INSTITUTIONS AND PRODUCTIVITY IN HISTORY

by
Douglass C. North, Washington University,
St. Louis

An economic historian trying to make sense out of the widely varied patterns of growth of economies over the past several millenia using the modern productivity literature or its quantitative basis, the growth accounting literature should certainly be puzzled. Equally an explanation in terms of that literature of the performance characteristics of the Soviet Union over its some 70 years of existence should remain a puzzle. Institutions play no role in the stories at all.

The performance of an economy is a function of both the institutions and technology developed in a society but the literature on productivity is devoid of any discussion of the institutional foundations of the economy. Instead that literature and the growth accounting quantitative foundations of that analysis focus on the significance of the capital stock, technology, r&d, and savings in the rate of growth of productivity. The new growth economics literature equally equates growth with human capital invstment, physical capital, savings and sometimes as a negative function of population growth.

The argument of this essay is that productivity increases result from both improvements in human organization and from technological developments. Indeed it is probably true that the former is as important as the latter in economic growth. In the following sections I first lay out the theoretical justification for this arguement (I), then explore some of the key institutional changes in history that have laid the foundation for modern economic growth (II), explore the interplay between institutional and technological changes in the past 150 years,the era of the 2nd economic revolution (III), discuss the institutional and transaction cost changes that have characterized that revolution (IV) and conclude with some implications and questions that this analysis poses for future productivity change and economic growth (V).

I

Institutions are the incentive structure of a society and therefore the rules, norms, and enforcement characteristics that make up the institutional foundations of a society direct the allocation of resources of that society and economy. Economic growth throughout history could only be realized by creating an institutional and organizational structure that would induce productivity enhancing activity--a supply side argument; and equally that the consequent tensions induced by the resulting societal transformation have resulted (and are continuing to result) in politically-induced fundamental changes in the institutional structure to mitigate these tensions--a demand side argument. Both the

supply side and demand side institutional changes have been and continue to be fundamental influences on productivity change.

The supply side changes have improved productivity by lowering either transaction and/or transformation costs. Transaction costs are the costs of measuring and enforcing agreements and transformation costs are the costs of physical production. Together they make up the total costs of an organization. Institutions that lowered transaction costs are those that enabled the parties to exchange to specify and define the multiple valuable attribute of a good or service or the performance of agents or to enforce agreements at lower cost. Institutions that have lowered transformation costs have been those that have made possible the introduction or application of an improvement in technology.

Supply induced changes in organization and technology have altered the perceived costs and benefits to individuals and groups throughout history and engendered responses through group action or the political process that have in turn altered the productivity of the economy. From the age old opposition by groups to technological change to the modern environmental movement demand side changes continue to play a major part in influencing productivity.

II[1]

Institutional innovations that lowered transaction costs in the early modern western world were 1) those that increased the mobility of capital; 2) those that lowered information costs; 3) those that spread risks; and 4) those that improved the enforcement of contracts. The innovations had their origins in earlier times; most of them were borrowed from medieval Italian city states or Islam or Byzantium and elaborated upon in subsequent development. Since there is an immense literature in economic history on them my objective here is to focus on their implications for lowering transaction costs and hence improving productivity.

The variety of techniques that evolved to evade usury laws such as "disguising them as late payment" or by exchange rate manipulation (Lopez and Raymond, 1955, p. 163) were early and cumbersome ways to disguise interest payments and made the enforceability of such contracts problematic. As usury laws broke down and rates of interest were permitted the cost of writing contracts and enforcing them declined.

A major innovation that affected the mobility of capital was the evolution of the bill of exchange and particularly the development of techniques and instruments that allowed for the negotiability of the bill of exchange and the development of discounting methods. Negotiability and discounting in turn depended on the creation of institutions that would permit their use and the development of centers where negotiating and discounting could occur: first fairs, such as the Champagne fairs, then banks, and finally through financial houses that could specialize in discounting. Increasing volume obviously made such institutional developments possible. In addition to the economies of scale necessary for the development of the bills of exchange, improved enforceability of contracts was critical, and the interrelationship between the development of accounting and auditing methods and their use as evidence in the collection of debts and in the

---

[1]. This section draws from my essay "Institutions, Transaction Costs and the Rise of Merchant Empires" in Tracy (ed) 1991

enforcement of contracts was an important part of this process (Yamey, 1949 and Watts and Zimmerman, 1983).

