

CHAPTER 1

Institutions, transaction costs, and the rise of merchant empires

DOUGLASS C. NORTH

THE central purpose of this volume is to explore the role that merchant empires played in the evolution of Europe, from its position as a relatively backward part of the world to its preeminence in 1750. My assignment is not to focus on the evolution of superior military technology or on state-building, but to concentrate on the creation of institutions that permitted trading empires to exist and to be viable, that is profitable. I focus, therefore, on the costs of transacting, which are a function of the institutions and technology human beings create to interact with each other in repetitive dealings. But, of necessity, the analysis is concerned with politics, military technology, and so forth because they all interacted to determine the costs of transacting. In what follows, I state the issues from the perspective of an economist (Section I), explore the evolution of the state in its connection with trading empires (Section II), look at the innovation of commercial and financial instruments that lowered the costs of transacting (Section III), and then consider the development of institutions and enforcement procedures that made those instruments possible and effective (Section IV). Finally, I put this institutional evolution in historical perspective (Section V). An appendix elaborates the theoretical and technical issues.

I

The international trade model of the economist is the ideal foil against which to examine this evolution. The model is the ideal foil because

I am indebted to Michael Hauptert for research assistance; to Elisabeth Case for editorial assistance; and to Larry Neal and Jim Tracy for valuable comments on the draft of this chapter presented at the "Rise of Merchant Empires" conference.

it is frictionless: all of the actors have perfect knowledge, there are no institutions or government, and therefore people interact purely on the basis of comparative advantage. The result is that one can clearly see the importance of these assumptions when one contrasts this model with the actuality that we perceive in historical research. The international trade model is built on the principles of comparative advantage, in which differential endowments of the factors of production – land, physical capital, and human capital – would lead each area and region to specialize and then trade with each other. It would be a world of interdependence, of specialization and division of labor, much as Adam Smith described in *The Wealth of Nations*. The model's implications have been explored at length in international trade theory. Factor and product prices tend to be equalized by trade and factor movements among all of the world's players. In this theoretical world, economic growth is not a problem because it is only limited by the tastes and preferences of individuals with respect to the number of children they have and the amount of savings they wish to create, which determine the rate of capital formation. Now, if one wishes to complicate this pure model with the cost of transportation, it changes the model only to the degree that positive transportation costs become a drag upon the relative comparative advantage of players in different parts of the world. But then, of course, as transportation costs fall through history, one would expect that trade and specialization and in consequence growth and prosperity would occur.

Now, if we take a long enough view of economic history, something like this process has occurred. Indeed, we do live today in a world of specialization, where trade and comparative advantage play something like the role described by the model. But for an economic historian the implications of the model do not accord with the evidence over most of the historical period we are concerned with. For most of history, international trade has been a relatively small part of economic activity. Autarky and self-sufficiency, or local trade, have dominated most of history; and only in rare periods, such as during the heyday of the Roman Empire in the Mediterranean world and a few other periods in early history, have we observed extensive amounts of trade occurring. Moreover, declining transportation costs have themselves not been a sufficient reason to induce the growth of international trade, though at times they have helped. And, even if we look at the world today, we see enormous disparity between the rich and the poor nations of the world; in the latter, patterns of economic activity and standards of living still exist that were paralleled by the

Western world centuries and even millenia ago. What is missing from this neo-classical model are transaction costs and the implications of transaction costs for economic growth.

Transaction costs are all the costs of human beings interacting with each other. They involve not only the economic costs of making bargains and enforcing contracts and agreements, but also the political costs of devising a framework of rules and enforcement so that bargains can be extended over time and space and therefore allow us to capture the gains from trade. If transaction costs were zero, then the world of the economist's model that I have been describing would obtain, but, in the presence of positive transaction costs, the world has looked extraordinarily different. Transaction costs include the costs associated with capturing the gains from trade: the costs of measuring and defining the attributes of the goods and services being exchanged and the performance of agents as well as the costs of enforcing agreements with respect to the contracts that are made. They interact intimately with production costs in that it is the two together that make up the total costs of economic activity and define therefore whether trade, specialization, and production and interchange will occur.

