

PN-ABD-354

Best available copy -- pages 2 and 31 missing

INSTITUTIONS AND ECONOMIC PERFORMANCE*

by

Douglass C. North

Luce Professor of Law and Liberty
and
Director, The Center in Political Economy
Box 1208
Washington University
One Brookings Drive
St. Louis, Missouri 63130

Copyright by Douglass C. North 1989

* I am indebted to Barry Weingast for comments on drafts of this essay and to Elisabeth Case for editorial improvements.

INSTITUTIONS AND ECONOMIC PERFORMANCE¹

by

Douglass C. North

(draft of April)

From the most primitive tribes to modern societies, human beings have always devised ways to structure human interaction. Institutions, whether solutions to simple problems of coordination (conventions), or to more complex forms of exchange such as those that characterize modern societies, provide a set of rules of the game that (together with other constraints) define and limit the choice set. They are the humanly devised constraints that shape human interaction so that when we wish to greet friends on the street, drive an automobile, buy oranges, borrow money, form a business, bury our dead, or whatever, we know or can learn how to do these things. It is easy to observe that institutions differ when we attempt to do the same things in a different country -- Bangladesh for example.

That institutions affect the performance of economies is hardly controversial. That the differential performance of economies over time is fundamentally influenced by the way institutions evolve is also not controversial. Yet neither current economic theory nor cliometric economic history shows much sign of appreciation of the role of institutions in economic performance. What has been missing is the development of an analytical framework to integrate institutional analysis into economics and economic history. This essay and the larger

Why is it so costly to transact? The short answer is that it takes resources to define and enforce exchange agreements. Even if everyone had the same objective function (eg maximizing the firm's profits) transacting would take substantial resources; but in the context of individual wealth maximizing behavior, and asymmetric information about the valuable attributes of what is being exchanged (or the performance of agents), the costs arising from transacting are a fundamental influence on economic activity.

A longer explanation of the costliness of transacting requires a more thorough examination of the nature of exchange. We owe to Lancaster (1966) and Becker (1965) the insight that a good or service is composed of a bundle of valuable attributes. It is only a short additional step to recognize that some attributes are physical (size, shape, color, location, taste, etc); others are property rights attributes (the right to use, to derive income from and to exclude others). To the extent that these attributes are separable they must be defined, that is measured, in order to be transferable in exchange. It is costly to measure and protect the rights over them (Barzel, 1982). This argument holds equally for the performance of agents in hierarchical organizations. It is also costly to enforce agreements. If exchange consisted of the transfer of a unidimensional good at an instant of time (implicit features of neo-classical theory) then these issues would be of trivial importance. But enforcing the exchange of multidimensional goods across space and time poses fundamental dilemmas of cooperation.

Let me illustrate the problems involved in complex exchange by briefly summarizing in very oversimplified fashion some of the implications derived from game theory. Wealth maximizing individuals will usually find it worthwhile cooperating with other players when the play is repeated, when they possess complete information about the other players past performance, and when there are small numbers of players. Such a crude summary disguises the richness (and ingenuity) of the results of an army of game theorists who have extended, elaborated, and modified (as well as found exceptions to) each of those qualifications to squeeze a great deal more out of them.

But let me turn the game upside down. Cooperation is difficult to sustain when the game is not repeated (or there is an end game), when information on the other players is lacking, and when there are large numbers of players. Now these polar extremes do in fact reflect real life contrasts. We do usually observe cooperative behavior when individuals repeatedly interact, when they have a great deal of information about each other and when small numbers characterize the group. But at the other extreme, realizing the economic potential of the gains from trade in a high technology world of specialization and division of labor characterized by impersonal exchange is rare because one doesn't necessarily have repeat dealings, know the other party, nor deal with small numbers. In fact, the pure essence of impersonal exchange is the antithesis of the conditions for game theoretic cooperation. The reason is that the costliness of measurement and enforcement as described above foreclose complex forms of exchange. Successful solutions have entailed the creation of institutions that in

game theoretic terms raise the benefits of cooperative solutions or raise the costs of defection; and that in transaction cost terms lower transaction plus production costs per exchange so that the potential gains from trade become realizeable. Regardless of the approach, the key is institutions.

-2-

Institutions consist of informal constraints and formal rules, and of their enforcement characteristics. Together they provide the rules of the game of human interaction. As I have defined institutions they could include organizations since organizations also provide a structure to human interaction. Indeed when we are examining the costs that arise as a consequence of the institutional framework, they are the result of both the basic institutional framework and the organizations that arise in consequence of the institutional framework. A great deal of this essay will blur the distinction between them. But conceptually they should be separated for reasons that are essential to understanding institutional change and will be elaborated in section 4 below.

