies are in reality quite far reaching, in arguing for the abandonment of static welfare-analytic approaches to the problem of where government should intervene in the economy, and its replacement by explicitly dynamic analysis that asks whether 'now' is the time in this or that specific market. Moreover, the general thrust of the recommendations regarding issues of technology policy that emerge from considerations of path dependence, will more often than not turn out to be entirely opposite in nature to those that seem to be most worrisome to the concept's *laissez-faire* critics.

In sum, I am unable to find any compelling reasons why economic analysis should remain 'locked in' to an ahistorical conceptual framework, apart from the unfortunate hysteresis effects of 'intellectual sunk costs'. But those effects are real, and must be countered. Therefore, drawing upon the analogy offered by field models of physical systems that have multiple basins of attraction, I suggest (in section 6) that some injection of further, intellectual 'energy' is likely to be necessary in order for our discipline to free itself from the local region of 'low potential' in which it has too long remained trapped.

2. Almost everything you wanted to know about 'path dependence' – but were always afraid to ask

Path-dependence, as I wish to use the term, refers to a dynamic property of allocative processes. It may be defined either with regard to the relationship between the process dynamics and the outcome(s) to which it converges, or the limiting probability distribution of the stochastic process under consideration.

At the most intuitive level we may draw a distinction between dynamic processes that are path dependent, and the rest. The latter, path-*independent* processes, may be said to include those whose dynamics guarantee convergence to a unique, globally stable equilibrium configuration; or, in the case of stochastic systems, those for which there exists an invariant (stationary) asymptotic probability distribution that is continuous over the entire feasible space of outcomes – that is, a limiting distribution that is continuous over all the states that are compatible with the energy of the system.

Stochastic systems possessing the latter properties are said to be *ergodic*, and have the ability eventually to shake free from the influence of their past state(s). In physics, ergodic systems are said to be connected, in the sense that it is possible to transit directly or indirectly between any arbitrarily chosen pair of states, and hence, eventually, to reach all the states from any one of them.

Path dependent processes thus may be defined negatively, as belonging to the class of exceptions from the foregoing set of processes, in which the details of the history of the systems' motion do not matter – because they cannot affect its asymptotic distribution among the states. This leads us immediately to

A negative definition: *Processes that are non-ergodic, and thus unable to shake free of their history, are said to yield path dependent outcomes.*

In this connection, it may be worthwhile to notice that the familiar homogeneous Markov chain invoked in many applications in economics – models of population migration and spatial distribution, of income and wealth, and occupational and social status distributions, firm size distribution, and so forth – is characterized by an invariant set of state-dependent transition probabilities that are finite (positive), and for convenience in many applications contexts, are specified so as to ensure that the process is *ergodic*. The distributions of the individuals or firms whose motions among the states are governed by Markov chains of this kind will each converge to their respective, invariant asymptotic probability distribution – a distribution that is continuous over the entire feasible state space. (This unique limiting distribution is the one that emerges as the transition matrix operator is repeatedly iterated.) When there is an absorbing state or subset of connected states (from which the probability of escape to the subset of transient states is zero), the system will converge weakly to that single attractor. Obviously, such a system's behaviour is not deterministic, but it may be said to be 'pre-destined', in the sense of being governed from the outset by a unique asymptotic probability distribution.

However, when a state-dependent process has two or more absorbing subsets (that is, distinct regions of equilibria that are locally stable), the homogeneous Markov process becomes *non-ergodic, and its outcomes can be said to be path dependent.* In the trivial case in which the initial condition of the system was one or the other of the absorbing states, it is plain that whatever governed that selection would fix the limiting position of the system. Further, it is no less self-evident that if there is at least one transient (non-absorbing) state from which the multiplicity of absorbing states can be reached, directly or indirectly, then the realization of the random process at that point in the system's history (on its path) will select one rather than the other outcome(s) to which the system eventually must converge.

For many purposes, however, we would like to say what a path dependent process *is*, rather than what it is not. Help from the probability theorists can be invoked in order to do so in a precise way. Focusing upon the limiting patterns generated by a random process (thus characterizing a dynamic system), we have

> ***A positive definition:*** *A path dependent stochastic process is one whose asymptotic distribution evolves as a consequence (function of) the process's own history.*

This broader definition explicitly takes in processes that possess a *multiplicity* of asymptotic distributions, as generally is the case for *branching processes* – where the prevailing probabilities of transitions among states are functions of the sequence of past transient states that the system has visited. Branching processes that are subject to local irreversibilities share the property of non-ergodicity. The latter therefore characterizes the processes of biological evolution, because speciation constitutes a non-reversible event.

Transition probabilities that are not invariant functions of the current state are also the characteristic feature of so-called non-homogeneous Markov chains. Rather confusingly, however, probability theorists sometimes refer to the latter as having *path dependent transition*

probabilities, thereby contrasting them with the more familiar class of homogeneous (or first order) Markov chains whose transition probabilities are (current) *state* dependent.[4] But, as has been seen from the negative definition discussed above, path dependence of the transition probabilities is not a necessary condition for a process that generates path dependent outcomes.

The foregoing account of what the term 'path dependence' means may now be compared with the rather different ways in which it has come to be explicitly and implicitly defined in some parts of the economics literature. For the moment we may put aside all of the many instances in which the phrases 'history matters' and 'path dependence' are simply interchanged, so that some loose and general connotations are suggested without actually defining either term. Unfortunately much of the non-technical literature seems bent upon avoiding explicit definitions, resorting either to analogies, or to the description of a syndrome – the set of phenomena with whose occurrences the writers associate path dependence. Rather than telling you what path dependence *is*, they tell you some of the symptomology – things that may, or must happen when the condition is present. It is rather like saying that the common cold *is* sneezing, watering eyes and a runny nose. I can illustrate this with the following two passages:

> Path dependence is the application to economic systems of an intellectual movement that has lately come into fashion in several academic disciplines. In physics and mathematics, the related idea is called chaos – sensitive dependence on initial conditions. As chaos theory has it, a hurricane off the coast of Florida may be the fault of a butterfly flapping its wings in the Sahara. In biology the related idea is called contingency – the irreversible character of natural selection. Contingency implies that fitness is only a relative notion: survival is not of the fittest possible, but only of the fittest that happen to be around at the time. (Liebowitz and Margolis, 1995c, p. 33)