Still another innovation affecting the mobility of capital arose from the problems associated with maintaining control of agents involved in long distance trade. The traditional resolution of this problem in medieval and early modern times was the use of kinship and family ties to bind agents to principals in ways that provided some assurance to the principal that the orders and directions of the principal were safely carried out (the church`s greater success with agents probably reflected ideological commitment). However, as the size and scope of merchant trading empires grew, the extension of discretionary behavior to others than kin of the principal required the development of sophisticated accounting and auditing methods and more elaborate procedures for monitoring the behavior of agents.

A second general source of improving productivity that lowered transaction costs was developments that lowered information costs, including the printing of prices of various commodities; the printing of manuals that provided information on weights, measures, customs, brokerage fees, postal systems, and, particularly, on the complex exchange rates between monies in Europe and the trading world. Obviously these developments were primarily a function of the economies of scale resulting from the volume of international trade.

Then there were the institutional innovations that transformed uncertainty into risk. We think of insurance and portfolio diversification in the modern world as methods for converting uncertainty into risks and thereby reducing, through the provision of a hedge against variability, the costs of transacting. When we look at the medieval and early modern world, we find the same innovations. That is, marine insurance evolved from sporadic individual contracts covering partial payments for losses to contracts issued by specialized firms.

Marine insurance was one example of the development of actuarial, ascertainable risk; another was business organization that spread risk through either portfolio diversification or institutions that permitted a large number of investors to engage in risky activities. The commenda itself, from its Jewish, Byzantine, and Muslim origins (Udovitch, 1970) through its evolution at the hands of Italians to the English Regulated Company and finally the Joint Stock Company, provides an evolutionary story of the institutionalization of risk (although as discussed below, the developments created new problems of agency for the principals involved).

. When we look at the development of enforcement mechanisms, we see that the process was a long one. Greif(1989) describes the development of a reputation mechanism among Jewish traders that permitted long distance trade in the Meditteranean as early as the tenth century. Commercial law appears to have had its beginnings in the internal codes of conduct in fraternal orders of guild merchants; those who did not live up to them were threatened with ostracism. More specialized the law merchant (mercantile law) evolved and came to govern most commercial transactions throughout large areas of Europe. A uniform set of standards was conveyed through long distance trade codes of conduct, so that Pisan laws passed into the sea codes of Marseilles. Oleron and Lubeck

gave laws to the north of Europe, Barcelona to the south of Europe, while from Italy came the legal principle of insurance and bills of exchange. (Mitchell, 1969, p. 156)

The development of more sophisticated accounting methods and the use of such methods and of notarial records for evidence in disputes permitted evidence to become the basis for ascertaining facts in disputes. The gradual blending of the voluntaristic structure of enforcement of contracts via internal merchant organizations with coercive enforcement by the state is an important part of the story of increasing the enforceability of contracts. The long evolution of merchant law from its voluntary beginnings and the differences in resolutions that it had with both the Common and Roman law are a part of the story. The three types of law did not accommodate each other very well to begin with. This was particulary true in cases of moral hazard and assymetric information in insurance contracts as well as those associated with fraud in exchange. The law merchant was assumed by the court of common law but continued to be administered in the original spirit of the law merchant, that is as a law based on custom. At first it still applied only to proven merchants, whether they were the plaintiff or defendant. Cases seldom laid down a particular rule because it was virtually impossible to separate custom from the facts. The habit was to leave the jury with the custom and the facts and the judge would charge the jury to determine and apply the custom when supported by the facts.[2]

The law merchant, besides providing a much needed court of law especially suited to the unique needs of the merchant, also fostered some significant developments which aided in decreasing transactions costs of exchange. Among such developments can be included the recognition of the responsibility of the principal for his agent. This spawned both a benefit and a cost. It allowed the merchant to expand his scope of operation via a series of agents. At the same time it created a principal-agent problem. Initially this legal recognition was in effect only for well known agents of the principal. The fact that credit was generally given to the agent because it was generally believed he was acting for his master provided an obvious opportunity for the agent to benefit himself. At the same time, however, the privilege was also used to control the principal-agent problem. By extending to his agent the privilege of using the merchant`s credit for his own personal trading, the merchant was able to increase the opportunity cost to the agent of losing his position. If the agent abused his position, he would lose not only his job, but a valuable line of credit as well.