Once we introduce transaction costs into the story, we have to ask ourselves an entirely different set of questions than we find in the economist's model of pure international trade theory. How does a transaction-costs framework alter the way in which we explore the historical issues? First, we must continually look at both transaction and production costs to see when it becomes worthwhile for production to take place and for trade and exchange to occur. Throughout most of history, the insecurity and consequent risks have made the costs of transacting so high that trade has been very limited. Second, the role that the state plays then becomes an important part of our examination. Throughout most of history, the state has acted more like the Mafia than an organization that was concerned with economic growth. What made rulers of some states perceive that their interests lay in promoting trade rather than holding trade and economic activity to ransom? During the period we are examining in this volume, the role of the state in enforcing agreements and lowering the costs of transacting among merchants and individual actors is a central part of the story we wish to explore. At some point the state realizes economies of scale in violence or in enforcing agreements (the flip side of the same coin) and therefore begins to assume the role of contract enforcer between parties on an international level. Notarial records suggest that, at the local level in Italian city-states, the state

played a role in enforcing local contracts and supplemented the role of merchants' legal relationships and merchant courts in enforcing property rights and agreements. How the state widens its role, what induces it to widen its role, and the degree to which it helps or frequently hinders the evolution of international trade and the fall of transaction costs, are central to the issues that we must explore.

A central actor in the evolution of international trade was the trading company, a quasi-independent voluntary organization that retained some coercive aspects. It was quasi-voluntary because it was certainly a trading company but frequently was endowed, empowered, and sometimes delegated by the state to have coercive force. Although it at times engaged purely in trade, it was not simply a passive trader, but frequently engaged in plunder and piracy. How did this particular form of voluntary organization evolve; and as it evolved, how did the costs of transacting change? Here the story becomes more complex because, as voluntary organizations evolved, they faced a complex of issues, from external threats of piracy or the inability to get agreements with traders in foreign countries, but also from the internal problems of agency, that is, the problems of constraining and controlling their own agents in international exchange. The agency problem became a central issue as voluntary organizations evolved through time. But they did evolve not only in ways by which to protect themselves against external threats, or achieve some degree of compliance in enforcing agreements in foreign political systems, but also in ways by which they could control the problems of internal organization. They did so, through the development of a succession of organizational and financial innovations, which over time lowered the costs of transacting.

II

What set off the expansion of Western Europe, which led to its ultimate hegemony in the world? A proximate part of the explanation, though certainly not a complete one, is the revolution in military technology that occurred in the late Middle Ages; the cross-bow, the long-bow, the pike, and gunpowder had implications for the organization and capital costs of warfare. The costs of warfare rose. So, accordingly, rose the costs of survival of political units. Because kings were supposed to live on their own, they were faced with devising ways to increase fiscal revenues. This "Crisis of the Tax State," to use the title of a celebrated essay by Joseph Schumpeter published in 1919, led to radical changes in the polities and economies of Western

Europe. A king's revenues from his own estates and other traditional rights he held from feudal obligations were nowhere near enough to be able to pay the armies and mercenaries necessary for survival. Of necessity, the king tapped the wealth and income of his constituents. This led to the creation of representative bodies of wealth-holders, who then traded tax revenues for rights received from the Crown.

At this point our story diverges radically. The Estates General of France and the Cortes of Castile in Spain gradually gave up the powers they had in initial bargains with the king and their role was reduced. In consequence the Crowns of France and Spain (and Portugal too) evolved into centralized monarchies. In order to control the economies and polities, and their overseas empires, they developed large, centralized bureaucracies. The result was an institutional path that was to shape the subsequent history of these polities and economies. The problems of control within this vast bureaucracy present a classic dilemma of agency.

