

WHERE HAVE WE BEEN AND WHERE ARE WE GOING?

Douglass C. North*

Washington University, St. Louis

I

Improving our understanding of the nature of economic change entails that we draw on the only laboratory that we have--the past. But "understanding" the past entails imposing order on the myriad facts that have survived to explain what has happened--that is theory. The theories we develop to understand where we have been come from the social sciences. Therefore there is a constant give and take between the theories we develop, and their application to explain the past. Do they improve our understanding--is the resultant explanation broadly consistent with the surviving historical evidence?

But can the gradual accretion of "sound" explanations of the past help in understanding where we are going? That depends on the degree to which there are lessons from history. If there are lessons they are not the kind that are the bread and butter of politicians, statesmen and soothsayers. Rather they are of two kinds: 1) the deep underlying interrelationships between the rules of the game that humans devise to structure human interaction and the way those rules evolve in the interaction between humans and their ever changing environment--an environment that is changing not only as a result of external natural forces but also in consequence of the changes induced by the players themselves. Which brings us to 2) the significance of path dependence, properly understood, in shaping the present and the future.

*. I am indebted to Elisabeth Case for editing this essay.

To begin we need to assess what we have learned from the past and then assess the usefulness of the tools at hand--ie the rationality assumption and growth theory we employ in economics? We will then go on to explore in subsequent sections some recent development that offer the promise of improving our understanding of the past and of where we are going.

II

The overall pattern of evolution involves 3 stages: (1) The physical and chemical evolution of the stars and planets, (2) biological evolution starting with DNA and development of living species, and (3) societal evolution starting with homo sapiens.(see Boulding, Journal of Evolutionary Economics, Vol I, #1, pp9-17). While stage one is a given in our analysis stage two plays an important role since the evolving genetic structure of humans--particularly the way the mind has evolved and the development of language play a major role in the third stage of societal evolution.

The first great discontinuity in societal evolution was surely the development of agriculture which took place almost four million years after human beings became separate from other primates--that is, only ten thousand years ago. It probably evolved independently in different parts of the world at different times and may have resulted from some combination of constriction of the environment and local population pressure (North and Thomas, 1979?). Whatever the source, agriculture created civilization both by expanding the productive potential of humans and establishing the sedentary conditions that made possible the growth of specialization and division of labor.

In the ensuing ten thousand years the rate of change appears to have been very slow for at least the first half of that period although the surviving archaeological evidence has been scarce. In the second, it has been episodic for most of the time; periods of what appear to have been economic growth in

particular geographic regions have been interrupted by stagnation and decline and sometimes those geographic areas have failed to recover. Athens in the fifth century B.C., Rhodes in the third century B.C. and Rome in the early Empire were examples of growth followed by decline and failure to recover for centuries. Following the demise of Rome in the west there was a long hiatus until the beginning of revival in the tenth century. With Mohammed came expansion of the Muslim world in North Africa and beyond. And in the far east Chinese dynasties produced enormously rich cultures but it remained for the relatively backward area of western Europe to be the incubator of modern growth.

If modern economic growth had its genesis in the tenth century revival of western Europe sustained growth appears to have begun perhaps four hundred years ago but confined to a small part of the earth for most of that time--western Europe and the overseas settlements of England. Widespread growth is a recent phenomenon mostly dating since World War two. Even today large parts of the world are not experiencing growth (eg, the republics of the ex Soviet Union and Sub Saharan Africa), or growth continues to be episodic (eg, Latin America).

III

The rational choice paradigm assumes that people know what is in their self interest and act accordingly, or at the very least that competition will weed out those who make incorrect choices and reward those who make correct choices. But it is impossible to reconcile this argument with the historical and contemporary record.

Growth theory as it has evolved from neo-classical theory is equally unhelpful in explaining this historical and contemporary record. Convergence of the growth paths of economies, such as Baumol(1986) maintained, has only tended to occur amongst the developed countries. Persistent divergence as argued among some of the new growth theory variants of the Solow model can't explain

the rise of the Asian tigers or China. In fact, to put it bluntly, the growth theory stemming from neo-classical economics, old or new, suggests not only ignorance of the empirical evidence, historical or contemporary, but a failure to recognize that incentives matter; surely a remarkable position for economists whose theory is built around incentives.