The effect of the merchant law on contracts and sales was especially encouraging to the expansion of trade. The existing Roman and Germanic laws did not give the security and certainty of bargains to merchants that was needed. Neither body of law protected them against the claims of the original owner of stolen or lost goods that the merchant had innocently purchased. The feudal lord recognized the value of fairs and markets as a revenue source and therefore the importance of protecting the honest purchaser. Under merchant law, the honest purchaser was allowed either to keep the goods or return them if the original owner refunded the purchase price.

Protection of the bona fide purchaser was not a part of the common law. However in commercial disputes the "good faith" principle was used earlier and on a much wider

---

[2]. For a game theoretic analysis of the role of the law merchant in solving organizational problems of exchange see Milgrom, North, and Weingast, 1991.

scope--indeed the basis of Roman contract law by 200 A.D. It evolved first out of the Fair Bonds, which validated sales at fairs by affixing a seal to the bond. Originally this was a voluntary measure--the custom of fairs allowed debts to be contracted by witness. Eventually though, the desire to avoid fraud and at the same time increase revenue led to a law requiring that all sales be recognized by a sealed bond. Once sealed, the bond could be invalidated only by proving that the seal had been forged.

The good faith principle was extended to the area of insurance. Extreme good faith was required when writing out a marine insurance contract. Because the person wishing insurance had more knowledge he must tell the underwriter the "whole truth and nothing but the truth". The law required this extreme good faith or the contract would be invalidated. Misrepresentation was a sufficient reason, even when not intended, to invalidate the contract, as opposed to ordinary contracts where intent to defraud was necessary in order to invalidate a contract.

A major player in this evolution of markets was the state, and there was continuous interplay between the fiscal needs of the state and its credibility in its relationships with merchants and the citizenry in general. In particular, the evolution of capital markets was critically influenced by the policies of the state, since, to the extent that the state was bound by commitments that it would not confiscate assets or in any way use its coercive power to increase uncertainty in exchange, it made possible the evolution of financial institutions and the creation of more efficient capital markets. The shackling of arbitrary behavior of rulers and the development of impersonal rules that successfully bound both the state and voluntary organizations were a key part of this whole process.[3] The development of an institutional process by which government debt could be circulated, become a part of a regular capital market, and be funded by regular sources of taxation was another. (Tracy, 1986).

It was in the Netherlands, and Amsterdam specifically, that these diverse innovations and institutions were put together to create the predecessor of the efficient modern set of markets that make possible the growth of exchange and commerce. An open immigration policy attracted businessmen; and efficient methods of financing long distance trade were developed, as were capital markets and discounting methods in financial houses that lowered the costs of underwriting this trade. The development of techniques for spreading risk and transforming uncertainty into actuarial, ascertainable risks as well as the creation of large scale markets that allowed for lowering the costs of information, and the development of negotiable government indebtedness all were a part of this story (Barbour, 1949).

III[4]

Modern productivity growth is a consequence of the second economic revolution--a revolution which began in the last half of the nineteenth century and is still continuing.[5]

---

[3]. North and Weingast, (1989) have an analysis of the transformation of the British polity.

[4]. This section draws from sections II and III of a paper entitled "Institutions, Transaction Costs, and Productivity in the Long Run" I presented at the Eighth World Productivity Congress in Stockholm in 1993

That revolution is the development of the scientific disciplines and the wedding of science and technology. It is a revolution because it is a fundamental change in the stock (and flow) of knowledge, which entails an equally fundamental change in the organization of human beings and the structure of societies to realize the productivity implications.

The development of the disciplines of physics, chemistry, biology, and genetics is a fundamental source of the growth in the stock of scientific knowledge. The systematic application of these disciplines to the basic economic problem of scarcity has not only purged the Malthusian spectre of diminishing returns from our purview but has created the vision of a potential world of plenty. To achieve that potential, however, entails a restructuring of economic, social, and political institutions and organizations in order to realize the increasing returns attributes of the technology in which this scientific knowledge is embodied.

