Cambridge University Press
978-0-521-08637-0 - Institutional Change and American Economic Growth
Lance E. Davis and Douglass C. North
Excerpt
More information

PART ONE

THE THEORY DEVELOPED

CHAPTER I

A THEORY OF INSTITUTIONAL CHANGE: CONCEPTS AND CAUSES[1]

(I) Introduction

Historians have traditionally displayed an interest in the institutions within which human action occurs, and much of their work has involved an examination of the interaction between people and these institutions. Economic historians, especially the 'new' group, have, on the other hand, focused their efforts on economically rational behavior as an explanation of past events; institutions have been taken as given, and the 'antiquarian' interests of the more traditional historians have sometimes been scorned.[2] Perhaps because of their concern with long-run change, traditional historians have recognized that institutions do have something to do with the speed and pattern of economic growth (a relationship that was obvious to them but one that economists have only gradually perceived). Much of written history is devoted to the study of the evolution and development of political, military, and social institutions; and just as these sophisticated institutions have evolved through history, so have complex economic institutions emerged to provide a part of the framework within which a highly technical society can survive and flourish. While there are few pieces of history that do not lean heavily upon some form of theory, unfortunately, there has been little theory to help understand the phenomena of institutional change. In the absence of such theory, history is limited to narration, classification, and description. There are relatively few historians who would willingly accept such a limitation.

[1] This chapter draws heavily on the work of J. Buchanan and G. Tullock, *The Calculus of Consent* (Ann Arbor, Mich., 1962); K. Arrow, 'Political and Economic Evaluations of Social Effect and Externalities', paper given at the NBER Conference on the Economics of Public Output, April 1968; and A. Downs, *An Economic Theory of Democracy* (New York, 1957).

[2] For a discussion of the 'new' economic history, its method, proponents, and critics, see either L. Davis, 'And It Will Never Be Literature', *Explorations in Economic History*, 2nd ser., vol. 6, Fall 1968, or R. Fogel, 'The New Economic History: Its Findings and Methods', *Economic History Review*, December 1968.

If the historian's explanation of the process of economic development has been less insightful than one might wish, a substantial part of the blame must rest on the set of blueprints of the causal structure that the economists have provided. The best of the historian's work has all too often been rooted not in sound logical deductions from explicit premises but in brilliant historical intuition. This triumph of intuition over mathematics rests not on the blind refusal of the historian to kneel at the altar of science, but on the fact that the theories that he might have used have been poorly specified, totally irrelevant, and, at times, marked by errors in logic. Until 'better' theories are created, no one can blame him for depending on the intuition that has served him so well in the past.

It is unlikely, however, that these theories will spring full-grown from the forehead of some ivory-towered theorist. It is more probable that theories capable of predicting the future and explaining the past will emerge in bits and pieces from some interaction between the theorist who worries about logic and the historian concerned with explaining the past. In his search for a theory that explains the process of economic evolution, the scholar must continually move from theory to fact and back again to theory.

This book is a 'day-by-day' account of an intellectual journey through American economic history. The journey was planned to provide a description of the processes that have produced the present structure of economic institutions. That description, in turn, is the basis for a first (and very primitive) attempt at the formulation of a specified, relevant, and logical theory of the birth, growth, mutation, and, perhaps, death of these institutions. The book is a study of the sources of institutional change in American history. It is specifically concerned with the relationship between economic organization and economic growth, but it is no more than a preliminary study. The theory is at some points woefully weak and the explanations at times incredibly simplistic. The book does, however, represent a first step towards a useful theory of economic growth, and it does provide some new interpretations of the American economic experience.

Since the book is written for historians (as well as economists), it may be well to digress briefly on the role of models and theories in the writing of history. Although it is technically inaccurate to do so, for simplicity's sake we will use the words 'model' and 'theory' interchangeably. They will both refer to a logical structure that relates a set of assumptions to a certain set of conclusions. In economics, it is initially assumed that a firm attempts to maximize its profits and that it is constrained in its production possibilities by a combination of its technological capabilities and existing resources, and in its sales opportunities by certain market conditions. From these assumptions it follows that if it pays a firm to produce at all, it will choose to operate at a level that will return the greatest possible profit – i.e. where the difference between total revenue and total cost is the greatest. This assertion is only a logical deduction, and like any such conclusion it has predictive or explanatory power in the real world only if, in addition to being logically valid, its initial conditions are met. It is under these circumstances that the theory is said to be 'operational'. If, for example, we were attempting to explain the production decision of a Soviet firm whose goals were the maximization of output, rather than profit, the theory would not be very useful.

