Knowledge and Competitive Advantage

The Coevolution of Firms, Technology, and National Institutions

Johann Peter Murmann
Northwestern University

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

http://www.cambridge.org

http://johann-peter.murmann.name

First published 2003

Printed in the United States of America

Typeface Garamond 10/12 pt. *System* LaTeX 2ε [TB]

A catalog record for this book is available from the British Library.

Library of Congress Cataloging in Publication Data

Murmann, Johann Peter, 1967–

Knowledge and competitive advantage : the coevolution of firms, technology, and national institutions / Johann Peter Murmann.

p. cm. – (Cambridge studies in the emergence of global enterprise)

Includes bibliographical references.

ISBN 0-521-81329-8 (hc.)

1. Dye industry – Germany – History. 2. Dye industry – Technological innovations – Germany – History. 3. Dye industry – Great Britain – History. 4. Dye industry – United States – History. 5. Competition, International – Case studies. 6. Comparative advantage (International trade) – Case studies. I. Title. II. Series.

HD9660.D843G35 2003
338.4′76672 – dc21 2003043048

ISBN 0 521 81329 8 hardback

Contents

CHAPTER

1 Introduction

[M]ost of the fundamental errors currently committed in economic analysis are due to a lack of historical experience more often than to any other shortcoming of the economist's equipment[.]

—JOSEPH A. SCHUMPETER (1954) in *A History of Economic Analysis* (Schumpeter and Schumpeter, 1986 edition, p. 13)

[The master economist] must reach a high standard in several different directions and must combine talents not often found together. He must be mathematician, historian, statesman, philosopher – in some degree.

—JOHN MAYNARD KEYNES (1933) in *Essays in Biography* (1971 edition, p. 173)

I concluded in my most recent research that detailed longitudinal case studies, covering long periods of time, were necessary to study [competitive success] . . . This style of research nudges strategy research, and indeed industrial economics, into the world of the historian.

—MICHAEL PORTER in *Strategic Management Journal* (1991, p. 116)

The Mecca of the economist lies in economic biology . . . But biological concepts are more complex than those of mechanics.

—ALFRED MARSHALL (1890) in *Principles of Economics* (1997 edition, p. xx)

[E]xplaining many things about the coevolution of populations in a community requires narrative history as a complement to statistical analysis.

—PAUL DIMAGGIO in *Evolutionary Dynamics of Organizations* (1994, p. 446)

There are few things – perhaps only one – that can arouse the passions of human beings as much as wealth. Humans need material objects to survive. But as social creatures, human beings desire wealth often not in an absolute sense of possessing more than before but in the relative sense of possessing more than one's neighbor. Thorstein Veblen (1899, p. 290) called this passion "the emulative predatory impulse," which he regarded as an evolved cultural modification of the basic instinct of workmanship that gave human beings a predilection for worthwhile achievement. It is not a coincidence that Adam Smith's classic text *The Wealth of Nations* (1776) was preceded by his *Theory of Moral Sentiments* (1761), in which he inquired into the passions that were necessary for creating a society capable of generating wealth. A close reading of Smith reveals that understanding economic inequality – a topic that became a key issue a century later in the development of sociology – was for him quite essential in identifying wealth-generating processes (see, for example, *The Wealth of Nations*, Book V, Chapter 1). Despite this longstanding human passion for wealth and the almost equally long-standing fear of destabilizing inequalities in modern societies (Hirschman, 1977), economics and sociology have not yet provided a complete understanding of how nations generate wealth and how they can distribute it relatively evenly. Even when we narrow the question considerably and inquire why nations differ dramatically in the performance of a particular branch of industry, existing theories do not provide us with an adequate explanation (Porter, 1990; Mowery and Nelson, 1999). Consider the following intriguing puzzle.

The Puzzle

London, 1856: William Henry Perkin serendipitously invents the first synthetic dye while trying to synthesize quinine, a medicine for malaria. Against the advice of his professor, August Wilhelm Hofmann, the nineteen-year-old Perkin leaves the Royal College of Chemistry and quickly commercializes his aniline purple dye, thereby launching the synthetic dye industry. From this time on, the industry continued to dazzle the eye with ever-new and appealing dye colors. Perkin, along with entrepreneurs from Britain and France, dominated the synthetic dye industry for the next eight years. During this period, British and French firms introduced most other innovative synthetic dyes onto the market and held the largest global market share.

Contrary to contemporary predictions, however, these firms were not able to sustain their leadership position in the new industry. German firms such as Bayer, BASF, and Hoechst[1] (some of the largest firms in the global chemical industry at the turn of twenty-first century) started to gain in market share. By 1870, Germany had about 50 percent of the global synthetic dye market. Britain fell to second place. By 1900, Germany's worldwide share climbed as high as 85 percent,[2] where it stayed with relatively minor fluctuations until World War I. From the 1860s on, American firms also tried to be successful participants in the U.S. market but could not compete with

1 Hoechst merged on December 15, 1999, with the French firm Rhône-Poulenc. The combined company was renamed Aventis.

2 This is Reader's (1970, p. 258) estimate. Thissen (1922) provides a lower estimate of 75 percent for Germany's global market share. If one includes German plants in foreign countries, the share may have been as much as 90 percent. Even though there is no agreement on the exact figure, everyone concurs that German firms collectively dominated the world market on the eve of World War I by a wide margin.

German and Swiss firms before World War I; they remained relatively small players or went out of business.

Any explanation of the shift in industrial leadership from Britain and France to Germany quickly becomes mired in an intriguing puzzle where the obvious suspects have surprising alibis. Possessing cheaper raw materials or a larger home market cannot account for why German firms left British and U.S. firms in the dust, because both latter countries had more raw materials and a larger home market. Why, then, did Britain lose its leadership position? Why did the American dye industry remain so small before 1914? This book takes a new tack to resolve the puzzle by engaging in a detailed historical analysis of what caused this transition in industrial leadership.

The purpose of trying to solve this particular puzzle of why industrial leadership shifted during the first fifty-seven years of the synthetic dye industry is to make a contribution to two important intellectual agendas pursued by scholars in a number of different fields. Adam Smith (1776) and David Ricardo (1817) and, more recently, Michael Porter (1990) and David Mowery and Richard Nelson (1999) are prominent examples of a wide array of social scientists who have tried to identify the factors that lead nations and firms to prosper. For economists and management researchers, the question of how economic success is generated remains a key intellectual challenge (Kogut and Zander, 1996; Landes, 1998; O'Sullivan, 2000). In tracing the development of one industry within the context of three countries, I hope to make a significant contribution toward formulating a much-needed dynamic theory of industrial leadership. At the heart of the theory lies the concept of coevolution, which has been employed with much success by researchers of biological (Kauffman, 1993, Chapter 6; Thompson, 1994) and cultural (Lumsden and Wilson, 1981; Durham, 1991) change. Recently, ideas of coevolution have been introduced in the discourse on industrial leadership (Nelson, 1995a), technological change and economic growth (Mokyr, 2002), and development of firms (Lewin, Long, and Carroll, 1999). What we need now is a theory that does more than explain industrial leadership at a particular time. I believe a coevolutionary theory that models firms as interacting with their social environment takes a significant step toward explaining how industrial leadership is gained and lost and how small initial differences in performance can translate into large differences over time.

In placing national institutions and technology into the center of my analytical framework, I continue an unduly neglected tradition that flourished around the turn of the twentieth century. Prominent social scientists such as Thorstein Veblen (1915) saw national institutions and their effects on technological development as a key to understanding why Germany, for instance, achieved higher rates of economic growth than Britain in those years. More recently, scholars from various disciplines have relied on institutional accounts to explain both Japan's rise to economic leadership after World War II (Fruin, 1992; Gerlach, 1992) and its recent economic troubles (Thomas, 2001). Nobel laureate Douglass North (1990) has argued in a recent book for an important role of institutions in shaping economic performance. However, institutional arguments have long been given scant attention in economic analysis (Hodgson, 1998; Hall and Soskice, 2001).[3] The goal is to help move these arguments

[3] To be sure, in some areas of sociology, institutional arguments have been more prominent in the last couple of decades (Hollingsworth and Boyer, 1997). Scholars who compare the development of

once again onto center stage and to focus our attention on a critical missing piece in institutional analysis, namely, how institutions are created in the first place.

The book also tries to make a contribution to a second important line of work that concerns itself with the rise and development of the large managerial firm as a new economic institution. The business historian Alfred Chandler (1962, 1977, 1990), who pioneered the study of this corporate entitity that appeared on the scene in the second half of the nineteenth century, identified this new organizational form as a key source of economic growth over the past century. On the Chandlerian model, large firms – those run by professional managers rather than owners – came to dominate industrial activity in modern industrialized economies because they could operate more efficiently by exploiting the scale and scope economies made possible by cheap transportation (railroads) and communication (telegraph). Sociologists Neil Fligstein (1990) and William Roy (1997) have argued that Chandler's analysis is incomplete because it leaves out the political context in which large managerial firms originate. I attempt to integrate the writings on the rise of the large managerial firm in business history and sociology by focusing on how collective action on the part of firms molded the social and institutional environment in which firms operate. I marshal considerable evidence to show how the rise of the large managerial firm required the construction of an institutional regime that would favor such firms over other forms of organization. As we shall see, German firms in the synthetic dye industry were much more successful in molding their institutional environment than were their British and American counterparts. In Chandler's analysis of the rise of large firms in Germany, Bayer figures prominently (Chandler, 1990, pp. 474–81) as an example of how a sophisticated managerial hierarchy was created that could organize more efficient production than smaller firms could. I will argue that Bayer could only realize its economic advantage precisely because it became a key player in lobbying efforts to create a favorable institutional environment. Bayer's leaders sought prominent roles in the chemical industry trade association and participated in collective action to improve the German education system in chemistry as well as change German patent laws to give large firms an advantage over foreign competitors and smaller domestic rivals.