Elsewhere, the same authors propose a kindred explanation, albeit one that is slightly more formal:

> The use of path dependence in economics is, for the most part, loosely analogous to this mathematical construction: Allocations chosen today exhibit memory; they are conditioned on past decisions. It is where such a mathematical process exhibits 'sensitive dependence on initial conditions', where past allocations exhibit a controlling influence, that it corresponds most closely to the concerns that economists and others have raised as problems of path dependency [sic]. In such a case, 'insignificant events' or very small differences among conditions are magnified, bringing about very different outcomes. It is that circumstance that yields both the 'non-predictability' and 'potential inefficiency'... (Liebowitz and Margolis, 1995b, p. 210)

Much could be said about the inaccuracies in the texts just quoted. For the present, however, it will be sufficient to notice one thing that they do not say, and three things that they do say.

That path dependence is a property of *stochastic* sequential processes is *not* mentioned, and only the allusion to 'contingency' provides any hint of the subject's probabilistic context. Of course, in order to pick up this clue, one would need to suppress the extraneous and misleading surmise that 'contingency' has a meaning that is specific to (evolutionary) biology, where it 'implies' something about the nature of selections made on criteria of inclusive

fitness.[5] Even that slender clue, however, is disguised by the statements that would have us associate path dependence with *deterministic* chaos, and the property of 'sensitive dependence on initial conditions' which characterizes that class of dynamic systems. The coupling of path dependence with chaos constitutes the first of the three positive assertions to which I previously referred, and it is incorrect. What it reflects is a too common predilection among mainstream economic writers for transposing concepts and arguments that are probabilistic in nature into simple deterministic models.[6] This habit is often seriously misleading, and must be especially so where neither certainty equivalence nor the operation of the central limit theorem of probability can legitimately be presupposed.

The second and third assertions disclose the authors' reasons why path dependence should be denounced as a problematic departure from the economic mainstream. They allege that a dynamic system in which there is 'memory' will be unpredictable, and worse, that it will be characterized by a potential for generating inefficient resource allocations. Like the first of the triad of assertions, these too are simply incorrect. There are some classes of non-ergodic stochastic processes whose outcomes are predictable, and I shall say more about these in due course. Further, it is vitally important to insist on logically distinguishing between systems that have the general property of path dependence, and that special sub-category of non-ergodic dynamic systems that may display (as an additional attribute) a susceptibility to one or another form of 'market failure'.

The latter condition, of course, is the one that adherents of strict neoclassical orthodoxy seem to find especially troublesome. Although I partake in the interest that most modern economists show regarding the efficiency of economic resource allocation, an obsession with the spectre of inefficiency was not what motivated me to inject the notion of path dependence into wide economic discourse, or to associate it with the application of insights from formal models of non-ergodic stochastic processes. This confession ought not to come as a surprise, especially to those who have encountered material that I have published before and since the pair of essays in which Clio, the muse of History, was coupled with the emergence of QWERTY as the *de facto* standard for typewriter keyboards (David 1985, and 1986).

The concept of path dependence and the associated framework of analysis is anchored in my long-standing quest to integrate *historicity* into economics. I think it important to distinguish between that peculiar aim, and the broader objectives of the 'new economic history' movement during the 1960s and 1970s, which saw the wholesale importation of the apparatus of modern economic analysis and econometric techniques into the study of economic history. Although the use of the economist's preferred methods of study of the past, undoubtedly has proved extremely illuminating in many contexts, it had become evident to some within the field that new constraints and analytical contradictions had been created by trying to understand economic history – which is to say 'economic dynamics' – through the assiduous application of *ahistorical* concepts and tools. It was the prospect of resolving those problems within the framework of path dependence that made the latter attractive from my vantage point. Imagine, then, my utter surprise to find this approach being attacked as a rival paradigm of economic analysis, whose only relevance consisted in the degree to which it could be held to represent a direct rejection of the normative, *laissez-faire* message of neoclassical microeconomics!

3. Significance: does path dependence mean there will be inextricable inefficiencies?

> Welcome to the world of path dependence, a world governed not by our stars, not by ourselves, but by insignificant accidents of history. In this unpredictable world, small seemingly inconsequential decisions lead inexorably to uncontrollable consequences....In the world of path dependence...our expectations for market outcomes are turned upside down. The Invisible Hand does not work in the world of path dependence. (Liebowitz and Margolis 1995c, p. 33)

This passage, from the article 'Policy and path dependence – from QWERTY to Windows 95', published in the Cato Institute's journal *Regulation*, ironically describes what is purported to be the essential message of those propounding the concept of path dependence. It is the authors' general contention that path dependence really cannot hold much interest for economists, because the world of market economies does not conform to the one that they construe the concept to be describing; because *remedies* for unsatisfactory situations generally will be available, and found quickly by profit-hungry entrepreneurs attracted by the potential 'surplus' that is implicit in any seriously inefficient state of affairs. Hence, on this reasoning, the only sorts of path dependent phenomena which would warrant the attention of economists must be extremely rare occurrences.

But, as has been seen, the core content of the concept of path dependence as a dynamic property refers to the idea of history as an irreversible branching process. One must logically distinguish from this the idea that it is possible that some branchings are 'regrettable' because they created inextricable inefficiencies that, in some counter-factual but equally feasible world, could have been avoided. Moreover, it is plainly a mistake to impute to the economic theory of path dependence *as such* the set of propositions that underlie the second of these ideas, for the notion of market failure has been long established in the literature of welfare economics.

Actually, it is within the context of static general equilibrium analysis that economists developed the concept of 'market failure' – namely, that the Pareto optimality of allocations arrived at via atomistically competitive markets is not guaranteed *except* under a stringent set of convexity conditions on production and preference sets; and, further, it requires the existence of markets for all extant and contingent commodities. One may or may not accept the usefulness for pragmatic policy purposes of defining 'market failure' in a way that takes those conditions as a reference ideal. Analytically, however, it remains a total *non sequitur* to assert that the essence of path dependence – a property defined for analyses of dynamical and stochastic processes – consists in asserting propositions regarding the possibility of 'market failure' that were proved first in the context of purely static and deterministic models.