Let me illustrate my definition by analogy with the rules of the game of a team competitive sport. They too consist of formal written rules and typically informal unwritten codes of conduct that underlie and supplement formal rules, such as not deliberately injuring a key player on the opposing team. These rules and informal codes are sometimes violated and punishment is enacted. Therefore, an essential part of the game is the likelihood of ascertaining violations and the severity (costliness) of punishment. Taken together the formal and

informal rules and the effectiveness of enforcement shape the whole character of the game. Some teams are successful as a consequence of (and therefore have the reputation for) constantly violating rules and thereby intimidating the opposing team. Whether that strategy pays off is a function of the effectiveness of monitoring and the severity of punishment. Conversely sometimes codes of conduct -good sportsmanship- constrain players even though they could get away with successful violations. It should be noted that it is one thing to analyze the rules that define the way the game is played but it is something else to model the organization and strategy that the team will develop as a response to the rules . Now let me return to institutions to elaborate on these common elements

Informal constraints include conventions that evolve as solutions to problems of coordination and that all parties are interested in having maintained; norms of behavior that are recognized codes of conduct; and self imposed codes of conduct such as standards of honesty or integrity. Conventions are self enforcing. Norms of behavior are enforced by the second party (retaliation) or by third party (societal sanctions or coercive authority) and their effectiveness will depend on the effectiveness of enforcement. Models of such exchange structures make up a large share of the game theory literature.

Self imposed codes of conduct, unlike conventions and norms of behavior, do not entail wealth maximizing behavior but rather the sacrifice of wealth or income for other values. Their importance in constraining choices is the subject of substantial controversy -- for

example in modeling voting behavior in Congress--(see Kalt and Zuppan, 1984). Most of the controversy has missed the crucial reason of why such behavior can be and is important. And that is, that institutions, frequently deliberately, sometimes accidentally, lower the cost to individuals of such behavior and can make ideas matter a great deal. Votes may not matter individually but in the aggregate they matter and they cost the voter very little; legislators commonly find enough ways by strategic voting to vote their personal preferences rather than those of the electorate; and judges with lifetime tenure are deliberately shielded from interest group pressures. In each of the above illustrations the institutional framework has altered the cost to the individual of expressing his or her convictions. In each case the choices that were made may be different than they would be if the individual bore the full cost that resulted from those actions. The lower the cost we incur for our convictions (ideas, dogmas) the more they contribute to outcomes (see Nelson and Silberberg, 1987 for empirical evidence).

Formal rules differ in degree from informal constraints. On a continuum from taboos, customs, traditions at one end to written constitutions at the other end of the scale the gradual transition in history has been uneven but unidirectional. Most conspicuously, both the sources and the rate of change are different as between formal rules and informal constraints. Formal rules are altered by deliberate action of political, judicial or economic bodies. While informal constraints are certainly influenced by alterations in formal rules the sources of change are much more complex, much less understood, and the rate of

change very different. This difference has important implications for institutional change.

Formal economic rules are typically nested in a hierarchy, from constitutions to statute and common law to specific contracts and by-laws of organizations; they are more costly to alter as we go higher on the ladder. Formal political rules specify the hierarchy of the polity from basic decision rules to agenda control. Economic rules define property rights, that is the bundle of rights over the use and the income to be derived from property, and the rights of alienation. Both political and economic rules are devised as a consequence of the bargaining strength of those making the decision rules; marginal changes occur with changes in bargaining strength (to be discussed below). But given the initial bargaining strength of the parties the function of the rules is to facilitate exchange, both political and economic.

The costliness of defining and enforcing agreements reflects the effectiveness of the institutions. The ability at low cost to measure what is being exchanged and to enforce agreements across time and space requires complex institutional structures; conversely, the inability at low cost to measure and enforce agreements has been a consequence of institutions that make it costly to transact (or produce). Successful economic growth is the story of the evolution of more complex institutions that make possible cooperative exchange relations extending over long periods of time, amongst individuals without personal knowledge of each other. Institutional reliability means we can have confidence in outcomes increasingly remote from our personal knowledge.

The combination of formal rules, informal constraints, and enforcement characteristics of institutions defines the humanly devised constraints and, together with the traditional constraints of standard theory, the choice set. The property rights literature has long since demonstrated that different property rights produce different outcomes; but because it has not taken into account both the effectiveness of enforcement and informal constraints, that approach is incomplete and at least partly misleading. The choices as reflected in contracts between exchanging parties actually will reflect not only formal constraints, but also the uncertainties arising from the effectiveness and costliness of enforcement. Equally conventions, informal community sanctions, will play a part in the exchange. Therefore, to understand the choices available in an exchange one must take into account all the dimensions that make up an institution.

Let me illustrate the relationship between institutions (formal rules, informal constraints, and enforcement characteristics) and transaction costs in a specific example--the exchange of a residential property in modern United States. I examine first the transaction costs incurred in the transfer and then the institutions that determined those costs of transacting.