One of the key propositions of this book is that the creation of German dominance in the synthetic dye industry before World War I cannot be understood without coming to terms with successful and unsuccessful patent law, science funding, and tariff lobbying efforts in the three countries. My analysis of the synthetic dye industry shows that we need to rediscover scholarship that recognized the importance of lobbying in industrial development, such as Galambos's (1966) *Competition & Cooperation: The Emergence of a National Trade Association* and Hirsch's (1975) *Organizational Effectiveness and the Institutional Environment*. Comparing the fates of firms in the three countries, this study shows that firms depended on their social environment for resources to prevail against foreign competitors. In the case of Germany, firms were able

professions across different countries (Gispen, 1989; Cocks and Jarausch, 1990; Jarausch, 1990; Lundgreen, 1990; Guillén, 1991) cannot but notice the importance of institutional differences. Similarly, scholars of comparative management such as Mauro Guillén (1994) frequently appeal in their explanations to institutional arguments. Comparative political scientists have also frequently resorted to institutional arguments in explaining differences in behavior across countries. See Thelen (1999) for a recent review of this literature.

to obtain more resources from their social environment than were their British and American counterparts. My analysis of why German firms overtook their foreign rivals and then cemented their leadership is consistent with Pfeffer and Salancik's (1978) theory of resource dependence, which highlights the political nature of creating successful organizations. One of the critical resources that firms in the synthetic dye industry needed to obtain was access to organic chemical knowledge and dye innovations. Because this knowledge, as well as dye innovations, was heavily concentrated at universities in the early period of the synthetic dye industry, firms needed to develop ties to university professors and their students. Examining the dependencies of a dye firm through Burt's (1992) more formal, network version of resource dependence theory makes apparent that firms were competing for access to the leading organic chemists in the world.[4] Those firms that were able to maintain ties to the best chemical talent of the day outperformed rivals that were not as well connected. After working in a professor's university laboratory, chemists often moved from academia to industry, from one firm to the next, and sometimes back to a university position. This created an informal network of ties that connected players in industry and academia. Mapping the network on a worldwide scale for the period before World War I reveals not only that this informal network was overwhelmingly constituted of Germans but also that the central positions were occupied by players from Germany. Explaining the shift in industrial leadership in the synthetic dye industry is intimately bound up with being able to account for the strong and weak ties in what I will call the academic–industrial knowledge network.

The informal network assumed a second function beyond simply transferring knowledge about chemical synthetic dyes. It served as a mechanism for organizing collective action. A key reason why German firms engaged in significantly more successful collective action to shape domestic patent laws and university policies is that they could rely on a much stronger network of actors spanning industry, academe, and government.[5] To be effective in orchestrating lobbying efforts, the informal academic–industrial network was enriched by ties to high-level government officials. Where this industrial–academic–government network was large and close-knit (Germany), collective action on behalf of the dye industry tended to succeed; where the network was small and distant (Britain and the United States), collective action was more likely to fail.

Is This Book for You?

Having made a careful investigation of the synthetic dye industry, I attempt to articulate a theory of coevolution of firms, technology, and national institutions so that other scholars can join the effort to formulate a rigorous theory of coevolution and to collect empirical support for it. My goal is to stimulate debate and new research rather than to provide the definitive word on the subject. I have not studied in detail all industries that ever existed from the industrial revolution until current times and hence it is an open question whether the theory I develop in this book is universal in scope. But my reading of other industry studies (e.g., Mowery and Nelson, 1999)

[4] My empirical analysis does not present a full-fledged test of Pfeffer and Salancik's (1978) or Burt's (1992) theories. Rather it draws on these theories to illuminate the historical dynamics in the synthetic dye industry.

[5] For an overview of the literature on organizational networks and alliances, see Gulati (1998).

gives me confidence that the theory provides a powerful analytical lens for industrial development in general.

Two communities – analysts of industrial leadership and scholars working on the rise of large managerial firms – are likely to find the study of the synthetic dye industry rewarding. If one surveys the literature in business strategy and organization theory, it becomes clear that both fields would benefit from devoting more attention to business history (Chandler, 1990; Galambos and Sewell, 1995; McCraw, 1997b; Jones, 2000; Lazonick, 2002).[6] Management scholars in business schools have become very competent over the past 25 years in conducting rigorous cross-sectional studies that relate differences in the structure of firms at a particular time to differences in their outcomes. But apart from organizational ecologists (Hannan and Freeman, 1989; Carroll and Hannan, 1995b; Rao and Singh, 1999) and several other scholars (e.g., Tushman and Anderson, 1986), not enough research on business strategy and organization theory is longitudinal in nature (Lewin and Volberda, 1999). Business historians, in contrast, take it as a given that understanding the actions of a particular firm requires the analyst to examine the development of the firm over time. Through a detailed study of the evolution of the dye industry, I hope to persuade management scholars that the field would be able to build more robust models of organizational development if it could draw on more historical studies.

This book also brings the sampling methods of organizational ecologists to business history. In doing so, I examine the entire population of dye firms that existed from 1857 to 1914 – large firms and small firms, short-lived and long-lived, failures and successes. Conclusions derived from the study of an entire industry become much more convincing if they are supported by detailed evidence from specific firms. Hence I also conduct matched comparisons of three pairs of firms – a winner and a loser from Britain, from Germany, and from the United States. These case analyses will demonstrate the benefit of examining causal processes with a higher resolution lens at the level of specific firms.

The resource-based theory of the firm (Wernerfeld, 1984; Dierickx and Cool, 1989; Barney, 1991; Teece, Pisano, and Shuen, 1997; Kraatz and Zajac, 2001) has received wide attention in the field of business strategy during the past decade and, given the current trend of viewing knowledge management as the key to success in a globalizing economy, the theory is likely to remain at the forefront of the research agenda. (See Kogut and Zander [1992] and Zander and Kogut [1995] for pioneering statements about the importance of knowledge in firm success, and Loasby's [1998; 2001] contributions on the cognitive foundations of organizational capabilities.) This literature identifies firm resources and capabilities that are hard to trade and hard to replicate as the key source of competitive advantage. Because empirical support for the theory has been collected generally within the context of one country, the resource-based theory of the firm has never dealt with the question of how the larger social environment rather than a particular firm itself may be an important source of

[6] Although many strategy scholars cite Chandler's (1962) book *Strategy and Structure: Concepts in the History of American Industrial Enterprise* as the pioneering study in the field of strategy, the share of historical analyses is much smaller than one would expect for a field that refers to *Strategy and Structure* as a foundational study.

competitive advantage. I show that bringing the resource-based theory together with ideas from institutional theory can refine the resource-based theory of the firm.

The new institutionalists in sociology and organization theory (Meyer and Rowan, 1977; Powell and DiMaggio, 1991; Dobbin, 1994; Scott, 1995) have argued against taking an atomistic view of organizations in which causal processes originate and end within the boundaries of the firm. Understanding why organizations fail and succeed, why they take on a particular form and not another, must be analyzed at the level of the field or, to use Marshall's (1923) term, the industrial district. The field comprises not only buyers, sellers, and suppliers but also regulatory organizations and a host of other supporting institutions. In bringing together the literatures around the resource-based theory of the firm and the new institutionalism, I develop the concept of "raw capabilities," which are created by the social environment and not within the boundaries of the firm. Comparing firms in the different national settings reveals that what firms do, to a significant extent, is to combine these raw capabilities into firm-specific capabilities that may be competitive in the global marketplace. The institutional environments in which firms are embedded frequently confer competitive advantage precisely because the institutional environments are hard to replicate and hard to imitate: They typically develop incrementally over long periods, their causal structures tend to be imperfectly understood, and changes in their makeup typically require agreement among a large number of actors, whose interests often do not coincide.

Institutional theorists in sociology have marshaled considerable evidence that institutions assert control over organizations (Dobbin and Dowd, 2000); with the notable exception of Holm (1995) and Ingram and Inman (1996), however, few studies have examined how institutions come about. My detailed account of how universities and patent laws were shaped through collective action should, therefore, prove valuable reading for the institutional scholar who wants to develop a deeper understanding of how institutions come about.

Evolutionary economists (e.g., Dosi, 1984; Metcalfe and Gibbons, 1989; Klepper and Graddy, 1990; Winter, 1990; Saviotti and Metcalfe, 1991; Witt, 1992; Nelson, 1995a) can profit from this study in at least two ways. Ernst Homburg and I argued in an earlier paper (Murmann and Homburg, 2001) that empirical studies of industry evolution have typically focused on one particular country. As a result, we have very little understanding of how patterns of industry evolution may differ from one social setting to the next. Are firms founded and dissolved at the same rates? Why are some national industries much more successful than others? Do industry shakeouts occur in every setting? Do industries run through the same stages of development in each country? Because the synthetic dye industry started at the same time in different national settings, a study of this industry can provide some answers to these questions. Although the specialist in evolutionary economics will read this book primarily for its empirical evidence, my formulation of a coevolutionary theory may provide the impetus for a vigorous debate about how to model coevolutionary processes of economic phenomena.

