97-10-080

What Should Policymakers Know About Economic Complexity?

Steven N. Durlauf

Department of Economics, University of Wisconsin and Santa Fe Institute

September 13, 1997

This paper was prepared for *The Washington Quarterly*. Financial support from the National Science Foundation and Santa Fe Institute is gratefully acknowledged. I thank Kim Sau Chung, Joshua Epstein, Susan Nelson and Christopher Wheeler for helpful conversations related to this work.

1. Introduction

This essay is written with two goals. The first is to outline the main ideas underlying the growing study of complex economic environments. The second is to suggest areas of public policy where those ideas might be important. Both goals are necessarily speculative. The study of complex systems, whether natural or social, is still in its infancy. While many insights and plausible conjectures have been generated, the long term importance of this work is still unclear.

In formal analyses, complexity denotes something quite different from merely "complicated" or "hard to analyze or solve." For our purposes, a system is said to be complex when it exhibits some type of order as a result of the interactions of many heterogeneous objects. When the interactions occur at a level of description other than that at which the patterns occur, these patterns are often called "emergent." Hence magnetism, which is a property of a large number of iron atoms (in this case with spins polarized in one direction) is an emergent property.

Economists have begun to use complex systems as the basis for formal modeling precisely in order to understand certain aggregate features of environments which are characterized by many heterogeneous actors. Examples are easy to construct. A stock market is comprised of many traders with idiosyncratic beliefs about the future. The actors exchange information and react to some types of common information, but stock prices are ultimately determined by a large number of decentralized buy and sell decisions. In a very different area, the composition of residential communities is determined by the individual preferences and location decisions of individual families.

Notice that in each case, there is a feedback between the aggregate characteristics of the economic environment under analysis and the individual actors which comprise that environment. Movements in stock prices influence the beliefs of individual traders and in turn influence their subsequent decisions. Families make community choices on the basis of expectations concerning community characteristics; these characteristics in turn evolve in response to individual choices. These feedbacks are in and of themselves not unique to complex environments. After all, elementary economics teaches us about the feedback between individual demand and supply decisions and market level prices.

Complexity deepens our understanding of economic phenomena by illustrating how various types of microeconomic structures lead to particular aggregate economic phenomena. For example, as illustrated in work by W. Brian Arthur, John Holland, Blake LeBaron, and Richard Palmer [1] booms and crashes in prices appears to be a frequent feature of stock market environments in which evolving rules-of-thumb behavior interact to determine individual purchasing decisions. Similarly, the emergence of racially segregated neighborhoods from a collection of individuals with different preferences for community racial composition has been illustrated by Thomas Schelling [2], in what is probably the first paper on economic complexity. Hence, the value of complex system thinking in the social sciences, for either researchers or policymakers, lies in its potential for enriching our understanding of the relationships between aggregate outcomes and individual decisions. While plausible theoretical complexity-based models have been developed to explain phenomena ranging from out of wedlock births (William Brock and Steven Durlauf) [3] to the distribution of city sizes (Paul Krugman) [4], these models have not been subjected to sufficient empirical scrutiny to assess their validity. In particular, the causal mechanisms which drive these models have yet to be shown to be empirically salient.

2. Complexity and Economics: Main Ideas

The existing literature on economics and complexity is surveyed in the recent collection of papers edited by Arthur, Durlauf and Lane [5]. Much of this new work has focused on the mathematical formalization of complex systems methods in economic contexts. Underlying this analytical work are several substantive economic assumptions.

i. interactions and positive feedbacks

Complex economic environments typically take as their microeconomic foundations direct interactions between economic actors. What this means is the individual actors are conceptualized as decisionmakers whose choice depend directly on the decisions of others. This conceptualization is quite different from the bulk of economic models in which individuals only interact via market clearing prices. Direct interactions approaches in turn allow economic models to accommodate a far richer class of phenomena than their neoclassical counterparts. For example, conformity effects, in which an individual's perceived benefit from a choice increases in the percentage of his or her friends who make the same choice, is a typical example of such a phenomenon. Such conformity effects are plausibly important in understanding such disparate behavior as the demand for certain types of clothing and the adoption of particular technological standards.