In the seller's utility function are the price, terms, and security of the contractual obligation; that is, the likelihood that the buyer will live up to the contract ex post. The value of the residence to the buyer is a function not only of price and credit terms but also of the attributes that are transferred with the sale. Some, such as the

legal rights that are and are not transferred, the dimensions of the property and house, are easily measured; others, such as the general features of the property, are readily observed on inspections. But still others, such as the maintenance and upkeep costs and the characteristics of neighbors, may be far more difficult to ascertain. Equally, the security of property against default, expropriation, uncertain title, or theft will vary according to the difficulty of ascertaining the likelihood of each and therefore its importance.

Now in the traditional neo-classical paradigm, with perfect information--ie. zero transaction costs--the value of the asset that is transferred assumes not only perfect information but perfectly secure property rights. In that case, since both buyer and seller have been able to costlessly ascertain the value of all the attributes (both physical and property rights) and there is no uncertainty or insecurity of property rights, the standard supply and demand models of housing with zero transaction costs would define the value of the asset. In fact, because all of the above-mentioned attributes influence the value of the residence to the buyer and seller, the smaller the discount (ie the smaller the transaction costs incurred) from the idealized neo-classical model, the more perfect the market. It is institutions in the aggregate that define and determine the size of the discount, and it is transaction costs that the buyer and seller incur that reflect that institutional framework.

The transaction costs of the transfer are partly market costs, such as legal fees, realtors fees, interest charges, title insurance,

credit rating searches; and partly the costs of time each party must devote to gathering information, the costs of searching, etc. Obtaining information about crime rates, police protection, security systems entail search costs for the buyer. To the degree that the buyer's utility function is adversely affected by noisy neighbors, pets, etc. it will pay to invest time in ascertaining neighborhood characteristics and the norms and conventions that shape neighborhood interactions.

The particular institutional matrix of this housing market consists first of all of a hierarchy of legal rules, derived from the U.S. Constitution and the powers delegated to the states. State laws defining the conveyance characteristics of real property, zoning laws restricting which rights can be transferred, common and statute law undergirding, defining, or restricting a host of voluntary organizations: all of these influence transaction costs. Realtors, title insurance, credit bureaus, savings and loan associations that affect the mortgage market all will be influenced. The efficiency of these organizations is a function of the structure of property rights and enforcement (such as title insurance costs) and the structure of the capital market (including both voluntary organizations and governmental organizations, guarantees, and subsidies). Equally important are a range of informal constraints -- conventions, norms, and codes of conduct that broadly supplement and reinforce the formal rules. They range from conventions of neighborhood conduct to ethical norms defining degrees of honesty in information exchange between the variety of parties involved.

My description has emphasized institutions that lower the costs of transacting but some --such as rules that restrict entry, require useless inspections, raise information costs, or make property rights less secure-- in fact raise transaction costs. Institutions everywhere are a "mixed bag" of those that lower the costs and those that raise them. The U.S. residential housing market is a relatively efficient market in which on balance the institutions induce low cost trasacting.

The fundamental implication of the foregoing illustration is that the discount from the frictionless exchange envisioned in economic theory will be greater to the degree that the institutional structure allows third parties to influence the value of attributes that are in the utility function of the buyer. These could be the behavior of neighbors, the likelihood of theft, the possibility of changes by local authorities in zoning ordinances that may affect the value of the property, etc. The greater the uncertainty of the buyer, the lower the value of the asset. Likewise, the institutional structure will equally determine the risks to the seller that the contract will be fulfilled or that the seller will be indemnified in case of default. It is worth emphasizing that the uncertainties described above with respect to security of rights are the critical distinction between the relatively efficient and secure markets of high income countries and the insecure and costly nature of these transactions in economies both in the past and in the present Third World.

Institutions play an even more decisive role in the production of goods and services since institutional structures affect both production

and transaction costs; the latter via the direct connection between institutions and transaction costs as described in the above illustration; the former by influencing the technology employed. All the usual problems of measurement and enforcement obtain; that is institutions shape the consequent transaction and production costs via the structure of property rights, the effectiveness of the courts and the judicial system, and the complementary development of voluntary organization and norms.

Specifically the firm's entrepreneur must be able to ascertain the quantity and quality of inputs and outputs. Since in the neo-classical firm these can be obtained costlessly, the contrast between a hypothetical neo-classical firm and a real firm is striking. The former was little more than a production function without any costs of organization, supervision, coordination, monitoring, metering, etc. However a real-life firm must purchase inputs that constantly require measurements and metering if it is to produce output of constant quality since variability in quality will, ceteris paribus, adversely affect demand for its product. Otherwise consumers (or if it is an intermediate good, producers) must (when quality is variable) devote resources to ascertaining quality; hence producers who can guarantee constant quality will be favored.²

These conditions (ie costless measurement and enforcement) are implicitly assumed in what we call efficient factor and product markets but their existence entails a complex set of institutions that encourage factor mobility, the acquisition of skills, uninterrupted production,

rapid and low cost transmission of information, and the invention and innovation of new technologies. Realizing all these conditions is a tall order never completely filled since, as with the institutions of exchange described above, the actual institutional framework is a mixed bag of those institutions that promote productivity raising activities and those that provide barriers to entry, encourage monopolistic restrictions, and impede the low cost flow of information.