Economic historians have debated for some time why Victorian Britain lost its economic leadership position to the United States and, in some industries, to Germany. The study of the synthetic dye industry before World War I can be read as another chapter in trying to come to terms with the relative decline of Great Britain. The theme that the economic historian may find most intriguing concerns the less powerful social network and the resulting collective action problems that the British dye industry faced in mounting successful lobbying campaigns against established industries. The British dye industry and the other new science-based industries were at an inherent disadvantage against powerful existing industries such as textiles. Germany industrializing later would, of course, encounter the same problems, but to a significantly smaller degree than Britain because the textile industry in Germany was not nearly as large in relative terms as in Britain. As Gerschenkron (1962) highlighted some time ago, it is crucial to pay attention to both when a new industry arises and how existing institutions may help or hinder the development of a new branch of industry.

Finally, a few words about what this book is not: It is not new history of the synthetic dye industry. The specialist historian of the synthetic dye industry will thus not find a revisionist interpretation of the synthetic dye industry before World War I. I have not uncovered new data that would call into question all existing accounts. Because a comprehensive database on all synthetic dye firms before 1914 has never been assembled, my data collection efforts with Ernst Homburg allow me to offer a systematic investigation of the organization of dye industry, shedding new light on how Germany was able to become and remain the dominant player in dye-making for so long. While I have made all efforts to get the history of the industry right, the purpose of the study is not to record all facts that are available on the synthetic dye industry or to provide the reader with a comprehensive handbook on the synthetic dye industry up to 1914. The reader who requires econometric analyses to be persuaded by an empirical argument will also be disappointed. I do not attempt to estimate an econometric model. The purpose of this book is to articulate a coevolutionary argument about the dynamics of industrial leadership and provide empirical data that give such arguments face validity. I hope the book can stimulate other scholars to test the arguments in a more systematic fashion.

Key Ideas in Evolutionary Theory

Let me preview the theoretical argument that I develop throughout the book. I have formulated the precise mechanisms of my coevolutionary theory after constructing a comparative narrative about the development of the synthetic dye industry in Britain, Germany, and the United States. All good theory development in the empirical sciences is a mixture of deductive and inductive reasoning that involves shuttling back and forth between the two modes of thought (Glaser and Strauss, 1967; Stinchcombe, 1968; Skocpol, 1984; Ragin, 1987; Eisenhardt, 1989a). By first studying in detail the evolution of the industry, I have placed more emphasis on the inductive aspect. I will now present a sketch of my coevolutionary theory, but a full articulation of the mechanisms that allow us to construe firms, technology, and institutions as coevolving will wait until I have given readers an opportunity to learn about the development of the synthetic dye industry. This will permit readers to develop their own notions about the causes of industrial leadership before they see the details of my model. Furthermore, those who are not predisposed to accept coevolutionary arguments can learn about the developments in the synthetic dye industry without having to see it through the lens of coevolution.

Evolutionary theory in the social sciences is sometimes misunderstood and seen as advocating social Darwinism, a view in which "successful" individuals in society have the right to trample and exploit its weakest members because nature works to let the strong prosper and the weak die out. Quite rightly, this so-called social Darwinism has been rejected as pseudoscience, and many people now greet the unreflective application of biological concepts to the social world with strong skepticism. Unfortunately, many people who for good reasons reject social Darwinist doctrines throw the baby out with the bath water by rejecting evolutionary arguments altogether. Moreover, many do not realize that the logical structure of an evolutionary theory is much broader than its biological versions (Campbell, 1969; Dawkins, 1976; Dennett, 1995; Hull, Langman, and Glenn, 2001). Evolutionary explanations have been applied to a diverse set of phenomena such as the development of the earth's geological features, economic change, and the development of languages.[7] Given that evolutionary theories of language existed long before Darwin, an evolutionary theory should not be interpreted as "biologizing" social theory before one has studied the specific arguments set forth.

The development of languages is perhaps the best example of evolutionary arguments that carry a much-needed neutral connotation and do not come with the heavy ideological baggage of biological concepts. The proposition that all languages spoken across the world today have evolved from one or a few common ancestor languages that branched out into families and subfamilies as people migrated to new places over the course of human history is widely accepted (Samuels, 1972; Ruhlen, 1994).[8] Similarly, the idea that today's vocabulary is evolving through a process of creating words for new phenomena – the World Wide Web, the Web, the Internet, electronic mail, e-mail, and so forth – (variation on existing words) and winnowing out a few of them that become most commonly used over the long run (selection) is not controversial (Müller, 1870).[9] In fact, in 1859, Darwin in *On the Origin of Species* (1964 facsimile, p. 422) appealed to the work of linguists to illustrate his ideas of a biological genealogy, underscoring the fact that evolutionary ideas existed before Darwin's application to the biological world. Hull (1995) recently examined the formal structures that characterize the evolution of language and biological life and concluded that they are essentially the same. Clearly, language is a better entry point to rigorous evolutionary theories in the social sciences than biology is.[10] One must

[7] Lewontin (1974, p. 6) has argued that an evolutionary perspective is equivalent to being interested in the change of the state of some universe in time, whether that universe consists of societies, languages, species, geological features, or stars. This definition is a bit too broad because it turns almost any development into an evolutionary account. As I will discuss later, the criteria for a rigorous evolutionary explanation are more narrow. But Lewontin also underlines the key point that evolutionary explanations are much broader than their biological versions.

[8] For an overview of the early connections between evolutionary thought in linguistics and biology, see Schleicher (1850) for the state of linguistics before Darwin. The essays by Schleicher, Bleek, and Haeckel presented in Schleicher (1850) give an early reaction in linguistics to the ideas of Darwin. For a comparison of intellectual developments in linguistics and geology in the nineteenth century, see the edited volume by Naumann, Plank, and Hofbauer (1992).

[9] For a contemporary evolutionary theory of language change, see Croft (2000).

[10] The best model for the evolution of firms, technology, and institutions is evolutionary epistemology, developed by Campbell (1974), Hull (1988), and others. As Hull (1995) points out, the evolution of languages has the property that evolutionary lines do not cross. Linguists report that two distinct languages

realize, however, that evolution does not necessarily imply progress or improvement, but rather cumulative and transmissible change. No society ever creates a new language from scratch without borrowing heavily from previous languages. Similarly, no society ever creates new industrial practices without drawing heavily on existing practices. As a general principle, novel things come about by changing and recombining existing things. The excellent summary description "descent with modification" crystallizes the key point of evolutionary theory into three words.

Identifying the formal structure that any evolutionary argument must possess to constitute a complete explanation[11] will make it easier to judge whether the theory I am proposing meets the formal requirements of an evolutionary explanation. More general than Nelson and Winter's (1982) articulation of evolutionary economics, evolutionary epistemology as articulated by Campbell (1974) and Hull (1988) provides, in my view, the most useful starting point for an evolutionary theory of industrial, institutional, and technological development. Expanding on Campbell's (1969) variation, selection, and retention model of evolutionary change, Durham (1991, p. 22) identifies five system requirements for an evolutionary theory of change:

R1. Units of transmission
R2. Sources of variation
R3. Mechanisms of transmission
R4. Processes of transformation
R5. Sources of isolation

An evolutionary explanation needs to identify clearly a unit of transmission – for example, genes, ideas, values, words, or even entire languages (R1).[12] It has to specify how these units are transmitted through time and space – for example, sexual intercourse in biology or social intercourse in culture (R3). It needs to say where variations come from – for example, gene mutation in biology or invention in culture (R2). And it has to articulate clearly the process that transforms the system through selection – for example, changes in the frequency of a trait in a population based on

have apparently never merged into one common language. Languages borrow words from one another, but linguists have not been able to find a true merger of two distinct languages. Industrial and conceptual evolution, by contrast, are replete with mergers of distinct lines of development. Industrial firms often merge and so do schools of thought. Because industrial evolution and evolutionary epistemology share key properties, evolutionary epistemology is, in my view, the best model for studying industrial phenomena.

[11] I draw here mostly on the excellent work of the philosopher David Hull (1988, 1989a) and the anthropologist William Durham (1991).

[12] Units of transmission come in different scales. Both individual words and entire languages can function as a unit of transmission. Later I will say more about how to deal with the different sizes of units of transmission and the relationship between them. The term "meme" is becoming increasingly popular to refer to nongenetic units of transmission. Susan Blackmore's *The Meme Machine* (1999) provides a useful introduction to memetics. I disagree with her insistence that memes are the only units of selection in cultural evolution. Sober and Wilson (1998) provide the most comprehensive argument for "group selection" as opposed to the "individual selection" that Blackmore, following Dawkins and other neo-Darwinians, is committed to. For a recent critical evaluation of meme literature, see the useful collection of essays for and against meme theory edited by Aunger (2000). Aunger (2002) lays out in more detail a neurological, brain-based model of memes.

natural selection through differential birth or death of variant organisms in biology or differential adoption of variant ideas in culture (R4). To account for stable differences in systems of the same kind (for example, how it is that two societies speak different languages), an evolutionary explanation needs to provide for sources of isolation (R5). If we are concerned with the evolution of one particular population, a source of isolation is not required. However, when we are trying to account for differences among populations, we need a source of isolation that introduces a boundary around each population and allows each to take its own specific path. Boundaries do not have to be sharp; they can operate as gradients, such that a population is very dense at a center but further out into the periphery becomes sparse. In principle, anything that fulfills these five systems requirements can evolve.