The technology requires occupational and territorial specialization and division of labor on an unprecedented scale and in consequence the number of exchanges grows exponentially. In order to realize the gains from the productive potential associated with a technology of increasing returns one has to invest enormous resources in transacting. In the United States,for example, the labor force grew from 29 million to 80 million between 1900 and 1970; during that period non-agricultural production workers grew from 10 million to 29 million, while white collar workers (the great majority of whom are engaged in transacting) increased from 5 million to 38 million. The transaction sector (that part of transaction costs that goes through the market and therefore can be measured) in the United States in 1970 made up 45 percent of GNP.[6]

If transaction costs were simply the costs of coordinating the increasingly complex interdependent parts of an economy they would be simply information costs or more specifically the costs of acquiring the information to measure the multiple dimensions of what is being exchanged. But they are also the costs of enforcing agreements and making credible commitments across time and space, necessary to realize the potential of this technology.[7] Let me briefly elaborate some of the measurement and enforcement problems that account for the size of the transaction sector. Necessary to be able to realize the gains of a world of specialization are control over quality in the lengthening production chain and a solution to the problems of increasingly costly principal/agent relationships. Much technology indeed is designed to reduce transaction costs by substituting capital for labor or by reducing the degrees of freedom of the worker in the

---

[5]. The first economic revolution was the development of agriculture in the eighth millenium B.C. This section is further elaborated in Ch. 13, "The Second Economic Revolution" in North, 1981. I neglect in this essay discussion of a major source of pre-modern (and in less developed economies, modern) productivity growth stemming from improvements in health and nutrition. See Fogel (1991)

[6]. John Wallis and Douglass North, "Measuring the Transaction Sector in the American Economy" in S.L. Engerman and R.E. Gallman, eds. Long Term Factors in American Economic Growth, Chicago, University of Chicago Press, 1986

[7]. The recent essay by Becker and Murphy (1992) emphasizes the coordination costs associated with capturing the gains from trade but ignores the enforcement costs.

production process and by automatically measuring the quality of intermediate goods. An underlying problem is that of measuring inputs and outputs so that one can ascertain the contribution of individual factors and the output at successive stages of production. For inputs there is no agreed upon measure of the contribution of an individual input. There is room for conflict over the consequent payment to factors of production. For output, not only is there residual unpriced output, that is waste and pollutants, but also there are complicated costs of specifying the desired properties of the goods and services produced at each stage in the production process.

Firms using the technology of the second economic revolution have large fixed capital investments with a long life and (frequently) low alternative scrap value. As a result the exchange process embodied in contracts has to be extended over long periods of time, which entails uncertainty about prices and costs and the possibility of opportunistic behavior on the part of one of the parties to the exchange. A number of organizational problems emerge from the use of this new technology.

First, increased resources are necessary to measure the quality of output or the performance of agents. Sorting, grading, labeling, trade marks, warranties, licensing, time and motion studies and a variety of other techniques to measure the performance of agents are all, albeit costly and imperfect, devices to measure the characteristics of goods and services and the performance of agents. Despite the existence of such devices the dissipation of income is evident in the difficulties of measuring the quality of automobile repairs, of evaluating the safety characteristics of products or the quality of medical services, or of measuring educational output. The problems of evaluating performance are even more acute in hierarchies because of the difficulties of achieving low cost measurement of the multiple dimensions of the agent's performance.

Second, while team production permits economies of scale to be realized, it has done so at the cost of worker alienation and shirking. The "discipline" of the factory is a response to the control problems of coordination and shirking in team production. From the perspective of the employer the discipline consists of rules, regulations, incentives, and punishments essential to effective performance. Innovations such as time and motion studies are methods of measuring individual performance. From the viewpoint of the worker they are frequently viewed as inhuman devices to foster speedups and exploitation. Since there is no agreed upon measure of output that constitutes contract performance, both are right.

Third, the potential gains from opportunistic behavior increase and lead to strategic behavior both within the firm (labor-employer relations, for example) and in contractual behavior between firms. Everywhere in factor and product markets the gains from withholding services or altering the terms of agreement at strategic points offer large potential gains.

Fourth, the development of large scale hierarchies produces the familiar problems of bureaucracy. The multiplication of rules and regulations inside large organizations to control shirking and principal/agent problems results in rigidities, income dissipation, and the loss of flexibility essential to adaptive efficiency.

Finally there are external effects: the unpriced costs reflected in the modern environmental crisis. The interdependence of a world of specialization and division of labor increases exponentially the imposition of costs on third parties.

It is the institutional framework and the consequent costs of transacting (a product of both the institutional framework and the technology employed) that determine the degree of success of an economy in overcoming these obstacles to realizing the productivity implications of any given technology. Well specified and enforced property rights, decentralized political and economic decision making, and effective competition have been the underlying institutional structure for the organizational changes allowing modern economies to reap the productivity gains of this technology.