The conformity effects described above are an example of the more general phenomenon of positive feedbacks between economic actors, in which a behavioral choice by one agent makes similar choices on the part of other agents more likely. Another example is role model effects, in which choice by members of one generation influence the choices made by the next generation, an intertemporal interaction which is relevant to issues of schooling and occupation. Indeed, it is no exaggeration to say that many of the efforts to introduce more realistic psychology and sociology into economic theory operationally are efforts to incorporate positive feedbacks of various types.

Systems with positive feedbacks in turn can exhibit multiple types of self-reinforced behavior at an aggregate level. The intuition for this result is straightforward: positive feedbacks means that while there is a tendency for members of a population to make similar decisions, such feedbacks do not say *which* is generally made. The implication of positive feedbacks for aggregate behavior have proven very fruitful for the study of an array of issues. For example, positive feedback effects can explain how two communities with identical distributions of individuals can exhibit very different levels of drug usage, school enrollment or crime rates; these differences can be explained by the aggregate consequence of peer group pressures which induce conformity effects. Alternatively, when the demand for productive inputs between firms is interdependent through supply and demand linkages, multiple levels of aggregate output can occur. Of course, this multiplicity will depend on the inability of individual actors to coordinate their activities.

ii. increasing returns

A second feature of many complex economic environments is the presence of increasing returns to scale. Increasing returns can best be understood in a simple example. Suppose a firm employs 10 workers and 10 machines to produce 100 units of output. Now, ask what level of output the same firm would produce with 20 workers and 20 machines. If the answer to this question is that the firm will produce less than 200 units of output, then the firm exhibits decreasing returns to scale. If the answer is that the firm will produce exactly 200 units, then the firm exhibits constant returns to scale. Finally, if the firm produces more than 200 units, then the firm exhibits increasing returns to scale.

Complex economic environments often exhibit increasing returns of some type. One reason is that increasing returns can be the consequence of positive feedback relationships. For example, two scientists working together can, through the interplay of different perspectives, advance more rapidly than if each works in isolation. It is well known from neoclassical economic theory that increasing returns environments can exhibit various types of inefficiencies; for example, aggregate productivity in the economy may be lower than optimal because individual firms do not internalize the effects of their own research and development investments on the productivity of their competitors [6].

3. Implications

While complex economic environments are typically characterized by positive feedback effects and increasing returns, they also exhibit a typology of interesting aggregate behaviors.

i. evolution versus steady state.

The first important feature of complex economic systems is that they usually must be understood as evolving processes. This feature contrasts with the more standard approach of looking at equilibrium or steady state behavior of the system. Notice that by a steady state, we do not necessarily mean that the system is at rest; rather we mean that the system is in a state where its behavior is characterized by dynamics which are not themselves changing across time. When a social, economic or political system is explicitly conceptualized as a dynamic process involving the interrelationships of many actors who possess limited information about the intentions and objectives of others, it is clear that the behavior of such a system cannot be understood exclusively by asking whether it has any steady states. First, it is unclear that such steady states exist. Second, even if any steady states exist, the presence of positive feedbacks can mean that there exist several such states, so a theory of adjustment is needed to understand which steady state will actually come to pass as the system evolves.

The implication that complex economic environments require explicit consideration of dynamics is an important lesson and one which in some sense is robust to other implications of such systems. Natural complex systems exhibit many specific types of aggregate pattern formation, often referred to as self-organization. It is far from clear that such precise aggregate patterns should be expected to occur in social contexts. Natural systems exhibit various sorts of homogeneity assumptions with respect to the individual elements of the system and various symmetry assumptions with respect to the interaction structure by which the elements are collected. Such assumptions have little justification in social environments. (And claims by natural scientists that social systems exhibit such patterns have generally been grossly exaggerated.) However, a dynamics-based perspective is still crucial in understanding many issues of importance, be it the emergence of ghettos or the emergence of stock market crashes.

ii. nonlinearity.