Quite the contrary proposition holds: under full convexity conditions a *non-tatonnement* general equilibrium process can be shown to converge in a strictly path dependent manner on one among the continuum of valid 'core' solutions which satisfy the criterion of Pareto optimality (see Fisher 1983, and David 1997b). This should be sufficient to expose the logical error of claiming that the essential difference between models of path dependence and standard neoclassical analysis must be the former's insistence on the presence of 'market failure'. To be sure, there are some underlying connections between the existence of conditions that give rise to path dependence in economic processes, and the possibility that the workings of competitive markets in those circumstances would result in allocations that are inefficient.

But the circumstances in which competitive markets will not yield Pareto efficient outcomes are not in themselves either new, or arcane.

It might then be noticed that the taxonomy of path dependence proposed by Liebowitz and Margolis (1995b), and curiously described as 'definitions of path dependence', embraces a classificatory principle that is based entirely on static optimality criteria. Inasmuch as such criteria remain conceptually orthogonal to the nature of the dynamical processes under consideration, it is perhaps not surprising to observe that the definitions offered by Liebowitz and Margolis for 'first-degree' and 'second-degree' path dependence do not actually serve to distinguish between dynamic systems that are path independent and those that are path dependent. The first-degree form describes a situation in which all the outcomes among which selections might be made are not Pareto-ranked, such as would exist for the Nash equilbria of in a pure coordination game; the second-degree situation is one in which the outcome realized is dominated by a feasible alternative, yet represents the unavoidable *ex post* consequence of having taken an action that *ex ante* represented the 'best' strategy.[7]

In discussing the conceptualization of third-degree path dependence in which there is market failure leading to inefficiencies of an 'irremediable' kind, Liebowitz and Margolis (1995b) make reference to the test of 'remediability' suggested by Oliver E. Williamson. But, they entirely omit mention of the important distinction that Williamson's (1993) work drew between remediability through 'private ordering' and through 'public ordering'. Nowhere in the literature dealing with theoretical and empirical aspects of path dependent economic phenomena have I found it said that this property leads to outcomes for which remediation via public ordering is wholly *infeasible*. For the state to undertake to 'correct' a market outcome might become socially inefficient. But that is a different proposition from its being simply infeasible. So, it is not open to the critics to claim that path dependence would have empirical or policy substance for economists if only it did not exclude the possibility of remediation by public ordering in those circumstances where private ordering was unworkable.[8]

One certainly must agree that among economists at large most of the interest in path dependence results from the possibilities that sub-optimal equilibria will be 'selected' by a dynamic process. So it is understandable (and certainly to be expected) that brief treatments of points of controversy concerning theoretical contentions and empirical 'evidence' would tend to focus upon that question to the exclusion of everything else. Nevertheless, there is more to economic life than the possibility of welfare losses due to static inefficiencies. The identities of winners and losers in market rivalries is of interest to the owners and employees of the enterprises involved. The structure of industry itself may be of significance for dynamic efficiency through innovation and entrepreneurship. Indeed, the intense recent interest of the business press (and the Justice Department) in the positions of Microsoft and its present and future rivals in the market for web-browsers and related software, makes it plain that something more is perceived to be at stake than the comparative social rates of return on further incremental investment in their respective product lines.

More generally, all manner of political and social sequelae, as well as questions of equity, are attached to the dynamics governing the evolution of income and wealth distributions, and processes of socio-economic stratification. If analysis of positive feedback mechanisms that affect those aspects of life would significantly enhance economists' abilities to understand and predict the path dependent phenomena arising therein, does that not warrant at least some notice in assessment of the conceptual framework's significance?

4. The meaning of 'lock-in' in the historical context of path dependence

The current state of imprecision and confusion in discussions of the meaning and significance of the term 'lock-in' has not been alleviated by the use of 'lock in' as one among the taxonomic criteria applied to classify path dependent processes in the recent work of Professors Liebowitz and Margolis. Quite the reverse. I must begin by reiterating some doubts as to the coherence of creating a taxonomy for path dependent economic processes that turns upon whether or not it is possible to imagine a system being inextricably 'locked in' to a state that is locally and globally dominated by other allocative arrangements. Yet the latter would appear to be the very condition that is indicated, when the term is taken by Liebowitz and Margolis (1994, 1995b, 1995c) to refer to a situation where all the participating agents know they would derive a *net* gain by arranging by whatever means were necessary, collectively to exchange the status quo for some other available configuration.

By 'net gain', in this definition, is meant a surplus over and above the full costs of organizing and implementing the move to another state. *Ex hypothesis* there will be sufficient surplus in the new state to compensate everyone and leave someone better off after absorbing all the costs of negotiation, mechanism design, and insuring credible commitment that may be required to implement a collective escape. Therefore, in the circumstances thus posited, one would be hard put indeed to see how, if the agents involved were economically rational individuals, the status quo could have persisted long enough to be of interest. What is there in the imagined situation that would serve to lock in anyone to so unstable an attractor? Either we accept that people behave rationally and that such situations will be as scarce as hens' teeth, or this is a rendering of the notion of lock-in that would oblige economists to acknowledge that sometimes history that really matters is a result of the workings of the mysterious, the irrational, or the wildly improbable forces in economic life – or possibly all three.

By contrast, as the term 'lock-in' has been used in my work and that of Arthur (1989), it simply is a vivid way to describe the entry of a system into a trapping region – the basin of attraction that surrounds a locally (or globally) stable equilibrium. When a dynamic economic system enters such a region, it cannot escape except through the intervention of some external force, or shock, that alters its configuration or transforms the underlying structural relationships among the agents. Path dependent systems – which have a multiplicity of possible equilibria among which event-contingent selections can occur – may thus become locked in to attractors that are optimal, or that are just as good as any others in the feasible set, or that take paths leading to places everyone would wish to have been able to avoid, once they have arrived there.