But the economic restructuring itself entails a still more fundamental restructuring of the entire society in order to create efficient economic markets This technology and accompanying scale economies entail specialization, minute division of labor, impersonal exchange, and urban societies. Uprooted are all the old informal constraints built around the family, personal relationships, and repetitive individual exchanges. Indeed the basic traditional functions of the family--education, employment (the family enterprise), and insurance--are either eliminated or severely circumscribed. New formal rules and organizations and an increased role of government replace them.

The contention of Marxists was that the problems that resulted from this restructuring were a consequence of capitalism and that the inherent contradictions between the new technology and the consequent organization of capitalism would lead to its demise. The Marxists were wrong that the problems were a consequence of capitalism; they are ubiquitous to any society that attempts to adopt the technology of the second economic revolution. However, as the foregoing paragraphs have attempted to make clear, Marxists were right in viewing the tension arising between the new technology and organization as a fundamental dilemma.

It is surely one of the great ironies of history that Karl Marx, who first pointed out the necessity of restructuring economic and political institutions and organizations if a society was to realize the potential of a new technology, should have been responsible for the creation of societies that foundered on that specific issue. The demise of the centrally controlled economies of central and eastern Europe is a sober testimonial to the inability of economies to effectively reap the benefits of the technology of the second economic revolution with an institutional framework of centralized bureaucratic control of the economy.

For the less developed countries the radical alteration in economic, social, and political institutions and organizations that is necessary to realize the potential of the second economic revolution has in some cases been an insuperable hurdle. Indeed many third world countries have realized many of the social costs and tensions that have accompanied this revolution with few of the benefits.

These tensions have only partially been resolved in the market economies of the western world. The growth of government, the disintegration of the family, the incentive- incompatibility problems in many modern political and economic hierarchical organizations are all symptoms of the consequent problems besetting western economies.

However, the relative flexibility of the institutions of the western world--both economic and political--has mitigated these problems. Adaptive efficiency,[8] while far

---

[8]. For an elaboration of the concept of adaptive--in contrast to allocative--efficiency and the implications for economic performance through time see Pelikan, 1987

from perfect in the western world, accounts for the degree of success that such institutions have experienced. The basic institutional framework has encouraged the development of political and economic organizations that have replaced (however imperfectly) the traditional functions of the family; mitigated the insecurity associated with a world of specialization; evolved flexible economic organization that has encouraged worker participation, reduced shirking and induced low cost transacting; resolved some of the incentive incompatibilities of hierarchies and encouraged creative entrepreneurial talent; and tackled (again very imperfectly) the external effects that are not only environmental but also social in an urban world.

IV

A fundamental limitation of the traditional economics approach to the sources of productivity growth is that it fails to put the problems of growth and productivity in a dynamic institutional context. Even Baumol's study (1986) which has the virtue of "an appreciation of history" to use Jeffrey Williamson's felicitous phrase (J. Williamson 1991) misses the key element: that is, the growing size of the transaction sector that has been an integral part of the long run productivity increase of the American economy. Between 1870 and 1970 the transaction sector of the American economy grew from 25 to 45 percent of GNP (Wallis and North, 1986). A substantial part of that growth is accounted for by the increasing coordination costs associated with an economy of ever increasing division of labor. The development of specialized sectors of the economy to provide transaction services such as banking, finance, and insurance as well as the creation of specialized departments inside firms (such as accounting, legal, and finance departments) was an essential part of the coordination process. The largest source of growth in the transaction sector between 1870 and 1970 came within firms. In 1870 only 10 percent of the transaction sector, 2 percent of GNP, was within firms in the non-transaction industries (industries other than wholesale and retail trade; finance, insurance, and real estate; and government). In 1970 over one third of the transaction sector, or 10 percent of GNP, was within those firms(Wallis and North, 1986). By now we have a substantial literature detailing the institutional and organizational changes--the supply side changes--that made the productivity growth possible from this new technology (Chandler, 1977, North 1981, O. Williamson, 1985, Nelson, 1981 and Nelson and Wright, 1992). But that is not all. The growth of the transaction sector has also been a result of the growth of rules and regulations that have evolved to deal with every facet of economic activity in the past century. They are a mixed bag of those that reduce transaction costs per exchange and those that raise transaction costs per exchange. Those that have reduced transaction costs have stemmed from well- specified and enforced property rights that provide incentives for productivity-enhancing behavior--both directly in terms of patent and trade secret laws, for example, and indirectly in terms of incentives to acquire knowledge--decentralized economic and political decision making, and effective competition in both economic and political markets.