It has to be the incentive structure imbedded in the institutional/organizational structure of economies that is a key to unraveling the puzzle of uneven and erratic growth. But that entails a still deeper puzzle. Why don't economies that have institutional frameworks that are inhospitable to economic growth simply adopt the frameworks of the successful economies? They do, or at least they try to: the rush to create market economies is a ubiquitous characteristic of third world and transition economies. But look at the results. They vary enormously, from China and the Czech republic, which so far are successful; to the republics of the ex Soviet Union, which so far show few signs of success; to sub saharan Africa, which remains a basket case.

To make sense out of the historical and contemporary evidence, we must rethink the whole process of economic change. Current theory stems from the development of national income and growth accounting literature and explores the superficial aspects of economic growth--technology or human or physical capital--rather than the structure of incentives and disincentives that make up the institutional framework of an economy and polity. If we enquire why these incentives vary we are driven to the various belief systems that determine the institutional framework.

We are far from being able to understand the whole process but I do intend to suggest some explanation which offer the promise of dramatically improving our understanding of where we have been. Specifically I shall assert and support three arguments that are at wide variance with the received wisdom. The first is

that contrary to both the economic history literature and the economic growth literature--old and new--the primary source of economic growth is the institutional/organizational structure of a political-economy and until we focus on that subject we shall not advance knowledge on economic growth. Second that economic growth is dependent on stable political/economic institutions that will provide low costs of transacting in impersonal political and economic markets. Third, and central to the issues of this conference, that it is the belief systems of societies and the way they evolve that is the underlying determinant of institutions and their evolution. In the rest of this essay I shall support these arguments.

IV

Curiously enough institutions are front and center in all current explanations of growth or lack of it in third world or transition economies; but the explanations lack both analytical content and an understanding of the nature of institutions or the way they evolve. The implication is that they can be created at will or frequently that they are a dependent variable to getting the prices right. But recent experience provides convincing evidence that neither can they be taken for granted nor do they automatically evolve from getting the prices right. And historical experience makes clear that efficient economic institutions are the exception. To proceed we must understand what institutions are and how they evolve.

Institutions provide the structure that humans impose on human interaction in order to reduce uncertainty. There is nothing about that structure that implies that institutions are efficient in the sense of inducing economic growth. Sometimes they are created to facilitate exchange, encourage technological change, and induce human capital formation and in consequence reduce

transaction and/or transformation costs; at other times they are created to support monopoly, prevent technological change, thwart human capital development, and generally raise the costs of transacting and/or transformation. In fact the second of these has been the far more common pattern throughout history. To provide an understanding of why, I shall state five propositions that, I believe, underlie institutional change; but first let me dispel any confusion between institutions and organizations. Institutions are the rules of the game--both formal rules and informal constraints (conventions, norms of behavior and self-imposed codes of conduct)-- and their enforcement characteristics. Together they define the way the game is played.

Organizations are the players. They are made up of groups of individuals held together by some common objectives. Economic organizations are firms, trade unions, cooperatives, etc.; political organizations are political parties, legislatures, regulatory bodies; educational organizations are universities, schools, vocational training centers. The immediate objective of organizations may be profit maximizing (for firms) or improving reelection prospects (for political parties); but the ultimate objective is survival because all organizations live in a world of scarcity and hence competition.

Now to the five propositions:

1. The continuous interaction between institutions and organizations in the economic setting of scarcity and hence competition is the key to institutional change.

2. Competition forces organizations continually to invest in new skills and knowledge to survive. The kind of skills and knowledge individuals and their organizations acquire will shape evolving perceptions about opportunities and hence choices that will incrementally alter institutions.

3. The institutional framework provides the incentive structure that dictates the kinds of skills and knowledge perceived to have the maximum payoff.

4 Perceptions are derived from the mental constructs of the players.

5. The economies of scope, complementarities, and network externalities of an institutional matrix make institutional change overwhelmingly incremental and path dependent.

Let me expand on propositions 2 through 5.

2. New or altered opportunities may be perceived to be a result of exogenous changes in the external environment which alter relative prices to organizations or a consequence of endogenous competition among the organizations of the polity and the economy. In either case the ubiquity of competition in the overall economic setting of scarcity induces entrepreneurs and the members of their organizations to invest in skills and knowledge. Whether through learning by doing on the job or the acquisition of formal knowledge, improving the efficiency of the organization relative to that of rivals is the key to survival.