An alternative path was pursued by the Netherlands and England, though certainly not by any intention (in the latter case) of the Crown. Gradually the Estates General in the Netherlands and the Parliament in England achieved a degree of control over fiscal matters and therefore over the polities themselves. The story in England in the revolutionary seventeenth century is a very familiar one; however, the issue of control remained in doubt through much of that century. By 1689 Parliament achieved supremacy, and the divergent path, in England as in the Netherlands, led to a radically different form of organization. From regulated companies to joint-stock companies, increasingly voluntary organizations evolved, more and more independent of government control.

Moreover, throughout the entire Western world, competition was playing a critical role. Polities and economies struggled, not only inside Europe for hegemony but also in the growing empires of the rest of the world, where competition played the critical role (in the decentralized parts of the system) of inducing increased efficiency, as it had in the Netherlands and England. On the other hand, in Spain, Portugal, and France, colonies were run by bureaucratic decree.

The importance of these diverging paths for the analysis in the following sections is clear. In the case of the Netherlands and England, decentralized control produced an adaptively efficient set of institutions that adjusted to changing needs; thus, competition gradually forced the development of more efficient institutions and instruments that promoted commerce and trade, and, in consequence, lowered

transaction costs in these economies. The bureaucracies of Spain, Portugal, and France, on the other hand, lagged behind, stifling initiatives that would have induced increased productivity, and Spain and Portugal pursued a downward path that would continue for centuries.

III

Innovations that lowered transaction costs consisted of organizational innovations, instruments, and specific techniques and enforcement characteristics that lowered the costs of engaging in exchange over long distances. These innovations occurred at three cost margins: 1) those that increased the mobility of capital; 2) those that lowered information costs; and 3) those that spread risk. Obviously these are overlapping categories; however, they provide a useful way to distinguish cost-reducing features of transacting. All of these 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. Because there is already an immense literature on the specifics of these innovations and instruments, my task, as I see it, is to attempt to draw out the larger implications of the specific developments for economic growth. In this section, let me schematically order the innovations. In the next section, I will look at their implications.

Let me begin with innovations that affected the mobility of capital. The first of these were the techniques and methods evolved to evade usury laws. The variety of ingenious ways by which interest was disguised in loan contracts ranged from "penalties for late payment" to exchange-rate manipulation,¹ to the early mortgage, but they did increase the costs of contracting. The costliness of usury laws was not only that the need for disguising interest made the writing of contracts complex and cumbersome, but also that enforceability of such contracts had become more problematic. As usury laws gradually broke down and rates of interest were permitted, the costs of writing contracts and the costs of enforcing them declined.

A second innovation that improved the mobility of capital, and the one that has received the most attention, 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

¹ Robert S. Lopez and Irving W. Raymond, *Medieval Trade in the Mediterranean World* (New York; 1955), 163.

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 such events could occur, first in fairs, such as the Champagne fairs, then through banks, and finally through financial houses that could specialize in discounting. These developments were a function not only of specific institutions but also of the scale of economic activity. 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 as well as auditing methods and their use as evidence in the collection of debts and in the enforcement of contracts was an important part of this process.²

Still a third 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 authority 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.

When we turn to information costs, the major developments were the printing of prices of various commodities as well as 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.

The final innovation consisted of the transformation of uncertainty into risk. By uncertainty, I mean here a condition wherein one cannot ascertain the probability of an event and therefore cannot arrive at a way of insuring against such an occurrence. Risk on the other hand implied the ability to make an actuarial determination of the likelihood

² B. S. Yamey, "Scientific Bookkeeping and the Rise of Capitalism," *Economic History Review* ser. 2 1 (1949): 99–113; Ross Watt and Jacob Zimmermann, "Agency Problems, Auditing, and the Theory of the Firm: Some Evidence," *Journal of Law and Economics* 26 (1983): 613–34.

of an event and hence insure against such an outcome. 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. Indeed, when we look at the medieval and early modern world, we find precisely the same results. That is, marine insurance evolved from sporadic individual contracts covering partial payments for losses to contracts issued by specialized firms.