From this vantage point, Arthur's (1989) phrase 'lock-in by small historical events' is evidently a gloss that should not be read too literally; it is a convenient contraction of the foregoing reference to the way in which trapping regions may be entered – although somewhat unfortunate, in allowing a hasty reader to suppose that the antecedent events somehow have *created* the local stability, or locked-in state. To be more precise, albeit more cumbersome, one should say that such configurations are self-sustaining (Nash) equilibria; that in the case of a path dependent process some particular historical event caused – that is, initiated the sequence of transitions that effectively selected, one rather than another among such configurations to be realized as the system's emergent property.

In some circumstances, as in the case of pure coordination games (where there are strategic complementarities in the dynamic interactions among agents) there is no Pareto-

ranking of a multiplicity of available equilibria from amongst which a path dependent, branching process can make a selection. *Which* coordination point is reached is a matter of welfare indifference to the parties involved. A coordination equilibrium, thus, provides us with the paradigmatic situation in which individuals are content to remain doing something, even though they would be happier doing something else if everybody would also do that other thing too. The reason they don't change what they are doing is, generically, that there are information imperfections that make it unlikely that a decentralized process can get everyone coordinated to move elsewhere, collectively.[9] Now notice that while incomplete information may be critical in blocking spontaneous escapes from dominated coordination equilibria, it is not a necessary condition for decentralized market processes to select such states. This is another reason why presenting 'lock-in' as a particular (pernicious, and supposedly uncommon) form of 'path dependence' is an invitation to further analytical confusions.

This last, important point can be elaborated on by observing that the generic problems of escaping from lock-in of the system to a globally inferior (but locally stable) attractor are rooted in 'pure' coordination costs. Such costs may be very high, however, especially if the individual agents are expected to act spontaneously under conditions of incomplete information. Hence, the nature of the *ex post* coordination problem generally is not the same as the problem of arranging coordination with agents who do not yet exist, or who have yet to recognize the complementaries between their interests and capabilities and those initiating the action. The sources of *ex ante* market failure that allow the system to be led into a globally inferior equilibrium are not necessarily the ones that make it very hard to get out.

Of course, if and when the structure of economic incentives and constraints bearing upon the process under study is altered by events that, for the purposes of the analysis may reasonably be regarded as 'exogenous innovations' (in the state of relevant knowledge, or in the regulatory institutional regime), the previous attractor(s) may be destroyed, freeing the system to endogenously begin to evolve some new configurations. Thus, the advent of microwave transmission technologies in the 1950s may be seen to have undermined the prevailing regulatory regime governing the US telecommunications industry (which had itself emerged through a path dependent process); and the denouement, in the event of the AT&T divestiture, brought into being a liberalized regulatory regime and new market structure that may be said to have formed new 'attractive paths', for the evolution of digital telecommunications technologies. But to claim that the evidence of change itself is sufficient to dispose of the notion of a persisting inefficient lock-in is tantamount to supposing that Schumpeter's gale of 'creative destruction' is blowing continuously at full force, through every niche, nook and cranny of the economy. Indeed, it is a way of losing one's sense of the variations in the flow of events through time that makes it interesting to read histories.

Strategic re-definitions, playing with words to avoid the force of the concepts with which they were originally associated, is a form of rhetoric that is essentially obscurantist. By the purely semantic trick of re-defining path dependence to come in various degrees of seriousness, and by associating the most 'serious' form to be, not a process, but a particular outcome state gauged in terms of allocational efficiency, it is possible to give superficial plausibility to the claim that no serious economic consequences are associated with the phenomenon of path dependence. This has been the taxonomic gambit tried by Professors Liebowitz and Margolis, who reserve their 'most serious' form of path dependence (third-degree) to be the state in which the status quo is Pareto-dominated *even after all transition and adjustment costs are considered.* They then can ask, rhetorically, why should one suppose that we would ever find a situation of 'serious path dependence', where people refused to make

themselves individually and collectively better off, after paying all the bargaining, transactions and information costs of arranging their escape from a bad situation? Why indeed? If one insists that the only sort of sub-optimality worth worrying about is the kind so wasteful as to justify escaping at any finite cost, then one is implicitly accepting the actual or equivalent loss of all the remedial expenditures (the costs of undoing the effects of outcomes we collectively prefer not to live with). Yet, those remedial expenditures might not have been unavoidable *ex ante*. Is it not pertinent for economists advising private and public agencies to consider the likelihood that some substantial portion of those costs were consequences of the path dependence of the dynamic process through which 'regrettable' outcomes were 'selected'?

Suppose, for the moment, that the significant economic question to be addressed in regard to the possibility of 'lock-in' is this: How can we identify situations in which it is likely that at some future time individuals really would be better off had another equilibrium been selected *ab initio*? By that we must mean that an alternative outcome would be preferred in some collective sense (perhaps by application of a compensation test) to the one that they are now in, and that they also (collectively) should be ready to incur some substantial costs to rectify the situation – assuming it was feasible to do so. Were it possible to answer that question by saying that such conditions will never obtain, then economists could well afford not to bother with the distinction between dynamic processes whose outcomes were path dependent and those which were path independent. It would be a distinction that might interest students of history, but would otherwise be inconsequential for economic policy. But such would be true only if multiple equilibria could be shown never to exist outside the context of pure coordination games (that is, where none are Pareto-dominated), or if it could be shown that it would never be possible to identify the structural conditions that give rise to other multiple equilibrium situations. We have no impossibility theorems of this sort, and neither of these propositions is likely to be established empirically.

5. Path-constrained melioration, the burdens of counterfactual historical analysis, and some policy implications

There is, however, another way to look at the question. It may be that the selection of Pareto-dominated equilibria in positive feedback systems is never allowed to become serious enough (in the Liebowitz-Margolis sense) to impress the contemporary observer who can imagine clever, if costly, mechanisms for organizing collective escapes from locally sub-optimal situations. This, indeed, is a cogent point, and deserves closer attention than it usually receives from economists who challenge the champions of historical economics to look around and find a 'really important' example – by which they seem to mean, a case of path dependent dynamics leading to a grossly inefficient equilibrium. Instead of imagining that history is played out without anybody noticing what is happening, and then, when an equilibrium appears to be reached people gather round and assess its optimality, we must allow for the process to encompass possibilities and consequences of incremental *path-constrained meliorating actions* being taken by observant, intelligent agents.