Another, complementary, way to assess the analytical rigor of an evolutionary account is to examine it through the lens of the conceptual tools developed by the philosopher David Hull (1989c, p. 96), who argues that a successful evolutionary explanation has to specify two entities, an interactor and a replicator. Hull defines a replicator as an entity that passes on its structure largely intact in successful replications and an interactor as an entity that interacts as a cohesive whole with its environment in such a way that this interaction causes replication to be differential. Essentially, the concept of a replicator fuses together Durham's concepts of a unit and a mechanism of transmission. Because genes are the replicator in biological evolution, the term meme has been coined by Dawkins (1976) to designate a nongenetic replicator for sociocultural evolution. The concept of an interactor maps into the concept of a process of transformation in Durham's scheme. What is replicated is information that is contained in the structure of the replicator. In more concrete terms, an interactor is an entity that exists in an environment characterized by competition for resources, survival, and replication. In the realm of biology, the human organism would constitute an interactor and the human genes a replicator. Using these two technical terms, Hull then can define selection as a process in which differential extinction and proliferation of interactors cause the differential perpetuation of the relevant replicators. Selection can be interpreted as a filtering device of information in which some information is passed on while other information is discarded. Notice that Hull's definitions do not make references to concrete entities. Quite purposefully, he defines replicators, interactors, and the selection process as abstractly as possible so that they can be applied to all phenomena that may be generated by an evolutionary process.

In applying the evolutionary model of variation, selection, and retention, the challenge for the researcher of a concrete phenomenon – in the present case, the evolution of industrial structures – is to specify how variants are introduced, how selection leaves behind variants that were not as fit according to the prevailing selection criteria (criteria that in turn need to be identified), and how some variants are retained over time to create a historical trajectory or genealogy captured by descent with modification.

Selection processes can vary dramatically across different phenomena. A powerful evolutionary theory needs to specify the most important selection process among a host of processes that transform a population of particular entities. In 1859, Darwin (1964), for instance, brought a great deal of specificity to his theory of biological evolution by arguing that natural selection (birth and death rates of individuals), not other forces[13] that are clearly also operating, was the most important force behind organic evolution. Trying to model Darwin's precision, Durham (1991) in his theory of cultural evolution identified selection by internal choice of individual human beings and not imposition by powerful external people (e.g., in a conquest of a particular society) nor the birth and death of individuals (natural selection) as the key driver behind cultural evolution.

Analogously, a theory of industrial evolution needs to specify its primary selection processes. In my view, there are two key selection processes in the case of populations of industrial firms. Organizational ecologists (Hannan and Freeman, 1989; Carroll and Hannan, 2000) have focused on births and deaths of individual firms. Evolutionary economists such as Nelson and Winter (1982), in contrast, and scholars of organizational adaptation (Child, 1972; Nadler, Tushman, and Nadler, 1997; Kraatz and Zajac, 2001) have stressed that organizations possess some capacity to adapt to changing environment by changing their strategies and structures. But scholars differ widely in their theoretical positions on how readily organizations can adapt to changing environments. Evolutionary economists view firms as strongly constrained in their ability to change, whereas scholars of strategic management typically grant organizations a much greater capacity for change. On the adaptationist model, selection of particular organizational traits takes place through managerial decision making. The relative importance of these two selection processes is one of the big questions that has not been sorted out; in my mind, this should be a key item on the agenda of organization theory. Although addressing this question is beyond the scope of my book, I note, however, that selection by individual and collective choices played an important role in the evolution of the synthetic dye industry. The characteristics of national patent laws and university systems were clearly selected in part through the collective lobbying efforts of actors in industry, academe, and government.

For a long time researchers in many fields have debated what constitutes appropriate units of selection in evolutionary models. Hull (1989c) introduced the concept of replicators and interactors in part to clarify that there are two processes in selection, the making of copies of an entity (replication) and the differential survival of copies caused by environmental interaction. Hull et al. (2001) have argued persuasively that instead of asking, "What is the appropriate unit of selection?", researchers need to pose two distinct questions, namely, "What is the unit of replication?" and "What is the unit of environmental interaction?" Another key challenge, then, in constructing an evolutionary theory is to identify clearly what is the unit of replication and what is the unit of interaction with regard to the phenomenon to be explained. Before moving on to the evolution of industrial phenomena, I will identify a few other important features of a successful evolutionary explanation. We will see a little later how important it is to clarify what constitutes the unit of analysis to be able to study effectively the evolution of industrial phenomena.

In the debate about appropriate units of analysis, it is often forgotten that evolutionary theory is inherently a multilevel theory. Any evolutionary analysis requires at

[13] Darwin (1859) ended the introduction to *On the Origin of Species* with the line, "I am convinced that Natural Selection has been the main but not the exclusive means of modification" (1964 facsimile, p. 6). Later in the book he identified migration (p. 81) and sexual selection (competition for mates; pp. 87–8) as other, less important forces in the evolution of biological populations.

least two levels, a level identifying particular individuals that reproduce at differential rates and a level specifying a particular population that is the locus of evolutionary change. A simple abstract example illustrates this point. Consider a population of 100 individual balls, 50 being red and 50 being white (the example is chosen to underscore that evolutionary explanations apply broadly and not just to biology). Let red balls replicate at a greater rate than white balls but let both groups go out of existence at the same rate. If we examine what percentage of balls in the population is red from period to period, we would find that the initial 50 percent increases gradually to 100 percent. In evolutionary terms, the population of balls evolved with respect to its color distribution. If the relative rate of the replication and destruction of red and white balls is different from the one in our example, we would see very different evolutionary outcomes in the relative frequencies of white and red balls. Note that none of the individual balls "evolved" in any sense, because none changed its color. The first key point is that populations, not individuals, evolve. The second is that an evolutionary explanation requires the analyst to specify who are the "individuals" and what is the population in the analysis.[14]

Just as in other sciences, researchers of industrial evolution have debated extensively as to what constitutes appropriate units of selection and what units of selection are more important than others for industrial change (Aldrich, 1999, pp. 35–41, 336–40). Most studies have focused on routines or entire organizations as the units of selection (Aldrich, 1999, p. 35). But there are clearly other possibilities. Social activity is analogous to the physical and biological world, in that it is organized in multiple levels (Simon, 1981). Starting at the microlevel are the actions and reactions of individual human beings (Weick, 1979); at the next higher level are work groups and teams in organizations; and then business units within organizations; entire organizations; entire industries; communities of industries; the organization of national economies; and finally, at the most macrolevel, the world economy.[15] Given that an evolutionary analysis always requires two levels – individuals and a population – the most macrolevel, the entire world economy, logically cannot serve as a unit of selection. In principle, all lower levels could function as units of selection.

The primary units of selection used thus far in organization theory have been either entire organizations or routines/competencies (Aldrich, 1999, p. 35). There is of course no theoretical reason why other units of analysis at intermediate or higher levels of aggregation, such as groups within organizations or organizational divisions or even entire industries, cannot serve as units of selection in evolutionary analyses. As anthropologist Durham (1991, p. 428) remarks regarding organic evolution, "In principle, of course, the potential exists for natural selection to operate at any level in the organizational hierarchy of life where there are variable, reproducing entities (see for example, [. . .] Vrba and Eldredge 1984; Eldredge 1985)."[16] Just as is the case

[14] Organic evolution traditionally was conceived as a process that involved three distinct levels of organization. The standard formula for organic evolution went as follows: Genes mutate, organisms compete and are selected, and species evolve. Hull (1989b, p. 83) reports that biologists now know that this formula is too simple. The central idea in the present context is that evolutionary explanations require the researcher to track at least two levels.

[15] This list is not exhaustive. Other possible units of analysis lie in scale between the ones I mention.

[16] Evolutionary biologists have not as yet reached a consensus on whether selection processes occur at any level in the organizational hierarchy of life.

with organic evolution (and I suspect with other domains too), we have very little systematic understanding of the relative importance of the selection process at each of the different levels of analysis in shaping social and industrial evolution. In my analysis of the synthetic dye industry, I operate at different levels of analysis – routines within specific firms, individual products, firms, national industries, the global industry – without systematically sorting out the relative importance of each level. Unfortunately, I can again only highlight what is still missing in a complete theory of organizational evolution and must leave the task of filling this gap for future research.

An important step toward collecting the necessary data that will make it possible to sort the relative importance of different levels of analysis in industrial change is that researchers become much more careful in specifying their level of analysis. Michael Tushman and I (Tushman and Murmann, 1998) recently undertook a systematic analysis of the literature on technological evolution and dominant designs. We found that researchers more often than not confused these different levels. As a result, empirical investigations of a large number of technologies have not led to a cumulative body of knowledge. What is needed in the study of industrial change, just as in the study of technological change, is that researchers clearly specify their level(s) of analysis. In particular, this requires that researchers identify the level of analysis that lies directly above and directly below the level that is the focus of their investigation (Murmann and Tushman, 2001).