By the fifteenth century marine insurance was established on a secure basis. The wording of the policies had already become stereotyped and changed very little during the next three or four hundred years . . . In the sixteenth century it was already current practice to use printed forms provided with a few blank spaces for the name of the ship, the name of the master, the amount of the insurance, the premium, and a few other items that were apt to change from one contract to another.³

Marine insurance was one example of the development of actuarial, ascertainable risk; another was business organization that spread risk through either portfolio diversification or through institutions that permitted a large number of investors to engage in risky activities. The *commenda* itself, from its Jewish, Byzantine, and Muslim origins⁴ 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 (though, as discussed below, the developments created new problems of agency for the principals involved).

IV

The specific innovations and particular institutional instruments briefly described in the previous section evolved as a result of the interplay of two fundamental economic forces. One was the economies of scale associated with a growing volume of trade; the other was the development of improved enforcement mechanisms that made possible the enforcement of contracts at lower costs. Surely the causation ran both ways. That is, the increasing volume of long-distance trade raised the rate of return to merchants of devising effective mechanisms for enforcing contracts. In turn, the development

³ Florence De Roover, "Early Examples of Marine Insurance," *Journal of Economic History* 5 (1945): 198.

⁴ Abraham Udovitch, "At the Origins of Western Commenda, Islam, Israel, Byzantium?" *Spectrum* 37 (1962): 198–207.

of such mechanisms lowered the costs of contracting and made trade more profitable, thereby increasing its volume.

When we look at the development of enforcement mechanisms, we see first that the process was a long one. Although a number of courts handled commercial disputes, the development of enforcement mechanisms by merchants themselves is of particular interest. Enforceability appears to have had its beginnings in the development of 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 evolved and was conveyed through long-distance trade codes of conduct, so that Pisan laws passed into the sea codes of Marseilles. Oleron and Lübeck gave laws to the north of Europe, Barcelona to the south of Europe, and from Italy came the legal principle of insurance and bills of exchange.⁵

The development of more sophisticated accounting methods and of 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 those of 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 two types of law did not accommodate each other very well to begin with. This was particularly true in cases of moral hazard and asymmetric 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. Eventually, this was changed. The turning point could be designated as 1756, the year Lord Mansfield became chief justice of the English court of King's Bench. He gave form to the existing customs. He established general principles that were to be used to rule future cases. He was not too fond of English common law and

⁵ William Mitchell, *An Essay on the Early History of the Law Merchant* (New York: reprint edition, Burt Franklin, 1969), 156.

as a result he derived many of these principles from the writings of foreign jurists. "For instance, in his judgement in *Luke v. Lyde*, which raised a question of the freight due for goods lost at sea, he cited the Roman Pandects, the *Consolato del Mare*, laws of Wisby and Oleron, two English and two foreign mercantile writers, and the French *Ordonnances*, and deduced from them a principle which has since been part of the Law of England."⁶

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 that aided in decreasing transaction 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 that merchants 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. But in commercial disputes the "good-faith" principle was used earlier and on a much wider scope (the basis of Roman contract law by A.D. 200). It evolved first out of the Fair Bonds, which validated

⁶ Thomas E. Scrutton, *The Elements of Mercantile Law* (n.p.: Wm. Cloves & Sons, Ltd., 1891), 15.

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.

Many rules of merchant law developed because common law interfered with trade. For example, the common law’s failure to protect bona fide purchasers forced examining the title of goods all the way back to the original owner. This presented an obvious problem for merchants: the cost and time required to carry out such a search were prohibitive.