The static framework of welfare analysis within which too many economists are still being taught to do their thinking tends to suppress the natural disposition to conceptualize the whole flow of current economic life as contingent upon the results of antecedent choices. Seen in truly historical perspective, a great deal of human ingenuity, especially the sort that is said to be 'mothered by necessity', is devoted to trying to cope with 'mistakes' that are threatening to become 'serious' in their economic consequences; to assuring, somehow, that their more

pernicious effects will be moderated, if not abated altogether. This is done *ex post*, by contriving technological 'fixes' and 'patches', by commandeering temporary task forces to handle emergencies that established organizational structures are discovered to be handling badly, by sustained efforts at 'reforming' (not reinventing) long-standing institutions, and, yes, by concerted educational campaigns to untrain people who have acquired dysfunctional habits of one sort or another.

We like to refer to all of that activity as 'progress' and, in a historically local sense, that is just what it is: melioration. But the meliorative options are more often than not quite tightly bounded by the existing critical situation: it is the existing software code that threatened to malfunction badly when the year 2000 dawned, not some other programs and data formats that were not implemented, although they might well have been trivial to modify. The resources spent in such perceived loss-avoidance activities are part of what we are happy to consider productive investments, adding to the net product, whereas some part of it could equally well be thought of as the deferred costs of regrettable decisions made in haste to be remedied at leisure, and sometimes for great profit. They might equally be called regrettable economic opportunities (see David 1999).

Most of the situations in which the discomforts of remaining in a bad coordination equilibrium could be really large are those in which the institution, or technology, or behavioural norm has become highly elaborated and deeply embedded in numerous activities throughout the economy. One must then contemplate a counter-factual world in which the whole general equilibrium course of evolution would have been very different. Consideration of the implications of general purpose technologies is one of the ways in which economists today are coming to grips with this sort of systems analysis. Little wonder that economic historians have been and should be concerned primarily with such questions.

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. Such an interpretation is not only logically unwarranted; it obfuscates an important but widely overlooked feature common to the histories of many network technologies, and one that has some bearing upon the way public policy might be approached in that area.

The facts of all the technological instances recently under re-examination – QWERTY, 640K lower memory in the IMB PC, AC vs. DC electrical current, light-water reactors, and VCR formats too – are quite consistent with the view that the behaviour of the initiating actors of the drama, generally, was quite deliberate (not at all random in the sense of remaining inexplicable to the historian), and furthermore 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 concerned with whether the larger system that might (and was) being built around what they were doing would be optimized by their choice.[10] 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 from the flourishing electrical supply systems industry altogether – appears to have been driven by quite different, rather myopic, but nonetheless rational economic considerations.

In general, what were 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 integrally. 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; Edison had not anticipated 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, 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 pre-recorded movies and video rental stores.[11] Nor were the IBM engineers in Texas, as they rushed to create a readily producible personal computer, concerned with the amount of random access memory that would be needed to load a word-processing program like WordPerfect whilst keeping an Excel spreadsheet and a LAN-modem open and running in the background.

The point here is not that these folks 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 they coped with the legacies of those initial decisions, taken quite deliberately, but with quite other and in some measure more evanescent considerations in mind.

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 contrary, as I began trying to make clear more than a decade ago.[12] One thing that public policy could do is to try to delay the market from committing to the future inextricably, before enough information has been obtained about the likely technical 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 the generic 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 cease 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, start to ask what kind of public actions would be most appropriate to take at different points in the evolution of a given market process.

The 'first best' public policy role in these matters, therefore, is not necessarily the making of positive choices, but instead the improvement of the informational state in which choices can be made by private parties and government agencies. In the context of the recent literature on sunk cost hysteresis and options theory, one may see that the more history matters – because complementaries create irreversibilities in resource commitments – the more worthwhile it is to invest in being better informed prior to leaping. There is an evident opportunity cost in giving priority to investments in further information acquisition; quite standard economics can be relied on to balance the expected value of waiting (searching) for further 'news', against the anticipated costs to the current generation(s) of not allowing markets to make choices on the basis of the knowledge that is presently available. Obviously, some assessment of the rate at which the relevant information states are capable of evolving will turn out to be of critical importance in determining when a stage has been reached where it no longer is best to defer irreversible resource commitments.

6. Overcoming 'intellectual sunk cost hysteresis' and escaping from disciplinary 'lock-in' to ahistoricism

The cluster of ideas that are now identified with the concept of path dependence in economic and other social processes probably would not excite such attention, nor require so much explication, were it not for the extended prior investment of intellectual resources in developing economics as an ahistorical system of thought. For many economists, their own costs sunk in mastering that discipline have produced a facility for reasoning that suppresses natural, human intuitions about historical causation. They thus have a 'learned incapacity' (in Thorstein Veblen's apt phrase) to see how historical events could exert a causal influence upon subsequent outcomes that would be economically important. Perhaps unknowingly, such folk have fully internalized Aristotle's teleological principle of explanation, which rejected the method of reference to antecedents, and so escaped infinite explanatory regress by substituting forward-looking functionalism (as we would describe it). This was undoubtedly useful, even though it has had the intellectual side effect, in many disciplines, of encouraging the formal suppression of the intuitive impulse to refer to pre-existing states and intervening 'events' when asked to account for the way things are today.

Mainstream economics is not alone among the social sciences in providing a way to explain an existing state of the world by reference to the purpose or end (*telos*) that it serves, rather than to the conditions from which it may have evolved.[13] This has proved a source of deep insights into many matters, but not into all matters of concern to economists and students of broader cultural phenomena, such as the spread of languages and social communication norms.[14] Nor, for that matter, does it suffice to provide good accounts of biological phenomena. In modern Darwinian evolutionary theory there is a beautiful, productive tension between the teleological principle of natural selection according to inclusive fitness, and the antecedents principle, namely, that the possibilities of evolution are tightly constrained at every moment by the current contents of the gene pool, which is the product of species' history. Perhaps that is why we might be drawn towards evolutionary biology as 'the Mecca for economics'.