Until now the debate in organization theory has not incorporated Hull's (1989c, p. 96) insight that an evolutionary analysis becomes more transparent by specifying two units, an interactor and a replicator. Another step toward achieving greater precision in regard to the units of analysis in organizational analysis is to specify clearly the replicator and the interactor in a particular analysis. Hull (1989c, p. 95) explains that good candidates for replicators meet the general criteria of longevity, fecundity, and fidelity. Ideas, mental representations, models of management, organizational cultures, and rules in operating manuals in firms are examples of possible units of analysis for replicators in organization theory. Candidates for interactors, entities that interact with the environment and cause replication to be differential, are standard behavior (or routines, in the language of Nelson and Winter, 1982[17]) – individual human beings, work groups, divisions, entire organizations, and so on.

Many writers make the point that evolutionary theory is explanatory but not predictive. While true that an evolutionary theory cannot make point predictions – that is, foretell exactly, in every conceivable detail, what is going to happen in an evolutionary system tomorrow – in many instances knowledge about the path and present state of the system that is evolving allows one to make broad predictions. Just as we can predict with confidence that dinosaurs will not evolve by tomorrow out of today's existing animal species, it is safe to predict that the African nation of Uganda will not emerge tomorrow as the largest chip producer in the world. Evolutionary

[17] Nelson and Winter (1982) not only use the term "routines" to refer to expressed behavior but also liken the role of routines in organizations to genes in biological evolution. As the ambiguities in the subsequent literature shows, I believe it is better to reserve "routine" to indicate expressed behavior (the interactor) and not the mental representation (the replicator). See Cohen et al. (1996) for a good discussion of routines.

theory does not provide exact predictions of future outcomes, but it does provide scientifically rigorous explanations.

Finally, another important characteristic of evolutionary explanations is that they are probabilistic rather than deterministic (Sober, 1984, pp. 110–34; McKelvey, 1997; Aldrich, 1999, pp. 33–5, 200–1). While some theories predict that, given condition *x*, *y* is going to follow, evolutionary theory can only make statements of the kind that, given *x*, *y* will follow with a probability of *z*. Let *x* be the selection environment and *y* the fate of a particular individual in the population. Rigorous evolutionary theories will make probabilistic statements like this: There is a *z* probability that individual *y* will not replicate (die when the entity has a limited life span) under the selection environment *x*.

In economics, management, and organization theory, evolutionary theories have been developed for two important sets of reasons: (1) to explain industrial and organizational structures not solely in terms of the agent's intentions but in terms of the consequences of the agent's actions, and (2) to recast the fundamental assumptions of how human beings make decisions and behave. Unlike many other theories of action, an evolutionary account focuses on consequences and not intentions. If actions generate positive outcomes under the prevailing selection criteria (even if the outcomes are unintended), they are selected for by the environment and persist. If they generate negative outcomes, actions will be selected against and fade away.

Evolutionary theories try to overcome unrealistic assumptions about human beings in rational choice theories. Instead of viewing individual human beings as perfect calculating machines that can consider all possible alternatives for a particular decision, know the future perfectly (at least in probabilistic terms), and then calculate the optimal course of action, evolutionary theories see human beings as fundamentally limited in their ability to consider alternatives, foresee the future, and make error-free calculations (Potts, 2000). In Nelson and Winter's (1982) evolutionary theory of economic change, for example, behavior is less guided by calculating consequences than by following rules that have been developed for a specific situation. Because human organizations are composed of collections of boundedly rational individuals – to use Simon's (1981) famous term – they also make decisions largely by following rules or routines, Nelson and Winter's term for the predictable aspects in the behavior of organizations. (How preferences of individual human beings are integrated to form a "preference function" of an organization is, of course, one of the fundamental questions in organization theory.)

For the most part, organizations act on a simple principle: If a given routine works, let's do more of it; if it does not work, let's do less (March, 1999). In the language of evolutionary theory, actions or routines within organizations are selected for the perceived benefits they generate. Because the organization-wide implications of particular routines are often difficult to determine – especially if routines are highly interdependent with other routines – it is bundles of routines and not individual routines that are selected. The selection mechanism in this case is the market, which eliminates entire organizations that do not return an adequate profit.[18] When

[18] In line with the preceding discussion of units of analysis in social organization, it is, of course, possible that an organization is a division of a larger organization. Then the proximate selection mechanism

organizations fail because their collection of routines is not efficient compared with that of rival organizations, a particular routine goes out of existence together with all the other routines of the organizations.[19] Notice that we are dealing with two different units of analysis, an individual routine and an entire organization.

An evolutionary view of organizations has profound implications for how one will see industrial development, as will be detailed in Chapter 5. Here, I just want to make a few introductory remarks on how "the world of industry" looks through the perspective of an evolutionary theory. Because organizations consist of routines that interact in highly complex ways, managers more often than not find it difficult to figure out what makes their organization successful. This causal ambiguity is not a problem as long as the environment does not change. Managers can guide their organizations to replicate existing models of action as faithfully as possible without understanding the causal microstructure that makes their collection of routines successful. When the environment changes, however, this causal ambiguity makes it difficult to determine what the organization will need to do differently. The very process of constructing complex organizational routines that make an organization reliable poses an enormous challenge in the face of changing selection criteria. The evolutionary perspective of organizational change takes into full account the inertia that arises from existing routines. It highlights the fact that the historical path of a specific organization imposes significant constraints on where it can go (Aldrich, 1999, pp. 200–22).

The literature on organizations is filled with examples of organizations that were unable to transform themselves despite the best intentions of top management (Tushman and Anderson, 1997; Tushman and O'Reilly, 1997). After many unsuccessful attempts to change its organization, General Motors (GM), for instance, realized that introducing new work and management processes into its existing divisions for the purpose of producing small cars at high quality and low cost was likely to fail. Instead, it decided to create from scratch a new division to make small cars, Saturn (Woodruff, 1992). Saturn was given an enormous amount of independence because GM wanted to build an organization with entirely new routines and not allow the existing divisions to impose their routines, values, and structures on the small car operation. The GM example highlights that change within an organization often comes about by creating new, relatively autonomous, divisions, by removing resources from existing divisions, by selling off divisions, or by closing them down altogether.

Individual organizations are constrained not only because of their own history but also because they exist in larger social environments that impose additional limits on the directions an organization can take. Implicit in the framework of this book

involves managers at headquarters, and not the market, to decide whether to sell or close the division. Business units that vary in their profit rates contain variable reproducing units, which interact with an environment. Managers play the role of the proximate agents of selection by giving more money to invest in growth to (selecting for) profitable business units and by pulling resources out of (selecting against) relatively unprofitable units.

[19] Arthur Stinchcombe (personal communication, 2000) reminded me that strictly speaking this statement is not true if the same routine exists in another organization where it can reproduce. As long as the routine has a separate evolutionary dynamic, it can flourish even after the "death" of the organization. Stinchcombe pointed out that many of Napoleon's tactics and strategy routines survived because of the successful adoption by his enemies, which, for example, destroyed the French army that had conquered parts of Italy and much of Russia and Germany.

is that individual countries often are the appropriate unit of analysis for the impact of the larger social environment on a particular organization. Organizations typically borrow blueprints and management models from other organizations, whether they be competitors, suppliers, or customers. Practices that are prevalent or even taken for granted in a particular social environment are easier and cheaper to import into the organization than are practices that run counter to the way actors and organizations that are close in social space are organized (Tilly, 1998, pp. 76–83).

If individual organizations and clusters of organizations surrounded by the same social environment are limited in their capacity to adapt to an environment in flux, what are the public policy implications for change at the industry level? One key implication is that significant change at the industry (or population) level is likely to come from the birth of new organizations with different organizational routines and the exit of established organizations. Efforts to encourage existing organizations to transform themselves radically, as in the case of converting defense contractors to makers of consumer products, may be bound to fail. The routines that make an organization successful in getting repeat business from the military are quite different from the routines for marketing and selling to the private sector. Hence, public policy makers have to weigh carefully whether a better way to foster innovation and change in an industry may be to encourage the formation of altogether novel enterprises that are not constrained by historically accumulated and well-entrenched routines.

Evolution of Technology

Before offering a sketch of my coevolutionary theory, let me briefly lay out what it means for technology and national institutions to evolve. Having now identified the general requirements for a rigorous evolutionary theory, it is not difficult for us to identify the key features of an evolutionary view of technology and national institutions. Articulating an evolutionary framework for each domain will make it easier to spell out the more complex coevolutionary model that links the two with the evolution of industrial structures.

Scholars have long recognized that technological change plays an important role in industrial dynamics and economic growth (Usher, 1954; Landes, 1969; Rosenberg, 1982). Over the past two decades, empirical research has documented that technological innovations not only are able to create new products and industries but also can radically destroy the fortunes of existing firms or even eliminate entire industries altogether (Schumpeter, 1934; Gort and Klepper, 1982; Nelson and Winter, 1982; Tushman and Anderson, 1986; Anderson and Tushman, 1990). Many historians of technology have argued that the development of technology is best understood as an evolutionary process (Basalla, 1988; Vincenti, 1990; Petroski, 1992; Ziman, 2000).