A critical feature of any set of economic theories is the way in which they are translated into empirical statements. Complex systems are inherently nonlinear in terms of the way in which features of the system move together. Thus, the way in which empirical inferences are drawn about complex environments is intrinsically different from conventional approaches which, as any user of regression analysis knows, are highly linear. For example, suppose that based on recent history of one and two percentage point fluctuations in per pupil expenditures, one found that, on average, every one percentage increase in per pupil expenditure had no effect on high SAT scores among high school seniors. Would it be reasonable to conclude that a 20% percent increase in per pupil expenditures would also have no effect? The answer clearly depends on whether the observable historical experience can be extrapolated, which in turn requires that the relationship between changes in per student expenditure and SAT scores is linear.

An important lesson from complex economic models is that nonlinearities are endemic when positive feedback effects are present. To see how nonlinearity follows from the existence of direct interaction effects, consider the question of how a percentage change in the cost of college will influence a family's decision to pay for a child's attendance. While the example requires a somewhat elaborate formal representation [7], the basic intuition is straightforward.

For an individual family, the decision to pay for college is a zero/one decision. Therefore, for each individual family, the direct effect of a change will be discontinuous-for each family sending someone to college, the effect of each percentage change on the decision is either 0 or 100%. When different families have different thresholds in terms of the maximum tuition they are able or willing to pay, the possibility exists that in the aggregate there will exist a linear relationship between the cost of college and aggregate attendance. Now, ask what would happen if the decision of a family to pay for college exhibits conformity effects, in that the willingness of parents to pay is a function of whether other parents are making the same decision. In this case, the potential exists for a nonlinear effect even in the aggregate. Why? Because the interdependence in decisions means that the there is a feedback from parents who choose not to send their children given the current level of college enrollment, which reduces the enrollment, to a further reduction induced by parents changing their minds because the percentage has now gone down.

Does the presence of nonlinearity mean that one cannot identify the effects of large policy changes? Not necessarily. Rather, the correct lesson is that the identification of such effects requires that empirical analysis be conducted in the context of a theoretical framework which allows for the presence of interaction effects and hence nonlinearities. For example, it is in principle possible to identify the magnitude of the conformity effect on college decisions and infer from that the effects of a large change in college costs.

iv. path dependence.

A hallmark of complex systems is that "history matters," in their long run evolution. Intuitively, this means that very long run behavior of an economic environment is influenced by short run factors. This idea is encapsulated under the term "path dependence" which can mean either 1) that particular innovations in the economy have permanent consequences, or 2) that particular innovations are not self-correcting, so that they remain permanent in absence of some countervailing innovation. Notice that path dependence is really the dynamic version of positive feedback effects. When behavior at a one period in time is sufficiently likely to induce similar behavior in the future, a system can exhibit multiple types of long run behavior, each determined by the particular shocks which occur early in the history of the process.

Path dependence refers to both those economic environments in which the history of shocks have long run consequences as well as those in which the shocks are merely persistent. The important feature in path dependent economies is that the certain shocks to the system will not work their way out of the system due to some intrinsic dynamic within the system. However, they may dissipate due to future shocks. This feature is particularly useful in thinking about models which can generate phenomena similar to the Great Depression in the United States, where the output declines of 1929-1932 do not seem to have been self-correcting, yet the New Deal and Second World War were able to shift the economy away from sustained low-production. Hence while, as an historical matter, the Depression was persistent rather than permanent, it nevertheless is defensible that the pre-World War II economy exhibited path dependence.

3. Examples

i. high technology

A consensus has developed among scholars that a number of forms of increasing returns are fundamental to the structure of high technology industries [8]. Several reasons explain this consensus. First, high technology firms require, due to the interdependence of the demand for their products with the state of other products in the industry, continuous research and development so as to ensure that their products are competitive. Research and development is an activity that is subject to strong increasing returns as ideas and innovations are more likely generated by large and diverse workforces than in small ones.

Second, there is the presence of fixed costs in production. In many production contexts, a firm or company exhibits increasing returns because some part of its productive inputs must be used to begin production which is independent of the scale of activity. This creates a threshold effect in production as nothing can be produced unless these fixed requirements are met.