Modern economics in its ahistorical, convergence model formulation serves some intellectual purposes very well, and the perpetuation of the methodological status quo can be seen to serve still other rational private ends. Nevertheless, if that style of explanation was entirely satisfactory in accounting for all economic and social phenomena without reference to legacies from the past, some of us would not presently be so exercised by trying to adjust contemporary economic thinking to the notion that history matters – nor would others be strenuously resisting that adjustment. Path dependence is a concept requiring explication for many today, simply because so much of economics committed itself to theories that would make the results of choice behaviours consistent in the sense of being path independent. But there is no compelling reason to regard that as an exclusive commitment.

Path dependence, at least to my way of thinking, is therefore about much more than the processes of technological change, or institutional evolution, or hysteresis effects and unit roots in macroeconomic growth. The concepts associated with this term have implications for epistemology, for the sociology of knowledge, and cognitive science as well.[15] Nevertheless, it would be quite wrong to imagine that positive feedback dominates all aspects of economic life (let alone 'life'), just as it is unwarranted to proceed on the supposition that economic dynamics everywhere are intrinsically characterized by the operation of stabilizing, negative feedback systems. Considering the possibility that the former framework is the one most relevant in a particular context does not rule out the opposite conclusion, or preclude appropriate resort to the latter framework – the familiar convergence models of neoclassical economics. These really are not necessarily mutually exclusive tool-sets, or incompatible standards, that cannot be integrated into a larger intellectual system. Even though we should be aware of the workings of strong social processes, familiar in the sociology of knowledge, that can turn normal science procedures into exclusionary dogmas, it is not necessary for social and behavioural scientists to adopt positions that exacerbate and amplify those tendencies.

Once the concept and the ideas surrounding path dependence are properly understood, there can be no reason to construe them as necessarily corrupting the discipline of economics, or to fear that once admitted they would be subversive of all *laissez-faire* policies. There simply are no good grounds to go on actively resisting these ideas, which if accepted will lead us into previously little-explored regions of theoretical and empirical enquiry. Nor is there even a sound precautionary case for seeking to contain their spread until it can be determined what would become of the grand edifice of economic analysis as we know it, once the assumed global dominance of negative feedback processes were discarded. The logic of sunk cost hysteresis has a legitimate place in the conventional theory of optimal investment behaviour. Yet, when it is carried over and applied to the field of *intellectual* investments in new tools of economic analysis, the result is a self-defeating orthodoxy of thought and surely not the optimal progress of our discipline.

References

Works by Paul A. David

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.

(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.)

(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 D. Foray and C. Freeman (eds.), *Technology and the Wealth of Nations*, London: Pinter Publishers.

(1993b), 'Historical economics in the long run: some implications of path dependence', in G.D. Snooks (ed.), *Historical Analysis in Economics*, 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 G. Pogorel (ed.), *Global Telecommunication Strategies and Technological Change*, Amsterdam: Elsevier Science Publishers.

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

(1994b), 'Positive feedbacks and research productivity in science: reopening another black box, in O. Grandstrand (eds.), *Technology and Economic Change*, Amsterdam: Elsevier, chapter 8.

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

(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.

(1995), 'Dépendence du sentier et économie de l'innovation: Un rapide tour d'horizon', (with **Dominique Foray**), *Revue d'Economie Industrielle*: Special edition: 'Economie industrielle – développements récents, 1st trimester, 1995: pp.27-51.

(1997a), 'Making use of treacherous advice: cognitive progress, Bayesian adaptation and the tenacity of unreliable knowledge', (with **Warren C. Sanderson**), in J. V. Nye and J. Drobak (eds.), *Frontiers of the New Institutional Economics*, 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).
[On-line as: <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 G. B. Navaretti, P. Dasgupta, K.-G. Maier and D. Siniscalco (eds.), *Creation and Transfer of Knowledge: Institutions and Incentives*, Berlin-Heidelberg: Springer-Verlag.

(1998c), 'From the economics of QWERTY to the millennium bug', *Stanford University Economics Department Newsletter*, Stanford CA, Fall 1998/1999.

(1999), '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.

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

Other Works Cited

Arthur, W. B. (1988), 'Self-reinforcing mechanisms in economics', in *The Economy as an Evolving Complex System*, (Santa Fe Institute Studies in the Science of Complexity, **5**), Redwood City, CA: Addison-Wesley.

Arthur, W. Brian (1989), 'Competing technologies and lock-in by historical small events, *Economic Journal*, 99 (March): 116-31.

Arthur, W. Brian (1990), 'Industry location patterns and the importance of history', *Mathematical Social Sciences*, 19: 235-51.

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

Arthur, W. B, Yu. M. Ermoliev and Yu. M. Kaniovski (1983), 'A generalized urn problem and its applications', *Kibernetika*, 19: 49-57 (in Russian). Translated in *Cybernetics*, 19: 61-71.

Arthur, W. B, , Yu. M. Ermoliev and Yu. M. Kaniovski (1986), 'Strong laws for a class of path-dependent urn processes', *Proceedings of the International Conference on Stochastic Optimization, Kiev 1984*, Arkin, Shiryayev and Wets (eds.), New York: Springer (Springer Lecture Notes in Control and Information Sciences, p. 81.

Baba, Y. and K. Imai (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, June 3-5.

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

Brown, L. (1993), *The New Shorter Oxford English Dictionary, On Historical Principles*, Oxford: Clarendon Press.

Cowan, R. (1990), 'Nuclear power reactors: a study in technological lock-in', *Journal of Economic History*, **50**(3), September : 541-67.

Cowan, R. (1991), 'Tortoises and hares: choice among technologies of unknown merit', *Economic Journal*, **101**(407), July : 801-14.

Cowan, R. and P. Gunby (1996), 'Sprayed to death: path dependence, lock-in and pest control strategies', *Economic Journal*, **106** (436), May : 521-42.