We have only to contrast the organization of production in a third world economy with a first world economy to be impressed by the consequences of overall poorly defined and/or ineffective property rights. Not only will the institutional framework result in high costs of transacting in the former but also insecure property rights will result in using technologies that employ little fixed capital and do not entail long term agreements. Firms will typically be small (except those operated or "protected" by the government). Moreover such mundane problems as an inability to get spare parts or a two year wait to get a telephone installed will necessitate a different organization of production than in an advanced country. Now it is usually true that a sufficient bribe may exist that will get quick delivery of spare parts through the maze of import controls or get rapid telephone installation but the resultant "shadow" transaction cost does significantly alter relative prices and consequently the technology employed.

Even with the relatively secure property rights that exist in high income countries it is frequently the case that a technical combination that involves costly monitoring may be less "efficient" than a technique

that has lower physical output but less variance in the product quality or lower costs of monitoring the worker. Because much of the recent transaction cost literature implies that institutions only determine transaction costs and techniques only determine production costs let me illustrate three different choices arising from the interplay between techniques, institutions, production costs and transaction costs to make clear that the relationship among them is more complex.

1) A contention of Marxist writers is that deliberate deskilling of the labor force occurred during the early 20th century in the United States. That is, employers adopted capital intensive techniques which eliminated the demand for highly skilled workers and replaced them with semi-skilled or unskilled workers. The explanation for this choice is that the bargaining power of skilled workers enabled them to strategically disrupt the production process, which given the "high speed throughput" (Chandler's term, 1977) of modern technology was enormously costly. Long run total cost could be reduced by using less skilled workers who were without the bargaining power to disrupt production. In this case a new technique was introduced to lower transaction costs.

2) Unitizing an oil field, that is creating an organization with the coercive power and monitoring authority to allocate the output of the oil field, raises transaction costs (because of the resources that must be devoted to creating and maintaining an organization and then to monitoring compliance). At the same time it reduces production costs (the result of more efficient pumping and recovery) to an extent that

more than offsets the rise in transaction costs (Libecap and Wiggins, 1985). In this case an institutional change raised transaction costs which were more than compensated by lower production costs.

3) Andrea Shepard (Rand Journal, 1986) describes the deliberate policy of a semi-conductor manufacturer who licenses the design of new chips to competitors, so that customers can be assured that the chip manufacturer will not be able to hold up customers who adopt the new design. By alleviating customers' concerns, this policy enhances demand for the product. While this policy lowers transaction costs, it does so at the sacrifice of productive efficiency, since both scale economies and "learning curve" effects are lost to competing firms.

Informal institutional constraints frequently play a major role with respect to the quantity and quality of labor output. While Marxists long ago recognized that the quantity of labor input could not be mechanically transformed via a production function into the quantity and quality of output, this subject has only recently become a major focus of economists' concern (at least partially a consequence in recent years of the quality difference in labor output between Japanese and American automobile manufacturers). Conventions about output, forms of organization designed to encourage worker participation and cooperation, and attempts to select labor with an ideological commitment to hard work have all become recent research agendas in industrial organization. The unique feature of labor markets is that institutions are devised to take into account that the quantity and quality of output is influenced by the attitude of the productive factor (hence investing in persuasion,

morale building, etc. is a substitute at the margin for investing in more monitoring).

-3-

The major focus of the literature on institutions and transaction costs has been on institutions as efficient solutions to problems of organization in a competitive framework (Williamson, 1975, 1985). Thus market exchange, franchising, or vertical integration are efficient solutions to the complex problems confronting the entrepreneur in various competitive environments. Valuable as this work has been, it leaves out the most important contribution which institutional analysis can make to economics: to explain the diverse performance of economies. How do we account for the poverty of nations, the failure of some economies to grow, or for that matter the differential performance of sectors in an economy? Institutions structure incentives, which in turn determine the performance of economies. The formal economic constraints (property rights) are specified and enforced by political institutions and the literature described above simply takes those as a given. While there is a large literature on regulation and even modeling political outcomes (for example Becker, 1983, 1985), it is essentially a-institutional and therefore fails to recognise that different political institutions will affect the efficiency of political exchange and hence economic outcomes.

To explain the diverse performance of economies let me start with a simple model of a polity consisting of a ruler and diverse constituents.³ In this setting the ruler acts like a discriminating

monopolist, offering to different groups of constituents "protection and justice"--or at least the reduction of internal disorder and the protection of property rights--in exchange for tax revenue. Since different constituent groups have different opportunity costs and bargaining power with the ruler, different bargains result. There are economies of scale in the provision of these (semi) public goods of law and enforcement and total revenue is therefore increased. However, the division of the incremental gains between ruler and constituents depends on their relative bargaining strength; changes at the margin, either the violence potential of the ruler or the opportunity cost of the constituents, will result in redivision of the incremental revenue.