But what have been the consequences of a century of the Interstate Commerce Commission on the productivity of United States transport industries? Do the laws with respect to competitition in the communications industries in the context of the radical changes in technology have no effect on the rate of productivity growth? What are the productivity implications of unitizing oil fields? Has the growth of environmental

legislation had no effect on productivity growth? Does the growth of legislation dealing with safety, working conditions, seniority rules, etc. of the labor force influence productivity? Does the development of transfer payments affect incentives?

The frictionless, a-political framework of neo-classical economics does not lend itself to providing an integrated view of how economies operate. If pressed, most economists would agree that institutions make a difference but would not go beyond a gross distinction between a market economy and a planned economy or maybe a third world economy. If they did want to explore the institutional features of a market economy it would be in terms of rent seeking, which is described as though it were some kind of a disease that can afflict only such economies.[9] It's not a disease--it's a feature of the incentive structure of any economy. It is not possible to create an institutional framework with unidirectional incentives all favoring efficiency.[10] Even if it were economically possible, it would not be politically possible. There is no such thing as a neutral set of rules and their enforcement in an economy such as would be inferred from the use of the term, a laissez-faire economy. Everywhere and at all times the rules of the game are a mixed bag of incentives and disincentives with respect to their implications for productivity. And they are continually evolving. There is a complex on-going interrelationship between institutions, organizations, learning, the perceptions of the actors, and the choices they make.

The high pay-offs were embodied in the perceptions of entrepreneurs that the application of new scientific knowledge to solving economic problems would be fruitful but equally that the framework of rules and laws would be conducive to high rewards for those who were successful in harnessing the new knowledge to technology. The pioneering discipline was chemistry, which was applied to a wide range of industrial activities in the nineteenth century. It was German scientists who took the lead and German entrepreneurs who were the first to integrate chemical laboratories with chemical manufacturing(Rosenberg and Birdzell, 1986, and Mokyr, 1990). The systematic application of science to technology spread into every aspect of economic activity wherever the basic institutional framework provided a hospitible environment. The demand for educational investment and the growth of universities as the principal source of scientific knowledge was still another consequence. The restructuring of economic organizations to take advantage of this technology has been eloquently described by Chandler (1977). If these institutional changes were the whole story then our story would be one of sustained productivity expansion unleavened by other considerations.

An integral part of the second economic revolution, however, was a transformation into an urban, interdependent world with all the social problems I have briefly alluded to. In consequence that revolution is equally a social and political revolution. Its political and economic origins were observable in the late nineteenth century with the articulation of the discontent of those groups that perceived (rightly or

---

[9]. The study by Baumol et al (1989) concludes with a discussion of rent seeking as a potential source of productivity decline in the years ahead.

[10]. The study by Miller (1991) is an effective argument about the impossibility of creating unidirectional incentives in heirarchies.

wrongly) that they had not shared in the improved economic well-being. In the United States it was organizations such as the Knights of Labor, the Granger movement, the American Federation of Labor, the IWW, the Populist party that pressed for institutional change; change that was in part adopted by the Democratic party and eventually embodied in New Deal legislation and (at least partly as a result of the threat of court-packing by Franklin Roosevelt) judicial changes. The resulting increase in the size and influence of government altered the relative gains to be realized by organizations as between economic and political policies. The growth of government; the increased insecurity associated with an interdependent world; the decline of the role of the family; and the external diseconomies (both environmental and social) all have led to the rise of new interest groups and ongoing institutional changes that are an integral part of the world we live in. While the particular timing and characteristics of the United States institutional change have been uniquely a feature of the path-dependent pattern of its institutional/organizational structure, the overall pattern of institutional changes embodied in the growth of government, transfer payments, and government intervention in economies has been a universal feature of modern economies.