While idle curiosity surely is an innate source of acquiring knowledge among human beings, the rate of accumulating knowledge is clearly tied to the pay-offs.

Secure monopolies, be they organizations in the polity or in the economy, simply do not have to improve to survive. But firms, political parties, or even institutions of higher learning faced with rival organizations must strive to improve their efficiency. When competition is muted (for whatever reasons) organizations will have less incentive to invest in new knowledge and in consequence will not induce rapid institutional change. Stable institutional structures will be the result. Vigorous organizational competition will accelerate the process of institutional change.

3. There is no implication in proposition 2 of evolutionary progress or economic growth--only of change. The institutional matrix defines the opportunity set, be it one that makes income redistribution the highest pay-off in an economy or one that provides the highest pay-offs to productive activity. While every economy provides a mixed set of incentives for both types of activity, the relative weights (as between redistributive and productive incentives) are crucial factors in the performance of economies. The organizations that come into existence will reflect the pay-off structure. More than that, the direction of their investment in skills and knowledge will equally reflect the underlying incentive structure. If the highest rate of return in an economy comes from piracy we can expect that the organizations will invest in skills and knowledge that will make them better pirates. Similarly if there are high returns to productive activities we will expect organizations to devote resources to investing in skill and knowledge that will increase productivity.

The immediate investment of economic organizations in vocational and on the job training obviously will depend on the perceived benefits; but an even more fundamental influence on the future of the economy is the extent to which societies will invest in formal education, schooling, the dissemination of knowledge, and both applied and pure research which will mirror the perceptions of the entrepreneurs of political and economic organizations.

4. The key to the choices that individuals make is their perceptions, which are a function of the way the mind interprets the information it receives. The mental constructs individuals form to explain and interpret the world around them are partly a result of the genetic evolution of the mind, partly of their cultural heritage, partly a result of the local everyday problems they confront and must

solve, and partly a result of non-local learning.¹ The mix among these sources in interpreting one's environment obviously varies as between for example a Papuan tribesman on the one hand and an economist in the United States on the other (although there is no implication that the latter's perceptions are independent of his or her cultural heritage).

The implication of the foregoing paragraph is that individuals from different backgrounds will interpret the same evidence differently; they may, in consequence, make different choices. If the information feedback of the consequences of choices were complete then individuals with the same utility function would gradually correct their perceptions and over time converge to a common equilibrium; but as Frank Hahn has succinctly put it, "There is a continuum of theories that agents can hold and act upon without ever encountering events which lead them to change their theories." (Hahn, 1987, p. 324) The result is that multiple equilibria are possible due to different choices by agents with identical tastes.

5. The viability, profitability, and indeed survival of the organizations of a society typically depend on the existing institutional matrix. That institutional structure has brought them into existence; and their complex web of interdependent contracts and other relationships has been constructed on it. Two implications follow. Institutional change is typically incremental and is path dependent.

Why can't economies reverse their direction overnight? This would surely be a puzzle in a world that operates as neo-classical theory would have us believe. That is, in a neo-classical world abrupt, radical change should immediately result

¹. This argument is elaborated in a cognitive science approach to the development of learning and belief systems in Denzau and North, "Shared Mental Models: Ideologies and Institutions", *Kyklos* (1994).

from a radical change in relative prices or performance. Now it is true that on occasion accumulated pressures do produce an abrupt change in institutions akin to the punctuated equilibrium models in evolutionary theory. But it is simply a fact that the overwhelming majority of change is incremental and gradual. It is incremental because large scale change would harm large numbers of existing organizations and therefore is stoutly opposed by them. Revolutionary change will only occur in the case of "gridlock" among competing organizations, which thwarts their ability to capture gains from trade. Revolutions are extraordinary and even when they occur turn out over time to be far less revolutionary than their initial rhetoric would suggest.

Path dependence could mean nothing more than that yesterday's choices are the initial starting point for today's. But path dependence appears to be a much more fundamental determinant of long run change than that.² The difficulty of fundamentally altering paths is evident and suggests that the learning process by which we arrive at today's institutions constrains future choices. The institutional structure builds in a set of constraints with respect to downstream changes that biases choices.