This model of the polity can become one step more complicated when I introduce the concept of a representative body reflecting the interests of constituent groups, which bargains with the ruler. This concept is consistent with the origin of Parliaments, Estates Generales and Cortes in early modern Europe which evolved as a response to the ruler's need for more revenue. In exchange for this revenue the ruler would agree to provide certain services to constituent groups. The representative body facilitates exchange between the parties. On the ruler's side, this new relationship with constituents leads to the development of a hierarchical structure of agents. This is a major transformation from the simple (if extensive) management of the king's household and estates to a bureaucracy monitoring the wealth and or income of the king's constituents.

When we move from the politics in early modern Europe to modern representative democracy our story is complicated by the development of multiple interest groups and by an institutional structure much more complicated but still devised to facilitate (again given relative bargaining strength) the exchange between interest groups. This political transaction cost analysis is built on the recognition of the multiplicity of interest groups reflecting concentrations of voters in particular locations. Thus, to illustrate from the United States political scene, there are elderly in Florida and Arizona, miners in Pennsylvania and West Virginia, artichoke growers in California, automobile manufacturers in Michigan, etc. Each legislator's district has a concentration of no more than a few of the large number of interest groups. Therefore, legislators cannot succeed acting alone, and must make agreements with other legislators, with different interests.

What kind of institutions will evolve from such exchange relationships between legislators?. Previous work, beginning with Buchanan and Tullock, focused on vote-trading, or log-rolling. This work was certainly a step forward in recognizing the way by which legislators could engage in activities that facilitated exchange. However, such an approach was too simple to solve fundamental problems involved in legislative exchange. It assumed that all bills and payoffs were known in advance, and it had a timeless dimension to it.

In fact, a variety of exchanges arise in which today's legislation can only be enacted by commitments made for a future date. In order to lower the costs of exchange, one must devise a set of institutional

arrangements that would allow for exchange across space and time. Note the parallels with economic exchange as described above. How does credible commitment evolve to enable agreements to be reached when the payoffs are in the future and on completely different issues? Self-enforcement is important in such exchange and in repeat dealings a reputation is a valuable asset. But, as in economic exchange, the costs of measurement and enforcement, discovering who is cheating whom, when free-riding will occur, and who should bear the cost of punishing "defectors" make self-enforcement ineffective in many situations. Hence political institutions constitute ex-ante agreements about cooperation among politicians. They reduce uncertainty by creating a stable structure of exchange. The result is a complicated system of committee organization consisting of formal rules and informal constraints that together shape legislative choices. Its evolution in the American Congress is described in a recent study of the structure by Barry Weingast and William Marshall (1988).

While political institutions facilitate exchange amongst bargaining parties, there is no implication of economic efficiency as an outcome. In an earlier study (North, 1981) I argued that there were two basic reasons why rulers typically produced inefficient property rights (defined here simply as rules which do not produce increases in output). First, the competitive constraint on the ruler means that a ruler will avoid offending powerful constituents with close access to alternative rulers. He will agree to a property rights structure favorable to those groups regardless of its effects on efficiency. Second is a transaction cost constraint. While efficient property rights would lead to higher

societal income they may not lead to more tax revenues because of higher costs of monitoring, metering and collecting. Granting guilds monopolies in Colbert's France may not have been efficient but it did improve tax collecting as compared to an unregulated decentralized economy.

The same two constraints have obtained throughout history (and continue to obtain). Inefficient (as defined above) economic institutions are the rule not the exception. It is not that political entrepreneurs would not like to have economic growth; it is that the institutions that have evolved do not create conditions of credible commitment that makes low cost transacting possible. Moreover the process of institutional change does not result in evolutionary competition weeding out inefficient institutions in favor of efficient ones (or at least the tendency is so weak and diffuse as to permit the persistence of inefficient economies for very long periods of time). Let us see why.

4

Understanding institutional change entails an understanding of (1) the stability characteristics of institutions, (2) the sources of change, (3) the agents of change, and (4) the direction change.