Evolutionary approaches to technological change come in a variety of flavors. Let us first identify what they have in common. All theories share the idea that the invention process by scientists, engineers, and tinkerers provides the sources of variation (Durham's R2) necessary for an evolutionary account of technological change. Scholars also agree that social intercourse provides the mechanism of transmission (Durham's R3). The most important difference between theories lies in their object of analysis, that is, the "thing" that is seen as evolving. Some scholars such as Basalla (1988) and Ziman (1999) conceptualize technology primarily in terms of the physical artifact. Basalla studied in detail how physical attributes of artifacts, for example, the shape of axes, have changed over the centuries. On this materialist view, technological evolution occurs because the relative frequencies of artifacts with a certain shape change over time. Other scholars argue that what is central in technological evolution is not the physical artifact but the science, engineering, and design knowledge that goes into making a particular artifact. In the language of Hull (1989c) these scholars argue that technological evolution needs to be understood in terms of both a replicator (ideas, knowledge, etc.) and an interactor (the physical artifact). Behind this view lies the observation that particular technological principles often diffuse widely in the economy. Steam technology a couple of hundred years ago or integrated circuits during the past few decades, for instance, are seen as such important technologies because the knowledge behind these two technologies found their way into a great many different artifacts and thereby transformed the economy.

Following the pioneering work of Nelson and Winter (1982), Mokyr (1999) recently proposed a selectionist model of technological change that focuses on knowledge, not concrete physical artifacts, as the most relevant object of analysis. Mokyr's (1999) model distinguishes between two types of knowledge, knowledge of what and why (a type he calls omega) and knowledge of how (a type he calls lambda). Scientific and basic engineering theories, for example, would fall into the what and why category. In contrast, blueprints and instructions of how to make things (be they verbal, oral, or pictorial) would fall into the how category. Because the different pieces of knowledge in omega can be combined in a multitude of ways, they give rise to a virtually endless number of possible procedures (or techniques, as Mokyr calls them) for making things. From this set of all possible techniques, few are in use at any time. What evolves in Mokyr's theory is the relative frequency of techniques in existence at any time. Ideas serve as the units of transmission (Durham's R1). Given that there are always a variety of alternative ways to accomplish a particular technological goal, techniques compete with one another for adopters. The selection mechanism that transforms populations of techniques (Durham's R4) is the choice of users to adopt or abandon a certain technique. Those techniques that find a larger number of adopters become more prevalent in a particular technological domain, and those that lose users and do not gain new adopters shrink in importance or may go out of existence altogether.

Scholars of technological change also differ in their views of the selection criteria that lead to differential replication among technological variants. On the one extreme are those who believe that the best technology always succeeds against inferior alternatives. Let us call them the "strong technological efficiency school" (see Liebowitz and Margolis, 1999, for a statement of the technological efficiency position). On the other extreme are scholars who believe that social criteria, such as political and economic power of particular groups, determine which technological variants thrive and which ones fail. Let us call them the "strong social constructionists" (see Bijker, Hughes, and Pinch, 1987, and Bijker and Law, 1992, for a statement of the social constructionist position). Many scholars lie somewhere in between these extremes, holding the view that under some circumstances social factors dominate selection, and under other conditions technical factors dominate (see Vincenti, 1991, for a statement of this position). Our previous discussion of the requirements for a rigorous evolutionary theory makes it plain that both views can be incorporated into an evolutionary model. As long as the theory specifies clearly the selection criterion – be it efficiency or social power or something else – it can constitute a proper evolutionary explanation. A little later, in discussing the economics of dye-making, I will give evidence that technical efficiency was the most important selection criterion in the case of synthetic dyes.

Let us return to the debate about the appropriate object of analysis in the evolution of technology. Ziman (1999) is well aware that technological artifacts are embedded in bodies of knowledge and social practices, but he argues that focusing on artifacts avoids the enormous problems of operationalizing a theory, à la Mokyr, that is based on knowledge and ideas as the objects of analysis. Let us recognize again that the formal requirements of an evolutionary theory are flexible enough to make it possible to conceptualize the evolution of technology in terms of both artifacts and ideas. In the long run, I believe, we should strive to build a rigorous evolutionary epistemology of technical change as conceived by Nelson and Winter (1982) and Mokyr (1999). We need to work out the specifics of how to operationalize and measure the key concepts in such a theory. Fortunately, there is no need to start from ground zero because Hull's (1988) work on science as an evolutionary process can provide important guideposts on the road toward a rigorous evolutionary epistemology of technology that articulates precisely what would constitute a compelling replicator and what would constitute an appropriate interactor for an evolutionary model of technical change. In the short run, however, we should not be rigid in terms of what aspect of technology we regard as evolving because this would put a brake on doing empirical studies of technological evolution. In my analysis of synthetic dye technology I take such a pragmatic approach. At times I find it useful to talk in terms of knowledge evolving and at other times I find it convenient to point to dyes (defined by their chemical structures) as the things that evolve.

Evolution of Institutions

In this book, I will focus on the evolution of national institutions, but the arguments I develop can be applied equally to lower levels of aggregations such as states, regions, or cities. Generally, institutions are defined either in terms of persistent patterns of actions or in terms of enduring patterns of ideas and values. Because there are almost as many uses of the word "institution" as there are authors (see Nelson and Sampat, 2001, for a recent overview), it is important that I state clearly what I mean by the term. I use the term to denote actions, rules, social structures, and practices that persist over time and are features of social aggregates that are larger than a single organization. If one dye firm has certain practices that remain moderately stable for many years, I would not count that practice as an institution. Only when a practice exists in many dye firms operating in the same environment would it qualify as an institution by this definition. Remember that for an evolutionary explanation we need both a population and an individual exemplar. Patent laws and associated practices, for example, apply to all firms in a particular country and typically remain stable for considerable periods of time. Similarly, universities in the same country share many features (appointment requirements, department versus chair structure, kind of degrees offered, etc.), which allows me to talk about these features as (national) institutions on my definition.[20] Because alternative institutions differ in how efficiently they coordinate human

[20] The way I use the term "institution" may remind some readers of North's (1990, p. 4) useful distinction between institutions, which he defines as the rules of the game, and organizations, which he defines as the players in the game. I differ from North in emphasizing that organizations are very active in shaping the rules of the game.

efforts for a particular purpose, institutions can have a profound influence on the performance of a particular industry or economy (North, 1990; Nelson and Sampat, 2001).

As discussed earlier, a rigorous evolutionary analysis needs to operate at both the level of a population and the level of individuals that are selected. What, for example, are possible populations and individuals in the case of patent laws? One approach is to see all different formulations of patent laws that exist the world over at any one point in time as the population from which each country selects its particular patent law. Scholars of comparative patent law have documented that when a country is about to implement, change, or abandon its patent laws, it always surveys the different patent laws around the world for ideas of how to design its own system (Penrose, 1951). On this first approach, a country's patent law serves as the individual that is selected for or against based on the selection pressures that are prevailing at the time. A second approach moves the unit of analysis down a level and treats each national patent law and associated practices as a population. In this case the individuals that are selected for and against could be, for example, (1) the individual rules that make up a national patent law and practices or (2) court cases. One could track how the frequency of rules that favor inventors over other parties changes over time. Or one could track the frequency of court rulings that nullify patents after they were granted by the patent office. Judges in this case would be the proximate agents of selection, who select particular practices based on their understanding of the law and the prevailing social climate. Focusing on court rulings is analytically appealing because sometimes patent laws do not change in their wording but their interpretations do.

Why is it useful to conceptualize changes of patent laws over time as an evolutionary process? First of all, national patent laws are marked by substantial diversity, especially as we go back in time. The history of patent laws shows that there is a great deal of trial and error in countries' attempts to design their patent laws. The Netherlands, for example, had a patent law but abolished it in 1869 because it was seen as hurting the economic development of the country (Penrose, 1951). Just as The Netherlands was abolishing its law, newly unified Germany created a national patent system that had a dramatic impact on the innovation strategies of dye firms. Some countries grant patents to the first person to file the patent claim, whereas other countries grant patents to the first to invent. Typically, policy makers introduce new regulations without being able to foresee the consequences of implementing a particular rule (Hayek, 1973). Only over time does it become apparent whether a particular patent rule has desirable consequences for the people who have the power to make rules. Patent laws evolve over time precisely because desirable rules are retained and undesirable rules are selected out by the prevailing selection regime. But as we shall see, the evolutionary explanations for the development of patent laws are not nearly as well worked out as those for industrial development because in the former case it is much more difficult to specify as clear and simple a selection criterion as profitability in the case of industrial evolution.

Similarly, the development of the global population of universities or, at a lower level of aggregation, national populations of universities meets the criteria of an "evolutionary system" à la Durham. Universities or colleges shutting their doors does not make the headlines these days. But looking back in time, one finds that a sizable number of universities have closed, leading to change in the population of national universities. Of the 1,990 four-year colleges founded in the United States between 1636 and 1973, 515 had gone out of existence by 1973 (Marshall, 1995). A map of German universities in 1900 by Franz Eulenberg (1904) shows a surprisingly large number of universities that existed for some time and then were closed. Preceding the wave of the last two decades, private universities were already formed in Germany in the beginning of the twentieth century but they invariably failed or were taken over by governments and turned into public institutions. I will focus in my analysis on national university systems as the populations whose evolution we will track. I chose this unit of analysis because the organizational structure of universities typically displays more variations across than within countries. An evolutionary analysis of national university populations could track, for instance, the relative frequencies of private versus public universities. Or, as Aldrich (1999, pp. 177–80) points out, one could examine the relative frequencies of single-sex versus coeducational colleges in the United States. Recall that an evolutionary analysis can pick out any trait or characteristic and then trace how the frequency of that particular trait or characteristic changes over time in the population.