This problem caused the first exception to common law that the law merchant made. The evolution of the situation from the thirteenth to the sixteenth century can be measured by the manner in which purchasers of goods with fraudulent titles were treated. In the thirteenth century, the purchaser of such goods would be forced to return them upon the discovery of a discrepancy anywhere along the chain of ownership of the good. By the time Sir Edward Coke was appointed chief justice in 1606, the final (good-faith) purchaser of a good was recognized (in certain but not all courts) as having the only viable title to the good, making any legal purchase he made legal all the way back down the chain of ownership.

Another, similar example of the common law hindering the merchants was the common law’s ruling that a “chose in action” was not transferable. This is the right to a thing as opposed to the thing itself; the separation of the property right from the property itself. This was inconvenient if the bearer of a bill of exchange had to check every endorser in the chain’s credibility and furthermore was not able to sue in his own name, but would have to do so in the name of the man mentioned as payee on the bill. The law merchant established certain “choses in action” called negotiable instruments. The holder of a negotiable instrument could then sue in his own name. Also, he

was not affected by the previous lack of title. The laws of negotiable instruments were almost entirely built on custom.

A major player in this whole process 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 because, 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.⁷ In addition, 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 a key part of this process.⁸

It was of course 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.⁹

Equally well known is the evolution of this process in England. The political conflict of the seventeenth century that culminated in the triumph of Parliament in 1688 and the subsequent flowering of the capital market in the next twenty-five years in England,¹⁰ the expansion of long-distance trade, the improved enforcements of contracting, and the reductions of uncertainty that came with the development

⁷ Douglass North and Barry Weingast, "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in the Seventeenth-century England," *Journal of Economic History*, 49 (1989): 803–832.

⁸ James D. Tracy, *A Financial Revolution in the Hapsburg Netherlands: Renten and Renteniers in the County of Holland, 1515–1565* (Berkeley, 1985).

⁹ Violet Barbour, *Capitalism in Amsterdam in the Seventeenth Century* (Baltimore, 1950).

¹⁰ Peter Dickson, *The Financial Revolution in England: A Study in the Development of Public Credit, 1688–1756* (New York; 1967).

of commerce and the joint-stock company were equally a part of the process in Britain.

Let me conclude this section by briefly assessing both the improvements in the costs of transacting that occurred in this period and the ongoing dilemmas. Perhaps the most striking feature that one can draw from transaction-cost stories of companies, such as the English East India Company,¹¹ the Hudson Bay Company,¹² or the Royal African Company,¹³ all of which have been extensively documented in their historical evolution, was the degree of sophistication they brought to solving the problems they confronted in a world where information costs (at least by modern standards) were enormously high, where enforcement of contracts was, even at the end of the period, quite uncertain, and where problems of agency increased as the volume and size of the companies grew to previously unknown levels. These problems, particularly the latter, were major dilemmas.

Yet, what stands out in the history of these companies is how sophisticated they were in solving these problems. David Galenson's study of the slave trade and of the evolution of the Royal African Company provides us with a good case study of this process. He demonstrates conclusively that sophisticated techniques were used to make markets work efficiently, at least efficiently in comparison to previous notions of trading patterns. He also demonstrates that the Royal African Company, despite being granted a monopoly, was itself a victim of the dilemmas of the world of such uncertainty.

There were three major reasons for the failure of the Royal African Company. First was the lack of rapid communications technology, which hindered the ability of the firm to respond to changing market conditions and to monitor agents. Second, its charter imposed certain costs on the company in exchange for certain other benefits. The Royal African Company bore those costs, such as manning forts in Africa, but never did reap the promised benefits. These benefits included a grantable monopoly in the slave trade between England and the American colonies in the Caribbean. Third, England refused to allow the West Indian colonies to issue their own currency and the result was a chronic shortage that left the Royal African Company the major creditor of the colonies. All of the problems faced by the Royal African

¹¹ Gary Anderson, Robert McCormick, and Robert Tollison, "Economic Organization of the English East India Company," *Journal of Economic Behavior and Organization* 4 (1983): 221–38.