V

Historical success stories of institutional adaptation reflect conditions where the belief system underlying an institutional matrix filters the information the players got from current experiences and interpreted that information in ways that induced choices that led to the modification, alteration, or adoption of institutions that resolved existing problems or led to improvements in competitive performance.

². The concept of path dependence has been pioneered by Arthur (1989) and David (1985).

Take the successful development of the Netherlands and England in early modern Europe.³ It was the lack of large scale political and economic order that created the essential environment hospitable to political/economic development. In that competitive, decentralized environment lots of alternatives were pursued as each society confronted its own unique external environment. Some worked, as in the cases of the Netherlands and England; some failed, as in the cases of Spain and Portugal; and some, such as France, fell in between these two extremes. But the key to the story is the variety of options pursued and the likelihood (as compared to a single unified policy) that some would turn out to produce political/economic development. Even the relative failures in Western Europe played an essential role in European development and were more successful than other parts of the world because of competitive pressures.

The last point deserves special emphasis. It was the dynamic consequences of the competition amongst fragmented political bodies that resulted in an especially creative environment. Europe was politically fragmented; but it was integrated in having both a common belief structure derived from Christendom, and information and transportation connections that resulted in scientific, technological, and artistic developments in one part spreading rapidly throughout Europe. To treat the Netherlands and England as success stories in isolation from the stimulus received from the rest of Europe (and to a lesser degree Islam and China) is to miss a vital part of the explanation. Italian city states, Portugal, and Germanic states all fell behind the Netherlands and England; but banking, artistic development, improvements in navigation, and

³. This section on the development of early modern Europe is elaborated at length in North, "The Paradox of the West" in Davis, ed (1995)

printing were just a few of the obvious contributions that the former states made to European advancement.

The Netherlands and England pursued different paths to political/economic success but in each case the experiences were conducive to the evolution of a belief structure that induced political and economic institutions that lowered transaction costs. In both polities competition among the evolving nation states was a deep underlying source of change and equally a constraint on the options available to rulers within states. It was competition that forced the Crown to trade rights and privileges for revenue including most fundamentally the granting to representative bodies--variously Parliament, States General, Cortes--control over tax rates and/or certain privileges in return for revenue. But it was the evolving bargaining strength of rulers vis-a-vis constituents that was the decisive feature of their subsequent development. Three considerations were at stake: 1) the size of the potential gains the constituents could realize by the state taking over protection of property; 2) the closeness of substitutes for the existing ruler--that is the ability of rivals (both within and outside the political unit) to the existing ruler to take over and provide the same, or more, services; 3) the structure of the economy which determined the benefits and costs to the ruler of various sources of revenue.

Let me briefly describe the background conditions of the two polities that led up to the contrasting external environments that shaped the belief systems.

To understand the success of the Netherlands one must cast a backward glance at the evolution of the prosperous towns of the Low Countries such as Bruges, Ghent, and Liege; their internal conflicts; and their relationship to Burgundian and Habsburg rule. The prosperity of the towns, whether based on the wool cloth trade or metals trade, early on made for an urban-centered, market-oriented area unique at a time of overwhelmingly rural societies. Their internal

conflicts reflected ongoing tensions between patrician and crafts and persistent conflicts over ongoing efforts to create local monopolies which, when successful, led to a drying up of the very sources of productivity which had been the mainspring of their growth. Burgundian (and later Habsburg) rule discouraged restrictive practices such as those that developed in the cloth towns of Bruges and Ghent and encouraged the growth of new centers of industry that sprang up in response to the favorable incentives embodied in the rules and property rights. In 1463 Philip the Good created a representative body, the States General, which enacted laws and had the authority to vote taxes for the ruler. The Burgundians and Habsburgs were rewarded by a level of prosperity that generated tax revenues that made the Low Countries the jewel in the Habsburg Empire.

England evolved along a route different from that of continental polities. Being an island made it less vulnerable to conquest and eliminated the need for a standing army (and undoubtedly contributed to the different initial belief structure that Macfarlane describes in The Origins of English Individualism). The Norman conquest, the exception to British invulnerability to external conquest, produced a more centralized feudal structure than that on the continent. The political institutions, in consequence, differed in several important respects from those of the continent. There was a single parliament for the entire country; no regional estates as in France, Spain, and the Netherlands. There were also no divisions into towns, clergy, and nobility. But the more centralized feudal structure did not ensure that the Crown could not overstep the traditional liberties of the barons, as the Magna Carta attests.