A basic function of institutions is to provide stability and continuity by dampening the effects of relative price changes. It is institutional stability that makes possible complex exchange across space and time. A necessary condition for efficient markets which underlie high income societies are channels of exchange, both political

and economic, which make possible credible agreements. This condition is accomplished by the complexity of the set of constraints that constitute institutions; by rules nested in a hierarchy, each level more costly to change than the previous one. In the United States the hierarchy moves from constitutional rules to statute law and common law to individual contracts. Political rules are also nested in a hierarchy even at the level of specific bills before Congress. Both the structure of committees and agenda control assure that the status quo is favored over change. Informal constraints are even more important anchors of stability. They are extensions, elaborations and qualifications of rules that "solve" numerous exchange problems not completely covered by formal rules and hence have tenacious survival ability. They allow people to go about the everyday process of making exchanges without the necessity of thinking out exactly at each point and in each instance the terms of exchange. Routines, customs, traditions and culture are words we use to denote the persistence of informal constraints. It is the complex interaction of rules and informal constraints, together with the way they are enforced, that shapes our daily living and directs us in the mundane (the very word conjures up images of institutional stability) activities that dominate our lives. It is important to stress that these stability features in no way guarantee that the institutions are efficient (as defined above). Stability is a necessary condition for complex human interaction but it is not a sufficient condition for efficiency.

One major source of institutional change has been fundamental changes in relative prices (see North and Thomas, 1973 for illustration)

but another has been changes in preferences. I know of no way to explain the demise of slavery in the 19th century in an interest group model. The growing abhorrence on the part of civilized human beings of one person owning another not only spawned the anti-slavery movements but through the institutional mechanism of voting resulted in its elimination. It is not that interest groups did not use the abolitionist movement to further their interests. They did. But the success of the interest groups did entail the ideological support of the voter. The voter paid only the price of going to the polls to express his conviction and the slave owner had no feasible way to bribe or pay off voters to prevent them from expressing their beliefs. As noted earlier, institutions make ideas matter.'

The agent of change is the entrepreneur--political or economic. So far I have left organizations and their entrepreneurs out of the analysis and the definition of institutions has focused on the rules of the game rather than the players. As noted at the beginning of this essay this separation of institutions from organizations was deliberate. Left out was the purposive activity of human beings to achieve objectives which in turn result in altering constraints. Organizations and learning alter outcomes, but how?

Let me begin with organization. More than half a century ago Coase (1937) argued that transaction costs are the basis for the existence of the firm. That is, if information and enforcement were costless, it is hard to envision a significant role for organization. What is it about transaction costs that leads to organization? The answers have ranged from the firm being a form of exploitation

(Marglin,1974), to a response to asset specificity (Williamson, 1975,1985,) to a response to measurement costs (Barzel, 1982). Whatever the merits of these alternatives (and they are not altogether mutually exclusive), they all focus on the trees but not the forest.

Organizations are a response to the institutional structure of societies, and, in consequence, the major cause of the alteration of that institutional structure. Let me explain.

The institutional constraints together with the traditional constraints of economic theory define the potential wealth maximizing opportunities of entrepreneurs (political or economic). If the constraints result in the highest payoffs in the economy being criminal activity, or the payoff to the firm is highest from sabotaging or burning down a competitor, or to a union from engaging in slowdowns and makework, then we can expect that the organization will be shaped to maximize at those margins. On the other hand if the payoffs come from productivity enhancing activities then economic growth will result. In either case the entrepreneur and his or her organization will invest in acquiring knowledge, coordination and "learning by doing skills" in order to enhance the profitable potential. As the organization evolves to capture the potential returns it will gradually alter the institutional constraints themselves. It will do so either indirectly, via the interaction between maximizing behavior and its effect on gradually eroding or modifying informal constraints; or directly, via investing in altering the formal rules. The relative rate of return on investing within the formal constraints or devoting resources to

altering the constraints will reflect the structure of the polity, the payoffs to altering the rules, and the costs of political investment.

Let me briefly expand on this model of institutional change by reframing a familiar story in American economic history--the growth of the economy in the 19th century. The basic institutional framework that had been carried over from England had not only encouraged decentralized and local political autonomy but also provided low cost economic transacting through fee simple ownership of land(with some early exceptions in proprietary colonies) and secure property rights. The post revolutionary enactments of the Northwest Ordinance and the Constitution codified, elaborated, and modified colonial institutions in the light of contemporary issues (and the bargaining strength of the players) but created an institutional environment that broadly induced the development of economic and political organizations that promoted increased productivity and economic growth (both directly and indirectly by an induced demand for education, for example). But it should be carefully noted that this institutional framework also spawned some organizations and policies that raised transaction costs and hence reduced efficiency (the Know-Nothing party or tariffs for example). Moreover as these political and economic organizations evolved to take advantage of profitable opportunities they gradually altered the basic institutional framework. Sometimes these alterations made the basic institutional framework even more conducive to productive activity; sometimes however they raised the rate of return to unproductive activity; sometimes the results were unanticipated by the entrepreneurs (political or economic). Exogenous forces such as changes in political

or economic conditions in the rest of the world induced changes in the American economy by altering relative political or economic prices to domestic political and economic entrepreneurs and their organizations and hence leading them to actions that altered the institutional framework. The story is familiar, but, by focusing on the interaction between the rules and the players, this approach has, I believe, the promise of telling a far more interesting history than heretofore. More interesting because it can account for the path of historical change.