Even if the theory is conjoined with a relevant set of initial conditions and is in principle operational, the historian must realize that the 'laws' (i.e. predictive or explanatory statements) that can be derived from the theory are probabilistic, not mechanical. In the same way that a physicist cannot predict the behavior of an individual sub-atomic particle but is quite able to predict the average behavior of large groups of such particles, the economic theorist can predict the behavior of typical firms and consumers but cannot make meaningful predictions about the behavior of single decision-making units.

If the model is to be completely useful, ideally it ought to be able to predict two kinds of things:

(1) Given any established set of institutions and some disequilibrating force, the model ought to predict whether the newly emerging institutions will be purely individual (i.e. involve only a single decision maker), depend upon some form of voluntary cooperation, or rely on the coercive power of government.

(2) It should provide some estimate of the period of time that is likely to elapse between the initiating disequilibrium and the establishment of the new (or mutated) institutions.

In the remainder of Part One (Chapters 1–4), we attempt to spell out the model and the initial condition in their simplest form. In Part Two (Chapters 5–10), the model is applied to a number of facets of the American economy, and its ability to explain institutional developments in those sectors is evaluated. In the final part (Chapters 11–12), we summarize the impact that institutional innovation has had on the public–private mix and attempt to reformulate the model along the lines suggested by the experience of Part Two. This particular form of exposition was chosen to make it easier for the reader to follow the argument, as well as to show him why the modifications were important.

This book, then, attempts to specify a theory of institutional change and to apply that theory to certain facets of American development. It is hoped that the theory will contribute to our understanding of that process and that such an experiment might make it possible to modify the model so that in the future it can be used to explain change in certain non-economic institutions and in certain non-American environments. The model has been formulated in a manner that makes it in principle operational, although, like many models in the social sciences, it predicts much less than we would like. As our narrative unfolds it will become increasingly clear that the model yields particularly poor results: when the potential gains and losses are large and relatively equal, but received and/or incurred by different groups; when the predictions involve a mixed result – an institution that is not purely public or purely private; when the fundamental legal and social rules that govern economic and political behavior are altered by the 'predicted' changes in the institutional structure. Despite these qualifications and limitations, we argue that the exercise is worthwhile. It focuses attention on the need for some theory of institutional change if we are ever to have a useful theory of economic growth; and, even in its present crude form, the model has, we feel, allowed us to take a new and productive look at certain aspects of the American historical experience.

(II) Some definitions

While Humpty Dumpty was obviously correct when he said, 'When I use a word it means just what I choose it to mean', Alice also had a point when she complained that words should not mean so many different things. In deference to Alice, it seems appropriate to define some terms as we intend to use them in the remainder of this study and to draw some distinctions between concepts that are sometimes lumped together.

(1) The *institutional environment* is the set of fundamental political, social, and legal ground rules that establishes the basis for production, exchange, and distribution. Rules governing elections, property rights, and the right of contract are examples of the type of ground rules that make up the economic environment. In the American economy, the environment is established by: a written document, the Constitution, and the interpretations that the judiciary have placed on it in decisions dating back to the earliest years of the Republic; and the views of the nation's citizens about the type of institutions that they prefer.[1]

The environment can, of course, be altered. In the context of the American

[1] In other countries the environment is set in other ways and different rules for amendment apply. In the United Kingdom, for example, there is no written constitution, and court decisions alone establish the rules of the game. In a totalitarian country, the rules are established by political fiat and altered by decree. We distinguish between fundamental legal ground rules and other types of legislation. The former we treat as exogenous; the latter as endogenous. However, we admit that the line between the two is not always clear cut, and to that extent the analysis contains an element of ambiguity.

legal structure such changes can come from an amendment to the Constitution either by political action or a change in judicial interpretation, or from a shift in citizens' preferences. Thus, for example, property rights were fundamentally altered both by the Thirteenth Amendment and by the court decision in the case of *Ogden* v. *Saunders*. Similarly, election rules have been changed both by constitutional amendment (the Fifteenth, for example) and by judicial reinterpretation. (The twin decisions in *Baker* v. *Carr* and *Reynolds* v. *Sims* are cases in point.) In this study we make no attempt to explain changes in the economic environment. Such changes have certainly occurred and any study of their causes would be interesting. They are, however, exogenous to this model of institutional innovation.[1]

(2) An *institutional arrangement* is an arrangement between economic units that govern the ways in which these units can cooperate and/or compete. The institutional arrangement is probably the closest counterpart of the most popular use of the term 'institution'. The arrangement may be either a formal or an informal one, and it may be temporary or long-lived. It must, however, be designed to accomplish *at least one* of the following goals: to provide a structure within which its members can cooperate to obtain some added income that is not available outside that structure; or to provide a mechanism that can effect a change in laws or property rights designed to alter the permissible ways that individuals (or groups) can legally compete.