Cusumano, M.A., Y. Mylonadis and R.S. Rosenbloom (1992), 'Strategic maneuvering and mass-market dynamics: the triumph of VHS over Beta', *Business History Review*, 66 (Spring): 51-94.

David, P. A., R.C. Maude-Griffin and G.S. Rothwell (1996), 'Learning by accident? Reductions in the risk of unplanned outages in US nuclear power plants after Three Mile Island', *Journal of Risk and Uncertainty*, 12: 175-98.

Durlauf, S. (1990), 'Non-ergodic economic growth and fluctuations in aggregate output', *American Economic Review*, **80** (3).

Durlauf, S. (1996), Neighborhood feedbacks, endogenous stratification, and income inequality', in *Dynamic Disequilibrium Modelling*, Cambridge: Cambridge University Press.

Eldridge, N. (1985), *Time Frames: The Rethinking of Darwinian Evolution and the Theory of Punctuated Equilibria*, New York: Simon and Schuster.

Fisher, F. M. (1983), *The Disequilibrium Foundations of Equilibrium Economics*, New York: Cambridge University Press.

Fogel, R. W. (1964), *Railroad and American Economic Growth*, Baltimore: The Johns Hopkins Press.

Föllmer, H. (1974), 'Random economies with many interacting agents', *Journal of Mathematical Economics*, 1: 51-62.

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

Grimmet, G. (1989), *Percolation*, New York: Springer-Verlag.

Grindley, P. (1992), *Standards, Business Strategy and Policy: A Casebook*, London: London Business School.

Koot, G. M. (1987), *English Historical Economics, 1870-1926*, Cambridge: Cambridge University Press.

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

Krugman, P. (1994), *Peddling Prosperity*, New York: W. W. Norton and Company.

Liebowitz, S.J., and Stephen E. Margolis (1990), 'The fable of the keys', *Journal of Law and Economics*, 33(1), April: 1-25.

Liebowitz, S.J., and Stephen E. Margolis (1994), 'Network externality: an uncommon tragedy', *Journal of Economic Perspectives*, **8** (2), Spring: 133-50.

Liebowitz, S.J., and Stephen E. Margolis (1995a), 'Are network externalities a new source of market failure? *Research in Law and Economics*, **17** (0):1-22.

Liebowitz, S .J., and Stephen E. Margolis (1995b), 'Path dependence, lock-in, and history', *Journal of Law, Economics, and Organization*, **11**(1), April: 205-26.

Liebowitz, S. and Stephen E. Margolis (1995c), 'Policy and path dependence: from QWERTY to Windows 95', *Regulation: The Cato Review of Business & Government*, 1995, number 3: 33-41.

Liggett, T. M. (1985), *Interacting Particle Systems (Grundlehren der mathematischen Wissenschaftern **276**)*, Berlin: Springer-Verlag.

McCloskey, D. N. (ed.) (1971), *Essays on a Mature Economy*, London: Methuen.

McCloskey, D. N. (1974), 'Victorian growth: a rejoinder to Aldcroft', *Economic History Review*, Second Series, 27 (2), May: 275-77.

McCloskey, D. N. (1976), 'Does the past have useful economics?', *Journal of Economic Literature*, 14(2), June: 434-61.

Mokyr, J. (1990), *The Lever of Riches: Technological Creativity and Economic Progress*, New York: Oxford University Press.

North, D. N. (1990), *Institutions, Institutional Change and Economic Performance*, Cambridge: Cambridge University Press.

Ruelle, D. (1991) *Chance and Chaos*, Princeton: Princeton University Press.

Ruttan, V.W. (1997), 'Induced innovation, evolutionary theory and path dependence: sources of technical change', *Economic Journal*, **107** (444), September: pp.1520-47.

Solow, R. M. (1986), 'Economics: is something missing?', in William N. Parker (ed.) *Economic History and the Modern Economist*, Oxford: Basil Blackwell, pp. 21-9.

Steward, I. (1990), *Does God Play Dice? The New Mathematics of Chaos*, London: Penguin.

Teggart, F. J. (1977), *Theory and Processes of History* (second paperback printing of the 1941 Edition of *Theory of History* (1925) and *The Processes of History* (1918), published in one volume), Berkeley, California: University of California Press.

Teggart, F. J. (1939), *Rome and China, a Study of Correlations in Historical Events*. Berkeley, California: University of California Press.

Williamson, O. E. (1993), 'Transaction cost economics and organization theory', *Industrial and Corporate Change*, **2** (2): pp. 107-56.

ENDNOTES

[1] I am grateful for the pithy comments that I received on a related earlier paper (David 1997b), from Avner Greif, Frank Hahn, Joel Mokyr, Robert Solow, Edward Steinmueller and Gavin Wright. Stavros Ioannides contributed very helpful editorial corrections. None among them should be held responsible for the deficiencies or excesses that remain in the present text.

[2] I hesitate to write 'dis-information' at this point, as that connotes intentions rather than consequences. I prefer to proceed on the supposition that those who have repeatedly misrepresented the meaning of the term in the course of criticizing 'path dependence' as an erroneous economic theory, and those who have deemed it to be an empty concept (in the sense that it is essentially devoid of empirical relevance for economists), simply are confused about its meaning.

[3] For this purpose it is best that I confront the critical treatment of path dependence by Professors Stanley Liebowitz and Stephen Margolis (1995b, 1995c). I therefore put to one side a rebuttal of the specific factual allegations that have been adduced in Liebowitz and Margolis's (1990) riposte to the story of QWERTY as related in David (1985, 1986). That attack has recently been cited by Ruttan (1997), who refers to the emblematic tale of QWERTY as 'the *founding myth* of the path dependence literature' (emphasis added). Although Liebowitz and Margolis fail to substantiate their contention that QWERTY simply is 'a fable,' their rhetorical strategy of attacking that case as though it constituted the only economically interesting exemplar of path dependence, managed to raise a small cloud of doubt regarding the empirical significance of the more general phenomenon. On the latter issue, however, see David (1999) for another view.