Institutional change therefore is an incremental process in which short run profitable opportunities cumulatively create the long run path of change. The long run consequences are often unintended for two reasons. First, the entrepreneurs are seldom interested in the larger (external to them) consequences but the direction of their investment influences the extent to which there is investment in adding to or disseminating the stock of knowledge, encouraging or discouraging factor mobility, etc. Second, there is frequently a significant difference between intended outcomes and actual outcomes. Outcomes frequently diverge from intentions because of the limited capabilities of individuals and the complexity of the problems to be solved. The path of institutional change that determines the long run evolution of societies is shaped by constraints derived from the past and the (sometimes unanticipated) consequences of the innumerable incremental choices of entrepreneurs which continually modify those constraints. Path dependence means that history matters, that it is a consequence of incremental institutional change and that it can account for the divergent paths of economies.

Moreover given the tendency of polities to produce inefficient property rights, economic decline or stagnation can persist since there will not typically develop a feedback that will create organizations with the incentive to invest in productive activity. Instead the "perverse" incentives will generate organizations and hence entrepreneurs with economic and political bargaining strength who will find it profitable to pursue economically inefficient paths. The contrasting histories of England, its North American colonies and their subsequent development (briefly outlined above) with Spain (and Portugal) and subsequent Latin American development is striking. In the former the institutional framework that evolved was broadly conducive to the creation of organizations that induced political democracy, stability and economic growth. In the latter, centralized bureaucratic political controls and detailed regulation of the economies carried over to the colonies and persisted even after independence. The long run consequences were not only political instability but the relatively poor economic performance that has characterized two centuries of Latin American history. Typically the opportunities that were open to political and economic entrepreneurs were policies that reinforced the existing strictures. This was so because the opportunities consisted of marginal changes that were shaped by an overall institutional framework that overwhelmingly favored such institutional policies. But this is not the whole story. With the revolutions, formal rules were instituted to alter the polities and economies. The ideological winds from the American Revolution did induce laws patterned after the U.S. Constitution and economic legislation was enacted to reduce or eliminate

regulatory constraints. Yet the consequences were radically different. It is the complex of formal rules, the way they are enforced and the informal constraints that together define the institutional framework. In the Latin American institutional environment, changes in the formal rules alone were not sufficient to redirect politics and economies in a new direction.

While the contrasting "path dependent" stories are clear enough, the incremental institutional and organizational evolutions are far from being well understood. It is relatively easy to trace the specific evolution of organizations and the way by which their development influenced the institutional framework (see, for example, North and Rutten, 1987 on the evolution of U.S. land policy). But the organizational consequences were not unidirectional (as I have stressed above) and the aggregate effects on economic and political performance were almost always a complicated multi-dimensional story.

Institutional change is overwhelmingly incremental; but discontinuous institutional change does occur in the form of revolution. It would take me far beyond the limits of this essay to deal properly with this topic, but one important point follows from the preceding analysis. That is that revolutionary change is seldom as revolutionary as it appears on the surface (or in the utopian vision of revolutionaries). The reason is not just that the "half-life" of ideological commitment tends to be short but that the formal rules change while the informal constraints do not. In consequence there develops an ongoing tension between informal constraints and the new formal rules, many of which are inconsistent with each other. The long

run resolution tends to be some restructuring of both and an outcome that retains or even recreates some of the pre-revolutionary formal constraints.

Let me conclude by summarizing the contribution that institutional analysis can make to economics. I shall not elaborate on the already rich literature that has sprung up from institutional and transaction cost analysis in industrial organization, public finance, public choice but instead focus on the broader contribution that is still to be undertaken:

1. The most general contribution that institutional modeling can make to economic theorizing is to make clear and explicit the institutionally specific context within which the model holds. Implicit in most economic models are specific political rules, property rights and enforcement characteristics that are critical to the outcomes. Changes in these would produce different outcomes. Economists, however, seem seldom to be aware of just how specific their model is to the institutional constraints (how would it work in Bangladesh for example).

2. A self conscious incorporation of institutions into economic theory will force economists to question the behavioral models that underlie the discipline. If neo-classical economics has no institutions in the models, it is because the behavioral assumption, which incorporates characteristics about human behavior as well as about the information that the players have, does not require it. But since institutions really exist, it is incumbent on the economist to ask searching questions about the implications of institutions for the

behavioral model that the economist employs. The role of ideology, for example, plays an important part once we recognize that people have subjective perceptions about the world around them and that expressing convictions in various institutional contexts frequently can be done at negligible cost to the individual. Likewise when our behavioral models incorporate the incompleteness of our information and our limited ability to process that information, then we will understand why we need to develop the regularized patterns of human interaction that we call institutions and why they may be very inadequate or far from optimal in any sense of the term. The nascent cooperation between psychologists and economists offers the promise of enriching our behavioral models (see Hogarth and Reder, 1986).