Given my focus on the synthetic dye industry, I examine over time the relative importance of research and teaching in organic chemistry in the British, German, and U.S. university populations. Among the variables I track for each national population are the frequencies of chemistry professors and students in chemistry.[21] To support an evolutionary line of argumentation, one must show that certain variants fail because of the particular selection criteria for universities prevailing in a country at a given time. I will show that the differences in the characteristics of the three national university populations can be attributed to differences in the selection environment. But to reemphasize the key argument of this book, the selection criteria prevailing in each country have in part been shaped by the collective actions or inaction of dye industry participants. Because the evolution of the three national university populations were causally linked to the evolution of the three national dye industries, it is proper to speak about a coevolutionary process.

A Sketch of Coevolution

Such a diverse group of scholars as Kauffman (1995, e.g., pp. 279–98), Lewin et al. (1999), Nelson (1995a), and Ziman (1999) have argued that we need to develop coevolutionary models to better understand the dynamics of industrial change. This book responds to their call and takes a significant step forward in articulating a powerful coevolutionary theory that links industrial, technological, and institutional dynamics. To avoid any misunderstanding, it is important to clarify at the outset that I use the prefix "co-" in coevolution not in the restricted sense that two things are evolving together but in the broader sense that multiple things are jointly evolving. Unlike those writings in which coevolution means the parallel development of two entities, my study analyzes the coevolution of national firm populations, technology, and two different kinds of national institutions, namely, research and training systems and patent practices. I will now briefly introduce the key ideas behind my coevolutionary model. It is fleshed out in full detail in Chapter 5.

[21] Ideally, one would have tracked the funding provided for chemistry research relative to funding for other disciplines. Unfortunately, these data are not available on a systematic basis before World War I.

In the preceding pages we have developed a general theory of evolutionary change. But what does it mean for things to coevolve? Scholars have used the term coevolution in several different ways. To make a theory testable, one must be very clear about the meaning of its central concepts. Coevolutionary arguments are beginning to receive more attention in organization and management theory (Kieser, 1989; Yates, 1993; Baum and Singh, 1994; March, 1994; Levinthal and Myatt, 1995; Barnett and Hansen, 1996; Haveman and Rao, 1997; McKelvey, 1997, 1999; Coriat and Dosi, 1998; Koza and Lewin, 1998; Baum and McKelvey, 1999; Lewin and Volberda, 1999; Lewin et al., 1999; Van De Ven and Grazman, 1999; Eisenhardt and Galunic, 2000), but because researchers are using coevolutionary language often in a very imprecise or inconsistent manner, they have invited unnecessary criticism. Some observers of the present state of coevolutionary scholarship in organization theory jump to the invalid conclusion that in coevolutionary explanations everything seems to be coevolving with everything else and hence cannot provide a parsimonious explanation. In biology, for instance, it is not the case that every species is coevolving with every other species in the world. In many cases, coevolution takes place between two species, for example, a particular plant and a particular insect, the former serving as food and the latter as an instrument for spreading the pollen (Thompson, 1994). Coevolutionary relationships frequently also exist between a predator and its prey (Nitecki, 1983). Similarly, a particular industry coevolves to a significant extent only with a very restricted number of other industries and surrounding social institutions. Often a coevolutionary relationship exists between producers and user populations, as in the case of the tabulating industry and the life insurance industry documented by Yates (1993). At other times a coevolutionary relationship exists between two populations of competing technologies, such as propellers and jet engines that power airplanes.

Let me try, then, to provide a precise definition of my use of the term coevolution. Two evolving populations coevolve if and only if they both have a significant causal impact on each other's ability to persist.[22] Such causal influence can proceed through two avenues: (1) by altering the selection criteria or (2) by changing the replicative capacity of individuals in the population without necessarily altering the selection criteria. Kauffman (1993) uses the idea of coupled fitness landscapes[23] to express this conception of coevolution. In coevolution à la Kauffman, one partner deforms the fitness landscape of the second partner and vice versa. As a result, a coevolutionary relationship between entities can increase the average fitness of both populations, decrease the average fitness of both, or have a negative or positive impact on the average fitness of one but not the other. Whether a coevolutionary process is beneficial or harmful for the parties involved depends on the particular causal

[22] My definition of coevolution is very similar to Nitecki's (1983, p. 1): "Coevolution occurs when the direct or indirect interaction of two or more evolving units produces an evolutionary response in each."

[23] Borrowing from Wright (1931, 1932), Kauffman (1993, pp. 33–4) defines a fitness landscape as the distribution of fitness values over the space of possible genotypes for individuals in a population. According to Kauffman, each possible genotype has a particular fitness value. The population on this view is a tight or loose cluster of individuals located at different points in the landscape. In this model, adaptive evolution in a population amounts to a hill-climbing process. When a population evolves, selection will lead to the reduction of some genotypes and the proliferation of others. As a result, over time the cluster of individuals representing the population will "flow" over the fitness landscape.

relationship that links the parties; therefore, this relationship needs to be specified in the empirical analysis.

My definition differs from Durham's (1991, pp. 171, 205–13) meaning of coevolution. Durham uses the term more broadly to include the case in which two factors (in his case genes and culture) have an effect on a third factor (in his case human behavioral diversity) without necessarily standing in a causal relationship with one another. In Durham's scheme, genes evolve and culture evolves, and they both have independent causal effects on diversity in human behavior. Whereas Durham allows parallel evolution to count as coevolution, I restrict my definition to Durham's narrow meaning of coevolution in which two evolving entities interact causally with one another. This restriction helps make a theory of coevolution in the social sciences sufficiently precise to lead to compelling explanations.

What is central about a coevolutionary process, in my use of the term, is the bidirectional causality linking the two parties in the relationship. To understand why Germany's dye industry was able to capture a dominant market position and maintain it for a number of decades before World War I, it is necessary, I argue, to examine the causal links between the national populations of industrial firms and the national populations of universities. Coevolutionary dynamics can lead to self-reinforcing processes that translate small initial differences in the performance of national industries into large differences over time. The key challenge for such arguments is to establish that causal processes indeed do connect the two partners in a coevolutionary relationship. To claim, for example, that technology coevolves with national firm populations, one must establish a precise mechanism by which this reciprocal influence takes place.

Arthur Stinchcombe (2000) proposes that we need to examine the "physiology" of the evolving social system to do a proper evolutionary analysis. By extension, a coevolutionary analysis should examine the physiology that links two or more evolving systems. In the case of the synthetic dye industry, we need to investigate flows that link firms with universities and patent practices. I will argue that the transfer of synthetic organic chemists between firms and universities constituted one important such flow that allowed this reciprocal influence to take place.

A second important flow involved university professors, who offered their expertise and reputations in the realm of patent laws and practices. University professors were typically used by the state to help draft patent laws or they were hired as expert witnesses in patent litigation, providing the professors the opportunity to influence patent laws and practices. Because these cross-flows varied from one country to the next, the competitive strengths of the three national dye industries evolved along very different paths. In Germany, dye industry firms were much more able to alter selection criteria in their favor than were firms in Britain or the United States.

Using cross-flows to establish evidence of reciprocal influence between two coevolving partners is particularly important because not everything that looks like coevolution is really coevolution. As Nitecki (1983) points out, coevolution may be mimicked by such things as sequential adaptations from different causes or simultaneous adaptation to the same environment. The only way to establish true coevolution as opposed to spurious coevolution is to gather evidence of cross-flows among the alleged coevolving systems. I try to do so in the body of this book by examining closely the career biographies of academics, industrialists, and members of relevant political organizations. I uncover links between these different populations of actors to show that they influenced each other's evolutionary trajectory.

Before attempting to solve the dye industry puzzle, I need to equip the reader with sufficient knowledge about dye technology, the changing role of science in the innovation process, and the economics of dye-making. Without such knowledge, one easily could get lost along the way. For those readers who have no knowledge of dye chemistry and dyeing processes, it will be useful to read Appendix I, which provides a more detailed technological history of dyes and dyeing, before starting Chapter 2. I begin with a short overview of the economics of dye-making and then move into a discussion of dye technology.