We can now turn to examining the evolving bargaining strength (and the three underlying determinants) of ruler versus constituent that shaped the belief structure and the path of each polity. Take the Netherlands. The productive town economies stood to gain substantially by the political order and protection of

property rights provided by the Burgundians and then by Charles V. The structure of the economy built around export trades provided the means for easy-to-collect taxes on trade but not at a level to adversely affect the comparative advantage of those export trades. The liberty to come and go, buy and sell as they saw fit led to the evolution of efficient economic markets. But when Philip II altered the "contractual agreement" the Seven Provinces became convinced that they could prosper only with independence. The resistance was initiated by the States General which in 1581 issued the Act of Abjuration of allegiance to Philip II and claimed sovereignty for the Provinces themselves. The powers of the newly independent country resided with each province (which voted as a unit) and a unanimity rule meant that the States General could only act with the unanimous approval of the Seven Provinces. Cumbersome as that process was, this political structure survived. The polity not only evolved the elements of political representation and democratic decision rules but equally supported religious toleration. The belief structure that had evolved to shape the independent polity was more pragmatic than intellectual, a consequence of the incremental evolution of the bargaining strength of constituents and rulers.

As with the Netherlands, it was England's external trade that provided an increasing share of Crown revenue with taxes on wine, general merchandise, and wool cloth; but it was the wool export trade that was the backbone of augmented crown revenue. Eileen Power's classic story of the wool trade (1941) describes the exchange between the three groups involved in that trade: the wool growers as represented in Parliament, the merchants of the staple, and the Crown. The merchants achieved a monopoly of the export trade and a depot in Calais; Parliament received the right to set the tax; and the Crown received the revenue. Stubbs (1896, 3:599) summarized the exchange as follows: "The admission of the right of parliament to legislate, to enquire into abuses, and to share in the guidance

of national policy, was practically purchased by the money granted to Edward I and Edward III..."

With the Tudors the English Crown was at the zenith of its power, but it never achieved the unilateral control over taxing power that the Crowns of France and Spain achieved. The confiscation of monastery lands and possessions by Henry VIII alienated many peers and much of the clergy and as a consequence "Henry had need of the House of Commons and he cultivated it with sedulous care" (Elton, 1953, 4). The Stuarts inherited what the Tudors had sown and the evolving controversy between the Crown and Parliament is a well known tale. Two aspects of this controversy are noteworthy for this analysis. One was the evolving perception of the common law as the supreme law of the land--a position notably championed by Sir Edward Coke--and the other was the connection made between monopoly and a denial of liberty as embodied in the Crown grants of monopoly privileges.

VI

England and the Netherlands represent historical success stories in which an evolving belief systems shaped by external events induced institutional evolution to provide the beginning of modern economic growth. Key institutional/organizational changes were those that permitted the growth of impersonal exchange--both economic and political. By permitted I mean that, to use game theory terminology, they altered the payoff between defection and cooperation to favor the latter in both political and economic markets. That is, personal exchange provides settings in which it typically pays to cooperate. Impersonal exchange is just the antithesis and necessitates the development of institutions to alter pay-offs in favor of cooperative activity.

In the case of economic markets recent historical research has provided analytical accounts of the evolution of the institutions that undergirded long

distance trade in the Middle Ages (Greif,1993); that led to the development of merchant codes of conduct that became the foundation of commercial law (Milgrom, North, and Weingast, 1990); that converted uncertainty to risk, leading to the development of marine insurance (deRoover, 1945); and that provided the foundation of an impersonal capital market with the development of the bill of exchange and the growth of early banking organization. By the end of the sixteenth century these and other institutional/organizational innovations had created the first modern economy in Amsterdam and the Seven Provinces that formed the Netherlands.⁴

But the creation of efficient economic markets is only half, and the less puzzling half, of the story. It is much more difficult to account for and explain the growth of "efficient" political markets that are a necessary precondition to the development of efficient economic institutions. It is the polity that specifies and enforces the economic rules of the game, and our knowledge of the essential conditions for the creation of such political institutions has not progressed much from the insightful observations of James Madison in the Federalist Papers.