¹² Ann Harper-Fender and Elizabeth Mancke, "Hudson's Bay Company: Precursor to the Modern Corporation," unpublished manuscript.

¹³ David W. Galenson, *Traders, Planters, and Slaves: Market Behavior in Early English America* (Cambridge, 1986).

Company might still have been overcome had it not been for the competition the firm faced. Even though it was granted the Crown monopoly, several competing firms existed. Some of these were foreign competitors, some British. Because the Royal African Company faced constraints (discussed below), while their competitors were free to operate without them, they were eventually driven out of business by more efficiently operating firms. It is an indication of the sophistication of the competing firms that they were able to overcome the problems of agency, long-distance communication, and so forth that all firms faced, and still fare better than the company protected by the Crown. The three agency problems the Royal African Company had to solve were those associated with ship captains, agents manning the forts, and sales agents in the West Indies. Monitoring the ship captains turned out to be extremely difficult and indeed led to the captain's taking advantage of the company at every turn. Similarly, agents manning the forts in Africa were a continual problem, a deadly combination of incompetence and dishonesty; again, the problems of monitoring them turned out to be an insuperable dilemma. A similar dilemma existed with agents who sold slaves in the West Indies. Perhaps the Royal African Company is the ideal illustration with which to conclude this section or because it indicates on the one hand how far company organization had come in its ability to engage in long-distance trade in the face of the dilemmas of high information costs and uncertain enforcement and on the other hand the road that lay ahead before further extension of such trade and commerce could occur.

V

What distinguished Western Europe from other places in the world, where persistent economic growth failed to occur, was that there gradually evolved a set of adaptively efficient institutions that persistently tended to lower the costs of transacting, producing, and transporting in a way that produced a continuous evolution of productivity increases in these societies (see appendix). We know all too little about this process, but clearly merchant empires were a step along the way. They were a step from autarky, localized trade, to larger trade and specialization, which at least for some economies, notably the Netherlands and England, were steps along the route to a persistent evolution of more efficient forms of economic organization. Perhaps the best way to look at this process is to reverse the perspective of Section I. There we used the neo-classical model, where

all rights are perfectly specified and delineated and where parties engage in trade and exchange without facing problems of measurement or enforcement. The reverse is a world of complete anarchy and autarky, where no rights are recognized by the parties, and no enforcement mechanisms exist. Such a world never, in fact, really existed, because, even in tribal groups and primitive societies, some rights were at least implicitly recognized in small-scale exchange. Nevertheless, we do observe in early modern Europe a gradual growth in specialization and division of labor made possible by a variety of institutions and instruments that reduced the uncertainties associated with contracting and enforcing trade over long distance and with increasingly diverse and "unknown" trading partners.

As indicated above, two major influences were a part of this process. First were the economies of scale associated with the volume of exchange; and second was the development of improved methods of enforcing contracts. Although such economies of scale can be looked upon, in a sense, as an automatic consequence of this volume of exchange, political action by the state to encourage adaptively efficient instruments was anything but automatic. Indeed, the divergent paths of the Netherlands and England on the one hand, versus Spain and Portugal on the other, provide us with an important clue to resolving this issue. Clearly, the incremental change of institutions and the consequent path-dependent evolution, which take us down one road or another, were a major part of this process. Path-dependency suggests that we can learn as much from the dead-end path pursued by Spain and Portugal, with respect to institutional evolution, as we can from the successful paths to evolving more efficient institutions pursued by the Netherlands and England. We are a long way from completely understanding the interplay among institutions, transaction costs, and economic growth, but exploring changing transaction costs and their implications and consequences for institutional evolution is a major step ahead in improving our understanding.