Why is the presence of increasing returns in production important from the perspective of policy? One reason is that increasing returns make it especially likely that high technology sectors exhibit noncompetitive market conditions. Under increasing returns, the average cost per unit of production decreases with the level of production. Hence the cost advantages will accrue to those firms which are able to produce at high levels. Further, in the presence of substantial fixed costs to production, there is a minimum scale of production which is required for market entry. Taken together, increasing returns create conditions under which a single or small number of firms can dominate a market.

Increasing returns to scale in the production of goods or in the development of new technologies is not the only reason to suppose that high technology industries are uniquely susceptible to noncompetitive market behavior. A distinct issue which is important in computer-based activities has to due with lock-in of demand.

Demand lock-in can easily be seen in many contexts. For example, consider the demand for betamax versus VHS videotapes. Once a consumer has purchased a VCR which can only use one of the two types of videotapes, that consumer is, unless willing to buy a new machine, locked into using that type of videotape. This feature is, of course, commonplace to many goods. What is important in this case is the following. The desirability of a particular videotape standard will depend on the level of use of the same standard by others. One reason for this is that if one's friends use the same standard, it is then possible to share tapes. If such effects are strong enough, then independent of any increasing returns in the production of videotapes there will be forces which lead the economy to choose a single standard. Second, and more realistically, suppose that access to one type of videotape is facilitated by the level to which that type is used. An example of this interdependence is the percentage of videotapes of one type which are available in videostores. Once again, demand for a type of videotape machine therefore type of videotape will again depend on the level of demand by others.

Computer operating systems are an example of how lock-in is especially likely in high technology contexts. The desirability of a particular operating system by a given consumer depends, to a very large extent, on the availability of products which are compatible with that platform. This availability depends, in turn, on the size of the market for products with that platform. Hence, the demand for computer operating systems by consumers exhibits strong interdependences--interdependences which are very different from the textbook model of consumer demand in which interdependences among consumers only occur because of their collective participation in a market.

Interdependence in demand for a good can, through the lock-in effects associated with the need to commit to one type of standard or technology, create the potential for one firm to dominate an industry due to the scale of the use of its products. The standard argument that superior products will drive inferior products from the market does not necessarily apply in this case. In the presence of lock-in, it is difficult for a new market entrant to generate a level of demand such that per unit costs become low enough to make market entry profitable. This is historically what occurred in the personal computer industry, where it is reasonable to conclude that the installed user base of MS-DOS/Windows and Apple operating systems has rendered it highly unlikely that a new, superior operating system could generate sufficient profits to compete. The bottom line is that due to increasing returns and lock-in, market equilibria can be highly inefficient, either due to the emergence of market power for some firms, or the mutually reinforcing adoption of inferior technologies.

Finally, note that technology development, like the production of knowledge in general, presumably involves many positive feedbacks which are not compensated. A technological breakthrough in one area may stimulate breakthroughs in others without any compensation to the initiator. This is a textbook case of production externalities; what complexity theory suggests is that connections between different technologies-may make these externalities very large. Inefficiencies due to the divergence between the private and social return to research and development again suggest that at least in principle some government intervention may be desirable.

ii. inequality

Interactions-based models have been used to understand the emergence of persistent inequality and poverty. The basic ideas of these models may summarized as follows. "First, individual preferences, beliefs, and opportunities are strongly influenced by one's memberships in various groups. Such groups may be fixed, such as race, or may be determined by the economy or society, such as neighborhoods, schools, or firms. Second, positive interaction effects occur between members of a given group, so that group level influences generate common outcomes among group members. Third, greater societal stratification by income, race, education, or language leads to divergence in group characteristics which results in greater cross-section inequality and decreased social mobility" [9]. What this perspective means is that the relationship between between an individual's characteristics and his economic outcomes is mediated by the characteristics of members of those groups to which he is a member. Some of these groups are immutable, such as ethnicity. Others, such as community membership or even the membership in labor force of a particular firm are determined within the operation of the economy. While the way in which groups form and the way in which intragroup interactions occur will differ greatly in different cases, the implications for inequality can be quite robust.