The laws legalizing the corporation provide an example of an institutional arrangement that is designed to accomplish both of these ends. The corporate form provides an organizational structure that makes it possible for management to control a much larger and more diverse set of economic activities than could be effectively directed within a more primitive organizational form; and it gives the organization itself legal life. A business so constituted can, therefore, compete in areas that are closed to other types of organizations.

The arrangement may involve a single individual, a group of individuals voluntarily cooperating together, or the government (alone or in cooperation with one or more individuals). The last mentioned innovation frequently implies some legal change, but the first two, while resting on the legal structure that constitutes the environment, involve only the private sector directly; and it is possible that innovation could occur without a change in the law.[2] It is the process of innovation of these institutional arrangements

[1] A survey of changes in the environment is presented in Chapter 4.

[2] The government can act, for example, through executive fiat (as the President did in ordering black employment in the defense industries) and in these cases one might say no legal change has occurred.

that the model of 'institutional change' is designed to predict – specifically it is designed to predict their level (individual, voluntary cooperative, or governmental) and the timing of their emergence.

(3) A *primary action group* is a decision-making unit whose decisions govern the process of arrangemental innovation. The unit may be a single individual or a group of individuals, but it is the action group that recognizes there exists some income – income that their members are not presently receiving – that *they* could accrue, if only they could alter the arrangemental structure. At least one member of any primary action group is an innovating entrepreneur in the Schumpeterian sense, and within the context of this model the group initiates the process of arrangemental innovation. The action group always increases its income if its innovation survives the test of competition. The group pays a portion of the innovation costs, but it may or may not have to bear all or part of the operating costs of the new arrangement (if, in fact, there are operating costs).

The New York manufacturers who informally banded together in 1811 to lobby for the passage of a general incorporation law provide an example of a successful primary action group. They saw that income could be earned if easy incorporation were possible, they paid the costs involved in pushing the revised law through the legislature, and they reaped the profits from their innovation.[1]

(4) A *secondary action group* is a decision-making unit that has been established by some change in the institutional arrangement to help effect the capture of income for the primary action group. The secondary action group makes some of the tactical decisions that bring about the capture, but it does not accrue all of the additional income (it may, in fact, quite likely accrue none).

If the New York law had established the office of Commissioner of Corporations charged with the task of receiving, reviewing, and approving applications for corporate charters, the commissioner (together with his staff) would have constituted a secondary action group. In the normal course of events the secondary action group might accrue none of the income arising from the innovation, but if the law granted them some discretionary powers, they might be able to effect a transfer of a portion of that extra income from the primary action group to themselves. If one wishes, it is possible to view the American tradition of bribing public officials as an arrangement

[1] Although the history is not clear, it is possible that the lobbyists might have effected an even more profitable innovation had they been able to convince the legislature to pass a *temporary* general incorporation law. However, the costs of such an arrangement may have been prohibitive either because of the legislative response or because of the rules laid out in the institutional environment.

designed to redistribute income between primary and secondary action groups.

(5) *Institutional instruments* are documents or devices employed by action groups to effect the capture of income external to the existing arrangemental structures when those instruments are applied within the new arrangemental structure. The corporate charter granted to a manufacturing company under the New York general incorporation act is an example of an institutional instrument.

The arrangement, if it is a governmental one, will directly involve the coercive power of government; if it is a voluntary one, it may have underlying it the coercive power of an existing structure of property rights. The effectiveness of the instruments may depend on some fundamental legal concept that is part of the economic environment. An officer of a corporation may sign a contract and in so doing effect a decision to buy a machine. While the contract once signed can be enforced in the courts, the enforcement power does not rest with the institutional arrangement, but depends upon some fundamental constitutional rule.

In an attempt to make the reader more familiar with these definitions, consider for a moment the case of a factory that produces smog as well as products. The smoke is part of the production process; it would be costly to eradicate it, but the people living near the factory find it very disagreeable. Assume that the real cost to them of the smoke (as measured by the amount they would be willing to pay to eliminate it) is greater than the cost the factory owner would have to incur if he were to install a smog control device. Clearly total income could be increased if the smog were eliminated; however, it may well be that there is no way the bargain can be effected within the existing institutional arrangement (where the costs of the smoke accrue to one group, the costs of elimination to another). This problem is typical of those faced by residents of most every city in the United States, and often it appears that some type of government institution should be innovated to effect the smoke abatement.[1] At least two alternatives are open to those who seek the additional income. They could band together to form a political coalition (a *primary action group*), and, if successful at the polls, they (or their representatives) could enact a law (an *institutional arrangement*) that prohibited the factory from emitting smoke. Alternatively the successful political coalition could underwrite legislation establishing a zoning board (a *secondary action group*), and that board could, in turn, issue a cease and

[1] The explanation for the choice of government as opposed to some other form of institution is the thrust of the argument in the next three chapters. Here for simplicity we merely assume that that choice is 'best' in some sense.

desist order against excessive air pollution.[1] The cease and desist order is an *institutional instrument* backed by the coercive power of the government. Either plan, however, depends on an *economic environment* within which it is possible by political action to abrogate certain of the 'rights of private property'. If the fundamental rules of society prohibited such interference (as in fact they did in the United States until the late nineteenth century), either arrangement (and its complement of instruments and secondary action groups) would be ruled out unless (or until) the rules were changed.