APPENDIX: TRANSACTION COSTS AND PRODUCTIVITY INCREASES IN EARLY MODERN WESTERN EUROPE

The growth of Western Europe – indeed the rise of the Western world – was a consequence of the growing productivity of these economies and their relative rise was a consequence of the contrast in productivity growth between Western European countries and the rest of the world. The story I have recounted in the foregoing essay implies

that declining costs of transacting brought about by the innovations of institutions played a key role in the process of growth. In the essay, I was primarily interested in telling a "plausible story." In this appendix, I wish to lay out the logic of my arguments, provide evidence consistent with them, and, most important, invite researchers to develop evidence systematically that will critically evaluate these hypotheses.

Productivity increase results from getting more constant quality output from given amounts of inputs of the factors of production. We customarily ascertain productivity increases by contrasting the inputs of land, labor, and capital required for a given amount of output in two different settings. Economists have traditionally conceived of the inputs of raw materials, labor, machinery, and so forth as being transformed into an intermediate or final good or service. In fact, concealed in this simple formulation were transaction as well as transformation costs, and accordingly we should have enlarged the argument to include the real resource inputs that went into both transformation (as a consequence of technology) and transaction (as a consequence of institutions) in order to get the total input costs associated with any given output. That is, the costs of production (inputs) are a function of the resources used in both transforming and transacting. Moreover, it is important to note that, just as a new technology (for example, the computer) might as easily have reduced the cost of transacting as the cost of transforming, so might an institutional change have reduced the cost of transforming (for example, unitizing an oil field through the allocation of property rights) as well as the cost of transacting. Clearly, the interdependence between transformation and transaction costs and between technological change and institutional change is complicated. And because it is complicated – and new to our understanding – we need to rethink a great deal of our understanding of economic history.

Interdependence makes the task of unambiguously measuring the role of transaction costs in productivity change much more difficult. Moreover, the difficulty of measuring variation in the quality of inputs and outputs and the dilemmas (also common to standard national income accounting) posed by costs of transacting that do not go through the market or are otherwise measured (queuing, waiting, bribery, quality deterioration, and so forth) complicate the task even further.

These problems are difficult but not insurmountable, and a major task of the economic historian is to attempt to get as good a measure as possible. Let me make some suggestions and then discuss some

of the evidence we currently have to explore the issues raised in this essay. The ideal data would measure the cost-reducing consequences of a specific-institutional or technological change. Even more ideal for an economic historian is to know the sequence of changes and their cost effects because usually the consequence of a given institutional (or technological) change is to induce a technological (or institutional) change. The problems of imputing causation are always present. For example, to choose an illustration relevant to this essay, the triumph of Parliament in 1688 was followed in the next 25 years by what has been described as a financial revolution.¹⁴ Not only did interest rates on government securities decline but so did interest rates of nongovernmental securities. The new institutions that played a part in this process were the Bank of England (1694), the creation of regular sources of tax revenue to finance government securities, and the development and expansion of other financial intermediaries. But the assumptions in this argument are 1) that the triumph of Parliament did induce the creation of these institutions; and 2) that these institutions were responsible for lowering interest rates by reducing the costs of transacting. Obviously we cannot prove the sequence of hypotheses to be true; however, we can at least test our confidence in them by examining alternative explanations, for example, checking that the interest-rate decline was not a consequence of changes in the price level (comparing real versus nominal interest rates) or of some other nontransaction cost change. Similar problems of measurement and causation are involved in the secular decline in marine insurance rates during this period. Perhaps the best evidence is found in the decline in ocean freight rates after 1600 that appears to have been a consequence of both institutional and technological change.¹⁵

An even more difficult but still critical issue, particularly with respect to the focus of this volume, is to explain the initiation of production and exchange where none existed before because a basic assertion of the transaction-costs approach is that it has been the

¹⁴ Dickson, *Financial Revolution*.