The main context in which these abstract ideas have been applied is in the evolution of community formation and education [10]. This strand of the inequality literature is driven by the belief that an adult's occupation, and hence economic status, is determined by the interaction of three factors: the level of human capital investment received as a child through formal education, a range of role model and peer group effects which influence aspirations, efforts, and opportunities, and luck. When human capital is created as a local public good (which of course occurs to a great extent due to local finance of schools), this implies a direct relationship between the income distribution of a community and the education received by its children. A similar relationship between community economic characteristics and adult outcomes is induced when the occupational distribution of a community determines a youth's role models or the number and type of contacts a community member has in the labor force. This two-tier mapping of a community's economic characteristics into the economic prospects of its children represents an intergenerational form of the positive feedback effects which typically characterize complex systems.

Given these feedbacks from a community's income distribution to the future occupations and incomes of its children, families will generally want to live in those communities with the most affluent income distributions. Hence, strong incentives exist for economic segregation of communities. The extent of this segregation will of course depend upon factors such as mobility costs, preferences for various neighborhood amenities, and the level of fixed costs in education, costs which can vary from athletic facilities to computers. As a result, the stochastic process describing the evolution of communities will exhibit quite complicated dynamics.

One important feature of this theory is the possibility that some family dynasties can decline into persistent or permanent poverty while others are always nonpoor. As wealthier families are segregated from poorer families through some combination of zoning restrictions and housing price differentials, rich and poor families become isolated. This isolation means that their children are exposed to very different schools and labor market opportunities. Hence initial differences in income between adults can become magnified between their children. Ghettos or other types of poverty traps may emerge and perpetuate themselves as a consequence of this endogenous stratification. Across generations, family income can therefore exhibit path dependence when the economic status of parents has a sufficiently strong influence on offspring, due to its role in determining community-level influences. Notice that the importance of group memberships in the determination of economic success suggests that policies designed to promote equality should focus on this factor. In fact, policies such as affirmative action may be defended because the presence of strong positive spillover effects induced by the allocation of jobs across ethnic groups [11].

iii. national security

Complex systems methods have had relatively little application in security issues [12]. However, it is clear that the ideas and metaphors which motivate economic applications can apply here as well. One possible question is the probability of a nuclear weapons accident. As made clear by Scott Sagan [13], understanding the probability of an accidental launch requires understanding the behavioral outcomes of an organization (military of course) comprised of many decentralized, yet highly interdependent decisions. Indeed, a major argument in defense of so-called "normal accidents" theory (so dubbed by Sagan), at least as developed by Charles Perrow, is that there exist sufficiently many nonlinear interactions between elements of large organizations that mistakes will invariably arise which cannot be accommodated by safety features which can only accommodate foreseeable contingencies [13]. Gene Rochlin has further argued that the degree of computerization of defense capabilities determines the degree of interdependence within a defense organization, so this general concern about organizations seems especially applicable to defense [14].

Could complex system methods help clarify the probability of a nuclear weapons-related accident? It is certainly plausible to believe that the answer is yes. Formal modeling of command and control systems could be achieved with a great deal of accurate microlevel detail. I am willing to conjecture two features of such an exercise. First, it will be possible to produce scenarios under which accidents occur at unacceptably high frequencies. One message of complexity based studies of stock price movements generated by interdependent traders or the evolution of social

pathologies is that extreme outcomes have nontrivial probability due to positive feedback effects; hence a similar result in the context of nuclear accidents seems reasonable. Second, it may nevertheless be possible to design redundancies and safety mechanisms within the system to render this probability negligible. Why? Because the same interdependence which may make accidents seem relatively likely also induces sufficient order in the system to allow for small but common influences on individual decisions to have large aggregate effects. By making each actor in the system slightly more cautious, the feedback effects will render the system as a whole much more cautious. How this can be done naturally requires expertise in the details of the organization of interest, but the capacity for large heterogeneous systems to experience collective order due to positive interactions suggests that this goal can in principle be accomplished [15].

4. Conclusions

Complex economic systems have several messages useful to policymakers. First, interdependence among various actors can create multiple types of internally consistent aggregate behavior. As a result, economic environments can become stuck in undesirable steady states. Such undesirable steady states may include high levels of social pathologies or inferior technology choices. Second, the consequences of policies will depend critically on the nature of the interdependences. In particular, the effects of different policies may be highly nonlinear, rendering history a poor guide to evaluating policy effectiveness. Hence, even though many complex economic environments seem to be particularly likely situations in which there exist social welfare enhancing policies, it may prove especially difficult to identify such policies. At a minimum, detailed empirical studies which underlie conventional policy analysis should prove to be even more valuable in complex environments.