[4] Liebowitz and Margolis (1995b: pp. 209-210) fall into just this confusion on the one occasion on which they offer a formal definition of the meaning of 'path dependence'. They say, correctly: 'The meaning closest to current use in economics is that of stochastic processes that incorporate some concept of memory.' But, thereupon they draw from the *Encyclopedic Dictionary of Mathematics* (Cambridge, MA: MIT Press, 1987) the following definition of 'path dependence': Letting P(n) be the probability of event E(n) =A(1) on the n-th trial, and (1-P(n)) be the probability of the mutual exclusive outcome E(n) = A(2), then the general 'response probability' for the sequential process is: $P(n+1)=f\{P(n), E(n), E(n-1),...,E(1)\}$. When the function $f = f\{P(n), E(n), E(n-1),..., E(n-d)\}$, the response probability is said to be 'd-trial path dependent'. In the special case where d=0 it is 'path independent'.

The text in Liebowitz and Margolis (1995b: p.210) then goes on to assert, quite erroneously: 'The use of path dependence in economics is, for the most part, loosely analogous to this mathematical construction: Allocations chosen today exhibit memory; they are conditioned on past decisions.' One should notice that if 'allocations' are associated with 'events', E(i), and (probabilistic) decisions at moment n are characterized by the pairs [P(n); 1- P(n)], then the foregoing statement does not correspond to the mathematical construction of 'd-trial path dependence', any more than the latter corresponds to the generic usage of the concept of path dependence by David (1985, 1986, 1988, 1989, et seq.), or by Arthur (1988, 1989, 1990, 1994), by Cowan (1991, 1996), by Durlauf (1990, 1996), Krugman (1991, 1994), and others contriibuting to the economics literature.

[5] The reference in the passage quoted to 'contingency' as the conceptual counterpart in biology of the idea of path dependence is followed by Liebowitz and Margolis's (1995b: p. 33) statement that 'In *Wonderful Life*, Stephen J. Gould applies this intellectual revolution to paleontology.' But, it should be shiningly clear from that work by Gould (1989: pp. 282ff, esp.), and really no less from his earlier writings, that he is not drawing upon a recent intellectual revolution: 'I regard Charles Darwin as the greatest of all historical scientists. Not only did he develop convincing evidence for evolution as the coordinating principle of life's history, but he also chose as a conscious theme for all his writings...the development of a different but equally rigorous methodology for historical science. Historical explanations take the form of narrative: E, the phenomenon to be explained, arose because D came before, preceded by C, B, and A. If any of these earlier stages had not occurred, or had transpired in a different way, then E would not exist (or would be present in a substantially altered form, E', requiring a different explanation....I am not

speaking of randomness (for E had to arise, as a consequence of A through D), but of the central principle of all history – *contingency*' (Gould, 1989: pp. 282-3). Further on, Gould (1989: pp. 283-4) writes of the universal psychological appeal of the notion of historical contingency, in terms that leave no doubt that this is not a concept specific to evolutionary biology: 'Historical explanations are endlessly fascinating in themselves, in many ways more intriguing to the human psyche than the inexorable consequences of nature's laws....Contingency is the affirmation of control by immediate events over destiny....Contingency is a license to participate in history, and our psyche responds. The theme of contingency, so poorly understood and explored by science, has long been a mainstay of literature....Tolstoy's theme in all his great novels.' What Gould provides in *Wonderful Life* is a new interpretation of the record of life left in the Burgess Shale, but, as he takes pains to acknowledge, this interpretation 'is rooted in contingency' – a very old and far from revolutionary idea.

[6] The practice can be employed with potent rhetorical effect on an unsophisticated audience, because the deterministic reformulation may then be subjected to criticisms from which the original analysis would be immune. A striking instance of such a switch is to be found in Liebowitz and Margolis's (1995b: pp. 214-15) reproduction and critique of a deterministic payoff tableau, used by Arthur (1989) purely *as a heuristic device* – to convey the possibility that a sequence of myopic adoption decisions under increasing returns to adoption could result in the commitment of the ensemble of adopters to a dominated outcome. In the course of pointing out that the payoff tableau may be read in a way that is inconsistent with the results reported for Arthur's stochastic model, there appears the following commendably candid footnote (pp. 214-15, n. 15): 'Actually, Arthur states that this example does not exhibit any 'non-ergodicity', meaning that it is not path dependent in the sense that small differences in historical sequences play a role in the final equilibrium. In this example the end result is the same no matter the order of initial participants. But it illustrates lock-in very well.' I might note that this footnote is the only place I have found in Liebowitz and Margolis's publications on path dependence where the concept is explicitly defined with reference to non-ergodicity, and even so the passage omits explicit reference to probability.

[7] Furthermore, Liebowitz and Margolis (1995b) offer a description of 'third-degree' path dependence that would apply equally to deterministic chaos – which, as was noted above, the authors correctly acknowledged to be not really the same thing as path dependence.

[8] This, however, would seem to leave Liebowitz and Margolis in the position of having to insist that economists should not attach real importance to path dependence because its 'third-degree' form ignores the reality that, even when remediation would not occur via 'private ordering', it would most likely be achievable through 'public ordering'. That is hardly what one expects from defenders of *laissez-faire*.

[9] For discussion of this in the context of technical compatibility standards, see, for example, David and Greenstein (1990b); on social conventions, organizational routines and formal institutions, David (1994c), and David (1994d).

[10] See, for example, David (1987, 1990); David and Bunn (1988), Cowan (1991).

[11] Compare the detailed analyses of the VHS market in Baba and Imai (1990), Cusumano, Mylonadis and Rosenbloom (1992) and Grindley (1992), none among which are noticed in Liebowitz and Margolis (1994), or by the latter authors' subsequent references to this case.

[12] Especially in David (1987), David and Bunn (1988), David and Greenstein (1990) and, most forthrightly in David (1992b).

[13] See David (1993b) for more on the teleological mode of analysis in economics.

[14] For further discussion of the latter topics, see, for example, David (1993a, 1994c), David and Foray (1993d, 1994d).

[15] On these epistemological topics, see, for example, the stochastic models discussed in David and Sanderson (1997), and David (1998b, 2000).