(III) A theory of institutional innovation: a first approximation

Economic institutions and property rights are assigned distinct and constant values in most economic models, but in the study of long-term economic growth, these values are always subject to fundamental change. We postulate that economic institutions are innovated or property rights are revised because it appears desirable for individuals or groups to undertake the costs of such changes; they hope to capture some profit which is unattainable under the old arrangement.

An institutional arrangement will be innovated if the expected net gains exceed the expected costs.[2] Only when this condition is met would we expect to find attempts being made to alter the existing structure of institutions and property rights within a society. For example, if production can be carried on more cheaply by large firms than small, it may be cheaper for a corporation to operate than for a sole proprietorship; if prices differ widely between two markets, it may be profitable to organize a third market to move goods from the low-priced to the high-priced market; if theft and brigandage are widespread, the creation of an efficient police force will raise the value of private property. If an entrepreneur contemplates the construction of a dam designed to produce hydroelectric power that also reduces flood damage downstream, the builder might appropriate a share of these benefits by the prior purchase of some of the downstream property. On the other hand, he might appeal to government to impose a tax on the downstream beneficiaries to help subsidize his construction costs.

As to form, arrangements can range from purely voluntary to totally government controlled and operated. Between these extremes exists a wide

[1] Alternatively they could band together to pay off the factory owner. As we shall see, the alternative of choice depends, like the choice between government and private arrangement, on the costs incurred and the revenues accruing to each.

[2] To avoid confusion we will use the term 'innovation' to refer to any change in the institutional technology embodied in an institutional arrangement and the term '*new* innovation' to the first such application.

variety of semi-voluntary, semi-government structures. Voluntary arrangements are simply cooperative arrangements between consenting individuals, and any individual can legally withdraw.[1] This ability implies, of course, that decisions must be unanimous, as the costs of acceding to the decisions are less than those incurred by withdrawal. Government arrangements, on the other hand, do not provide the withdrawal option. Action, therefore, does not require unanimous consent but only conformity with some decision rule.[2] In a democracy, for example, a simple majority frequently determines the course of action.

Both voluntary and government arrangements have been innovated to realize economies of scale, gains from transaction costs, internalization of externalities, the reduction of risk, and the redistribution of income.[3] For example, the corporation has aided the realization of the benefits of the economies of scale sometimes inherent in large-scale operations, and the TVA has yielded similar benefits from power generation and distribution. The stock exchange is an example of a voluntary arrangement whose innovation has reduced transaction costs, and an insurance company is an example of a voluntary arrangement designed to reduce risk. At the same time government employment exchanges and the Federal Deposit Insurance Corporation are examples of parallel government innovations. The Union Pacific's development of Sun Valley, Idaho – a development that turned a primitive area into a major resort – is an example of a voluntary cooperative group effectively capturing the externalities associated with the development of a complex of diverse economic activities. The enactment of a zoning ordinance, a government solution, might be aimed at the same goals in an already established community.[4] Trade unions and the American Medical

[1] We recognize that the costs of withdrawal from both voluntary and government arrangements can vary from zero to infinity (i.e. death of the withdrawing individual). Typically, however, the costs of withdrawal from voluntary organizations have been substantially lower than from government ones. Therefore, for purposes of simplicity in exposition, we have followed Buchanan and Tullock in assuming that voluntary organizations abide by unanimity rule and the cost of withdrawal is zero, whereas government arrangements abide by some political decision rule and do not permit withdrawal.

[2] In line with the argument in footnote 1 above, since emigration is always an alternative, withdrawal from governmental decisions are possible, but the larger is the area of sovereignty, the higher are the costs.

[3] These sources of reorganizational profits are spelled out in detail in Section IV of this chapter.

[4] In William Baumol's *Welfare Economics and the Theory of the State*, 2nd ed. (Cambridge Mass., 1968), he explores the rise of government as the method by which externalities may be internalized in a society. But in his second edition he acknowledges the point developed by Buchanan and Tullock in *Calculus of Consent*; it may be equally possible to internalize externalities, in many cases, through the use of voluntary organization as well as via government.