While these social and political changes have been the subjects of a substantial literature, there has not, to my knowledge, been any successful effort to systematically integrate these demand-side institutional changes into the overall analysis of long run productivity change. It is not simply a consequence of neo-classical economics ignoring institutions; it is equally a consequence of the difficulty of developing macro-level data to measure their impact.[12] It is one thing to measure the direct impact of environmental or regulatory legislation on transformation or production costs. Christainsen and Haveman suggest that federal regulations are responsible for from 12 to 21 percent of the slowdown in the growth of labor productivity in U.S. manufacturing during 1973-77 as compared to 1958-65 (1981, p 324). It is something else to measure the indirect impact in terms of increasing uncertainty and hence foregone choices. There is no easy way to measure the foregone production as a result of the time, costs, and increased uncertainty arising from the rules and regulations that today govern every aspect of production and exchange. One only measures transaction costs of existing economic transactions; unmeasured are production and exchange that do not occur because transaction costs are so great as to foreclose such economic activity. Yet to ignore these foregone opportunities because of the difficulties of measurement is to miss an essential aspect of the growth process. It is the exchanges that don't occur because of the high cost of transacting (and therefore producing) that are the real underlying source of poverty of third world economies and obviously critical in successfully analyzing sources of productivity change in modern economies.

V

Let me conclude this essay by suggesting ways by which institutions and transaction costs have influenced and continue to influence productivity change.

1. The most important insight on this subject is to recognize the critical facilitating role of a favorable institutional environment--political and economic--for creating the necessary conditions that made possible the second economic revolution. Creating such an

---

[12]. Denison (1979) does attempt to examine the direct impact of such legislation

environment is the most fundamental initial condition for the productive transformation of third world and eastern European economies.

2. There is a complex interdependence of institutional and technological change as an economy evolves. This subject is much too large to be dealt with here but I just mention several implications. The firm is concerned to minimize total costs, the sum of transformation costs and transaction costs. Minimizing total costs may result from a technological change that reduces transformation costs or from an institutional change that reduces transaction costs. But it may also occur by a technological change whose main impact is to lower transaction costs--development of the telephone, for example--or by an institutional change whose main impact is to lower transformation costs--unitizing an oil field, for example. Moreover, the productivity gains from the second economic revolution were sometimes initiated by technological changes which induced institutional change and sometimes by institutional changes which induced technological change.

3. The growth of government and government regulation has surely had, and continues to have, a major impact on productivity even if we have no direct way of measuring its impact. While that impact would appear to be negative a more comprehensive measure of welfare might well modify that initial presumption if, in fact, regulatory and environmental legislation has had a positive effect on such a broader measure of welfare.

4. One consequence of the growth of government and government regulation has had an unambiguously negative effect on productivity and that is a shift in the locus of decision making of economic organizations from economic to political policies as the relative gains from using the political process have increased. It is this rent-seeking feature of institutional change that has received the attention of economists.

5. As economies have grown richer, they have increased the proportion of consumption of goods and, particularly, services--medical services being a conspicuous example-- with high transaction costs. The growth of the service sector of economies as they become richer has reflected this income elastic demand and has probably been a drag on productivity increase. This is another way of saying that the coordination costs of economies may increase as they become richer. However the advent of the computer should be a powerful force to help reduce transaction costs. The fact that it does not appear to have had a major influence on productivity, so far, has been a puzzle to economists concerned with the slowdown in productivity growth. Paul David, using the dynamo as an historical analogy, shows that the productivity consequences of that innovation were not realized until much later. In fact the striking advances in total factor productivity do not occur until the 1920s.[13] David attributes the delay to the necessity of making complementary redesigns of the entire factory structure before the full productivity gains would be realized. I would add the complementary changes in institutions and organizations that would also have to occur as briefly alluded to in point 2 above. Recent evidence suggests we may be on the verge of realizing substantial productivity increase from the restructuring and reorganization of firms that has been made possible by the

---

[13]. David discusses the literature on the apparent paradox in "Computer and Dynamo: The Modern Productivity Paradox in a not-too-Distant Mirror" (1990)

investment in data processing equipment.[14] Whether that "re-engineering" will be sufficient to overcome the overall possible rising costs of coordination is an important task of future research.

We have only just begun the serious study of institutions and of the interaction of institutional analysis with more traditional economic analysis. The objective of this essay has been to suggest a redirection of research on productivity growth to a synthesis integrating institutional analysis with the traditional analysis of sources of productivity change contained in the growth accounting literature.

[14]. See for example the Wall Street Journal, March 16, 1993 on the "Price of Progress".