¹⁵ In "Sources of Productivity Change in Ocean Shipping: 1600–1850," *Journal of Political Economy* 76 (1968): 953–70, I maintained that the productivity increase over the period was solely due to institutional and not technological change, but more recent research by Knick Harley, "Ocean Freight and Productivity, 1740–1913: The Primacy of Mechanical Invention Reaffirmed," *Journal of Economic History* 68 (1988): 851–76, has demonstrated I was wrong for the later part of the study and that technological change was more important. However, the evidence for the seventeenth century does, I believe, support my argument. See also the chapter by Russell Menard in this volume.

insecurities and uncertainties of production and exchange that are the major stumbling blocks to economic growth, and it is their reduction that has led to development. But how can we know that an institutional change has induced a decline in transaction/transformation costs and subsequently led to production or exchange where none existed before? After all, any downward shift in the supply curve or demand curve for the good or service could explain its beginnings. Delving into the immense literature that exists in economic history on the question of whether a given expansion was demand- or supply-induced would take me too far afield. However, much of that literature, though inconclusive, does suggest some partial answers. On the supply side we can explore whether alternative technological explanations are possible, as was done in the case of ocean shipping. We can try to sort out the sources of productivity increases, whether economies of scale, technological change, institutional change. The most plausible explanation may be a sequential interdependent process embodying over time all three sources. (Again, however, it is true that our task of definitive measurement will be made vastly more complicated by the latter.)

Before turning to some of the specific issues in the foregoing essay, let me add one important point that cliometricians have tended to obliterate or at least ignore with their voluminous data sets. Qualitative evidence is immensely valuable. Used carefully to reflect the same concerns that plague us even with numbers, such evidence is frequently going to have to be the "only game in town." Do not misunderstand me. I think that there is vastly more statistical data out there, which, by asking the right questions, can be unearthed; and I strongly believe that even for the early period we can develop much improved quantitative measures. But we should also – indeed we must – judiciously employ qualitative evidence. For example, regardless of the deficiencies of quantitative data, the accounts we possess of the growth of the Dutch capital market, the consequent increased mobility, and the institutional structure that underpinned it are important. Even though we may suspect some of the real interest rate data, the widespread agreement that Dutch interest rates were lower than elsewhere appears to be solidly based. Moreover, the qualitative contemporary evidence overwhelmingly links them to low costs of transacting as a consequence of political stability and the underlying capital market institutions.

What kind of evidence do we possess to assess the three sources of transaction cost declines discussed in these pages? Consider first

the mobility of capital. We do have interest rate series¹⁶ that, whatever their substantial drawbacks, are consistent. We have a large volume of qualitative evidence on the evolution of the bill of exchange and the increased mobility that were made possible by improved negotiability and discounting. We have evidence on the development of a regular government capital market.¹⁷ By the beginning of the eighteenth century we have quantitative evidence of the integration of London and Amsterdam financial markets.¹⁸ The decline in information costs is recorded with the production of merchant manuals, the beginnings of "prices current" and the growth of a financial press.¹⁹ The conversion of uncertainty into risk obviously is reflected in the growth of marine insurance, the development of a number of financial intermediaries, such as discount houses, and of course the development and spread of the joint-stock company.

I do not mean to imply that I have summarized all the evidence available. I do mean to conclude this appendix by inviting economic historians to attempt to systematically develop quantitative information and to pull together and synthesize qualitative information that will provide us with a surer sense of the roles transaction costs and institutions played in the evolution of early modern Europe.

¹⁶ Sidney Homer, *A History of Interest Rates* (New Brunswick, 1963); Carlo Cipolla, *Before the Industrial Revolution: European Society and Economy, 1000–1700* (New York, 1980).

¹⁷ Dickson, *Financial Revolution*; Tracy, *Hapsburg Netherlands*.

¹⁸ Larry Neal, "Integration of International Capital Markets: Quantitative Evidence from Eighteenth to Twentieth Centuries," *Journal of Economic History* (1985): 219–26; Robert Eagly and Kerry Smith, "Domestic and International Integration of the London Money Market, 1731–1789," *Journal of Economic History* 46 (1976): 198–212.

¹⁹ Neal, "The Rise of a Financial Press, London and Amsterdam, 1681–1796," unpublished manuscript, 1985.