However the historical evidence cited above provides essential clues to the evolutionary process in early modern Europe. It was the growing need for revenue in the face of the rising costs of warfare that forced monarchs to create "representative bodies"--parliament, estates general, cortes--and cede them certain rights in return for revenue. But at this point the stories diverge between the Netherlands and England on the one hand and much of the rest of Europe on the other. In the former the three considerations cited above (p. 16) led to the growth of representative government; in the latter they led to the persistence or revival of absolutist regimes with centralized decision making over the economy.

⁴. See North (1991) and Greif (1993) for a summary of the game theoretic and other analytical literature.

The contrasting subsequent history of the New World bears striking testimony to the significance of path dependence. In the case of North America, the English colonies were formed in the century when the struggle between parliament and the Crown was coming to a head. Religious and political diversity in the mother country was paralleled in the colonies. The general development in the direction of local political control and the growth of assemblies was unambiguous. Similarly the colonies carried over free and common socage tenure of land (fee simple ownership rights) and secure property rights in other factor and product markets.

The Spanish Indies conquest came at the precise time that the influence of the Castilian Cortes (parliament) was declining and the monarchy of Castile, which was the seat of power in Spain, was firmly establishing centralized bureaucratic control over Spain and the Spanish Indies. The conquerors imposed a uniform religion and bureaucratic administration on already existing agricultural societies. Wealth maximizing behavior by organizations and their entrepreneurs (political and economic) entailed getting control of, or influence over, the bureaucratic machinery.

One cannot make sense out of the contrasting subsequent history of the Americas north and south of the Rio Grande river without taking into account this historical background.⁵

VII

Whatever lessons there are from history come from the fundamentals of human interaction over time. Let me first draw some specific implications from

⁵. The downstream consequences in the New World are discussed in more detail in North (1990), chapter 12. Robert Putnam tells a similar story of the contrasting development of different regions in Italy from the 12th century onward in Putnam (1993).

the previous two sections. Then I shall tentatively suggest some of the deep underlying features distilled from historical experience that can form an agenda for future research as well as a backdrop for this conference.

We can briefly deal with economies that have failed to develop. These are cases in which the belief system, reflecting the historical experiences of that economy, fails to overcome the critical hurdle--the creation of institutions, economic and political, that will permit impersonal exchange. Markets remain local and/or poorly developed. As these economies became exposed to the larger world economy the institutional matrix typically spawns the development of organizations whose profitable opportunities predominantly favor redistributive activities.⁶ Political organizations in such cases will appear similar to those of the classic Virginia school of political economy, in which the state is little more than a theft machine. We do not have to look very far to see instances of such polities in Africa or, with some modification, in Latin America.

The obverse, the success stories, have been briefly described in the outlines above of the Netherlands and Britain--instances where external experiences reinforced evolving belief systems in the direction of productive activities by the economic organizations. Political organizations in these cases have evolved strong informal constraints--norms of behavior--that have constrained the behavior of the political actors. This point deserves special

⁶. The second economic revolution, the wedding of science to technology which began in the last half of the nineteenth century, is the source of modern economic growth and entails enormous specialization, division of labor, urban societies and global markets. The organizational restructuring--economic, political, and social--necessary to capture the gains from trade with this technology has necessitated fundamental changes in societies. The institutional adjustments are difficult enough for developed economies and much more difficult for transition and third world economies. See North (1981) chapter 13 for further discussion.

emphasis. The gradual development of informal norms of behavior that have become deeply imbedded in the society provides the stable underpinning to the adaptive efficiency characterizing the western economies with a long history of growth.

This last point can serve as an opening wedge to explore some fundamentals of human interaction. What underlies the formation and evolution of institutions? It is the ubiquitous human drive to reduce uncertainty--to create order from disorder that leads humans to scaffold both the mental models they possess--ie belief systems--and the external environment--ie--institutions.

VI

Let me conclude this chapter with two unresolved problems that have been on the surface or just beneath the surface throughout this essay; problems that are central to further progress in economic history/development. (1) Economic change is a process and it is that process which we must understand. The static nature of economic theory ill fits us to understand that process. We need to construct a theoretical framework that models economic change. (2) The belief systems of the players and the nature of human learning shape the evolving institutional policies that will be pursued. The rationality assumption of neo-classical economics assumes that the players know what is in their self interest and act accordingly. Ten millenia of human economic history says that is a wildly erroneous assumption.