Notes

1. George Cowan, David Pines, and David Meltzer, eds., Complexity: Metaphors, Models and Reality, Menlo Park: Addison-Wesley, 1994 provides a fascinating overview of the state of complexity research.

2. W. Brian Arthur, John Holland, Blake LeBaron, and Richard Palmer, The Economy as an Evolving Complex System II, W. Brian Arthur, Steven Durlauf and David Lane, eds., Menlo Park: Addison-Wesley, 1997.

3. Endogenous growth theory, which has dominated macroeconomic research for the last decade, is based on understanding how these positive types of productivity feedbacks can generate sustained growth. The seminal article is Paul Romer, "Increasing Returns and Long Run Growth," Journal of Political Economy, 94, 1002-1037, 1986.

4. Thomas Schelling, "Dynamic Models of Segregation, " Journal of Mathematical Sociology, 1, 143-186, 1971.

5. William Brock and Steven Durlauf, "Discrete Choice with Social Interactions I: Theory," Working Paper, Department of Economics, University of Wisconsin, December 1995.

6. Paul Krugman, The Self-Organizing Economy, Oxford: Basil Blackwell, 1996. This book is also a delightful introduction to the application of complexity to economics.

7. W. Brian Arthur, Steven Durlauf, and David Lane, eds., The Economy as an Evolving Complex System II, op. cit.

8. W. Brian Arthur is the originator of many of the important ideas concerning the implications of complexity for technological evolution. See especially his "Competing Technologies, Increasing Returns, and Lock-In By Small Historical Events," Economic Journal, 99, 116-131, 1989.

9. Steven Durlauf, "The Memberships Theory of Inequality: Ideas and Implications," Working Paper, Department of Economics, University of Wisconsin, 1997.

10. See Roland Benabou, "Workings of a City: Location, Education, and Production, " Quarterly Journal of Economics, CVIII, 619-652, 1993, "Equity and Efficiency in Human Capital Investment: The Local Connection," Review of Economic Studies, 62, 237-264 and Steven Durlauf "A Theory of Persistent Income Inequality," Journal of Economic Growth, 1, 75-93, 1996 and "Neighborhood Feedbacks, Endogenous Stratification, and Income Inequality," in Dynamic Disequilibrium Modeling, W. Barnett, G. Gandolfo, and C. Hillinger, eds., Cambridge University Press for examples of this work.

11. Steven Durlauf, "The Memberships Theory of Inequality: Ideas and Implications," Working Paper, Department of Economics, University of Wisconsin, 1997.

12. The few efforts I have seen have been too mechanical (in that they relabel variables in physical systems with social science names, without accounting for substantive behavioral differences) to represent serious substantive contributions. This is not to say that all mathematical security or conflict models have this shortcoming-see Joshua Epstein, Nonlinear Dynamics, Mathematical Biology, and Social Science, Menlo Park: Addison-Wesley, 1997 for a nice introduction to successful examples of the applications of mathematics to issues in this area.

13. Scott Sagan, The Limits of Safety, Princeton: Princeton University Press, 1993.

14. Charles Perrow, Normal Accidents: Living with High Risk Technologies, New York: Books, 1984.

15. Gene Rochin, Trapped In the Net: The Unanticipated Consequences of Computerization, Princeton: Princeton University Press, 1997.

16. One might argue that normal accidents theories are based on the claims that certain contingencies cannot be foreseen rather than on claims about the complexity of organizations per se. However, if the concern about complex environments is that there are contingencies whose nature we cannot characterize, let alone whose probabilities we cannot evaluate, then it is incoherent to talk about the probability of accidents being high or low. Also, the issue for system performance is not the identifiability of the range of possible shocks to the system, but rather the identification of the range of responses and interconnections. If the argument that complex organizations are likely to produce errors is to make sense, it is presumably a statement about how the interactions in such systems evolve, which is why complexity seems a natural language for analysis.