3. Incorporating institutions in their models should make economists aware that ideas matter. Institutions structure human interaction so that we frequently and in many critical choice contexts can express our ideas, ideologies, and dogmas at little or no cost to ourselves. The result is to frequently produce different outcomes than those derived from interest group models in economics and public choice.

4. The integration of political and economic theory is essential in a world where government plays such an immense role in choices. The key to such integration is the modeling of political and economic institutions that will permit us to explore in theoretical terms the interaction between these two institutional structures and to derive in consequence real political economy models in macroeconomics and other areas in which government plays a critical role.

away all the relevant issues. It is institutions that provide the key constraints and therefore shape incentives, and it is the interaction between the institutional framework and the organizations that are a response to that framework that shapes the evolution of economies. Institutional theory focuses on the critical problems for development of human organization and the problems of achieving cooperative solutions to human interaction.

It is only appropriate that an economic historian should conclude this essay by making what should by now be an obvious point. Institutional analysis should place history in a more central role in economics, not only because it sheds light on a critical parameter held constant by the economist--institutions, but also because the constraints within which choice making occurs are derived from the past, and with an appreciation of the way those constraints have evolved, we can have a far better understanding of the choice set today and the institutionally specific context within which the economist's model holds.

¹This essay is drawn from, and is a drastic condensation of, parts of a forthcoming book by the author tentatively entitled, INSTITUTIONS, INSTITUTIONAL CHANGE AND ECONOMIC PERFORMANCE.

²For a fascinating analysis of the significant resource costs that the producer must engage in to assure constant quality, see the detailed description of the production of peas, in Susan Sheehan's essay "Peas" in the New Yorker, June 17, 1973. The trouble that Green Giant went to in attempting to eliminate variability in size, tenderness and sweetness in the production of peas involved enormous monitoring and metering resources that began in the field and did not end until the cans went off to the retailer.

³. This model is elaborated in "A Neo-Classical Theory of the State" in North, 1981.

BIBLIOGRAPHY

- Barzel, Y. 1982. "Measurement Costs and the Organization of Markets," Journal of Law and Economics 25: 27-48.
- Becker, G. S. 1965. "A Theory of the Allocation of Time," Economic Journal 75: 493-517.
- Becker, G. S. 1983. "A Theory of Competition among Pressure Groups for Political Influence," Quarterly Journal of Economics XCVIII: 371-400.
- Becker, G. S. 1985. "Public Policies, Pressure Groups, and Dead Weight Costs," Journal of Public Economics 28: 329-347.
- Buchanan, J. M. and Tullock, G. 1962. The Calculus of Consent. Ann Arbor: University of Michigan Press.
- Chandler, Alfred 1977. The Visible Hand. Cambridge, Mass.: The Belknap Press.
- Coase, Ronald 1937. "The Nature of the Firm," Economica 4 (November): 386-405.
- Hogarth, R. M. and M. W. Reder, eds. 1986. "The Behavioral Foundations of Economic Theory," Journal of Business Volume 57.
- Kalt, J. P. and M. A. Zupan 1984. "Capture and Ideology in the Economic Theory of Politics," American Economic Review 75: 279-300.
- Lancaster, Kelvin 1966. "A New Approach to Consumer Theory," Journal of Political Economy 74: 132-157.
- Libecap, G. D. and S. N. Wiggins. 1985. "The Influence of Private Contractual Failure on Regulation: The Case of Oil Field Unitization," Journal of Political Economy 93: 690-714.
- Marglin, S. 1974. "What Do Bosses Do?" Review of Radical Political Economy 6: 60-112.
- Nelson, D. and E. Silberberg. 1987. "Ideology and Legislator Shirking," Economic Inquiry 25: 15-25.

- North, Douglass C. 1981. Structure and Change in Economic History. New York: W. W. Norton and Company.
- North, D.C. and R. Thomas. 1973. The Rise of the Western World: A New Economic History. Cambridge, Mass.: The University Press.
- North, D.C. and J. Wallis 1987. "Measuring the Transaction Sector in the American Economy, 1870-1979," in S. L. Engerman and R. E. Gallman, eds. Long-Term Factors in American Economic Growth. Chicago: University of Chicago Press.
- North, D. C. and A. Rutten 1987. "The Northwest Ordinance in Historical Perspective," in D. Klingaman and R. Vedder, Essays on the Economy of the Old Northwest. Athens: Ohio University Press.
- Shepard, A. 1987. "Licensing to Enhance Demand for New Technologies," Rand Journal of Economics 18: 360-368.
- Weingast, B. R. and W. J. Marshall. 1988. "The Industrial Organization of Congress; or, Why Legislatures, Like Firms, Are Not Organized as Markets," Journal of Political Economy 96: 132-163.
- Williamson, O. E. 1975. Markets and Hierarchy. New York: The Free Press.
- Williamson, O. E. 1985. The Economic Institutions of Capitalism. New York: The Free Press.