The Economics and Science of Dye-Making

The well-established market for natural dyes made it relatively simple to determine how synthetic dyes in 1857 might be competitive. For Perkin to win customers for the first synthetic dye, his novel purple aniline dye had to meet two economic criteria. First, the cost of the new dye had to be low enough for some segments of the dyeing and printing trade to find the product appealing. The actual cost of a dye was determined not only by what the dye maker charges for a certain quantity but also by how expensive it was to apply the dye to a textile and how long the color will resist wear and tear. (Not surprisingly, fashion concerns sometimes turn normal quality concerns on its head: Nowadays many blue jeans are colored on purpose with a dye that fades readily so that the fashion-conscious owner can enjoy worn-looking pants after a short while [Seefelder, 1994]). Second, if the cost were equal to or more than that of natural dyes, the dye had to offer greater quality either by delivering a shade that was not available in natural dyes or by offering a more consistent shading from one product shipment to the next. (Appendix I provides a detailed account of the differences between natural and synthetic dyes.) As Perkin found out by consulting with dyers, his aniline purple was clearly competitive along the quality dimension. Hence the crucial question in pioneering the production of the first synthetic dye became, "Can the product be made cheaply enough?" The key factor in this equation was how much the chemicals would cost that were required to make aniline dyes. The basic chemicals were widely available in the market, making their costs rather straightforward to forecast. But how expensive would it be to make nitrobenzene and then aniline, which Perkin converted into the purple dye? Because no market existed for nitrobenzene and aniline, two of the key organic raw materials, Perkin had to learn how to manufacture himself the two compounds on a commercial scale. By making a series of production innovations – for instance, using iron instead of glass vessels as was then common practice in chemical laboratories – Perkin was able to manufacture the intermediates cheaply enough to make aniline purple a big commercial success, especially once the color became popular in French fashion circles in late 1858.[24]

This short description of the first dye production already gives an indication of the multiplicity of production steps that were typically involved in converting the basic raw material, coal tar, into commercial dyes. For the lay person it helps to conceptualize a synthetic dye company by drawing an analogy to a kitchen in a large

[24] *Mauve*, a recent biography of Perkin (Garfield, 2001), portrays in detail (pp. 60–73) the fashion craze that surrounded aniline purple once it was popularized in France under the name of mauve.

restaurant. Basic foods and ingredients are bought from outside vendors. The chef mixes basic ingredients to make sauces and flavors that are stored for later use. In preparing a particular dish, the chef then draws on many basic foods and materials, as well as preassembled sauces and flavors, and combines them according to the proper recipes. Some chefs like to make the "intermediate" sauces and flavors themselves. Others buy them ready-made from vendors. Similarly, synthetic dye firms in the later stages of the industry had the option of buying dye intermediates from the market or making them themselves.

Perkin held a British patent on aniline purple and hence possessed a monopoly on its manufacture, but other aniline dyes, such as fuchsine, marketed only a year later, were made by several firms. As a result, price competition between firms became a regular feature of the synthetic dye industry early on. Market prices for synthetic dyes were constantly falling, similar to those for personal computers in our own day. A few datapoints illustrate this trend: By 1864, the price of fuchsine (aniline red) had fallen to about 10 percent of the 1860 levels (Morris and Travis, 1992, p. 65). Borscheid (1976, p. 132) reports that alizarin prices fell from 270 marks (1869–1871) to 40 marks in 1877 and to 9 marks in 1886. Prices for dyes in part came down because raw material prices declined. According to Hückstädt (1967, p. 387), British benzene prices declined in a short two-year period from 12 shillings (1883) to 3.5 shillings (1885). These constant price drops between 1857 and 1914 reflected both the improved methods of large-scale production and the high level of competition. To stay competitive, a dye firm had to be able to make its own production cheaper; those who could not were forced to close.

At the start of the synthetic dye industry in 1857, participants could have hardly foreseen that by World War I synthetic dyes would virtually replace natural dyes. Just as in the case of electric lighting and the internal combustion engine, which replaced gas lighting and steam engines, respectively, the new synthetic dye technology in time proved to be vastly superior in cost and often in quality, relegating natural dyes to small niche markets. The size of the market for synthetic dyes increased dramatically. Between 1862 (the first year for which data are available) and 1913, output increased by 3,800 percent in terms of monetary value. Because of continual price decreases, however, the increase in volume was substantially higher. The expansion in production from 1871 (the first year for which an estimate is available) to 1913 was 4,000 percent (from 3,500 to 162,000 tons). A larger market could, in principle, either accommodate more firms or, if no new firms should enter the industry, result in enormous growth for existing players, or anything in between.

This immense increase in output was made possible in part by the large proliferation of dyes sold in the market. By conservative estimates, on the eve of World War I, dyers and printers could choose from 900 distinct dyes made from about 270 chemical intermediates. From the very beginning of the synthetic dye industry, a frenzied search broke out to find synthetic routes to the commercially most significant natural dyes, alizarin and indigo. Perkin in Britain and the team of Graebe and Liebermann in Germany invented synthetic alizarin concurrently in 1869. Within a few years, farmers in Europe had to give up planting madder, the crop from which alizarin was made. But synthetic dyes not only competed with their natural counterparts, they also competed with one another: New synthetic dyes often made older ones obsolete because they were cheaper or better, as was the case with the first synthetic dye, aniline purple, which had a commercial life of only a few years. By the 1860s, many firms already made more than one particular dye. As the number of distinct synthetic dyes proliferated, some firms offered their customers a full spectrum of synthetic dyes. This product strategy transformed the economics of the dye business.

On the production side, firms not only could exploit some scale economies,[25] but even more importantly, they could organize to reap the benefit of scope economies. The demand for many dyes was not so large that a production line would be occupied all the time. To use production equipment continuously, firms like Bayer produced a wide range of dyes and intermediates in the 1890s, using the same production lines. To do so, firms needed to be able to make some predictions about demand for various dyes and then figure out an elaborate schedule as to when to make what dye or what intermediate and how to deal with unexpected downtimes in the production lines. The cost reduction made possible by using chemical intermediates for many different dyes and by cross-utilizing production lines depended on having relatively steady demand for a firm's product. Without sufficient orders, even the most efficient producer cannot stay in business. The economics of large scope and reasonable large-scale production required that competitive firms find a way to secure relatively steady demand for their product. We will explore later in more detail how a firm such as Bayer solved this problem by building a sophisticated marketing and distribution organization.

But where did all the novel dyes come from that would keep a large factory occupied? Dye innovations in the 1860s occurred in very different ways from those in the 1890s. Some chemical knowledge was always required to find synthetic dyes. Before Kekulé published his benzene ring theory in 1865, however, even leading organic chemists had little knowledge about how a particular molecule would create a particular color. Chemists would try out a large number of reagents on aniline without having a good idea of how the reagent transformed aniline to give it a coloring property. More trials and serendipity led to success. Because laboratory skills were involved in making dyes, most important dye innovations came from university-based researchers and were then commercialized by industrial firms.[26]

It took more than a decade until Kekulé's benzene ring theory and subsequent progress in organic chemistry had a significant impact on the development of new dyes (Ernst Homburg, personal communication, 2001). But by 1880 more precise understanding of the chemical structure underlying synthetic dyes provided scientists a powerful tool to search more systematically for dyes. When in 1877 German patent law protected dye innovations, a few German firms such as Hoechst, BASF, and AGFA saw the advantage of hiring organic chemists whose sole task would be to synthesize

[25] Compared to heavy chemical production, for example, of soda and sulfuric acid, scale economies were not nearly as important in dye making, but they clearly did exist, often giving larger firms an economic advantage over smaller ones. Scope economies, however, especially the ability to use the same production equipment for many dyes, were competitively much more significant in the synthetic dye industry.

[26] The empirical pattern that most major innovations came from university laboratories left traces in the socialization of chemistry students and their expectations. In 1883 Hoechst hired Dr. Eduard von Gerichten, a student of Emil Fischer, from the University of Erlangen. Bäumler (1988, p. 118) reports that "the young man had very ambivalent feelings toward his new job. On the one hand he would earn 300 Taler, and that meant much more than as an assistant at the University of Erlangen, his last job; but on the other hand, he had been taught that great inventions were to be made not by the chemical firms but by the universities and technical universities" [all translations of foreign language material are mine].

new dyes. After these research chemists turned out economically successful dyes, firms hired more and more chemists and pioneered an entirely new corporate function, formally organized research. The birth of corporate research and development (R&D), which today is a standard activity in high-tech industries (for overviews see Nelson, 1962; Freeman, 1982; Rosenberg, 1982), can be traced to the German synthetic dye firms in the 1880s. By the 1890s the vast majority of dyes were being discovered in the R&D laboratories of Bayer, Hoechst, and BASF.

Whereas in the early days of the industry a firm could exist by copying dyes invented somewhere else, patent laws made the systematic application of science within the boundaries of a firm a critical dimension of remaining a leader in the industry. The organization of innovation evolved from hiring one chemist to employing a cadre of chemists who would systematically search for new dyes to complement the existing product portfolio of the firm or would find a novel synthesis that could circumvent the patent protection a competitor had on a specific dye. From the mid-1880s onward, firms needed to master the forefront of synthetic organic chemistry to remain significant players.

The Road Ahead

To provide empirical support for a coevolutionary view of industrial leadership, I mainly operate at two distinct levels of analysis – the level of the national industry and the level of the individual firm. At the industry level the key questions are, Why was Germany able to surpass Britain and acquire an 85 percent world market share? Why did the U.S. dye industry fail to catch up with Germany? And, more generally, what was the mechanism that translated small initial performance differences into large differences over time? The record of the dye industry shows that the environment does not automatically bring about high-performance firms. Firms need to take advantage of favorable conditions.

To observe in greater detail how national institutions help or hurt the competitive position of domestic firms, I will examine the development of two companies in each of the three countries. At the level of the firm, the central questions are, How were individual firms able to take advantage of abundant resources in their immediate national environment or compensate for the lack thereof and become successful? Specifically, why were particular firms within the same country environment more successful than their domestic rivals? This study design makes possible another interesting question: Why was the German firm Bayer so much more successful than any British or U.S. competitor?

I begin the analysis at the level of national industries (Chapter 2), trying to answer the question of why the three national dye industries – of Britain, Germany, and the United States – followed very different development paths. The central puzzle is why Britain was unable to maintain its initial leadership position; for instance, the leading organic chemist Hofmann predicted in 1863 that Britain would dominate the industry for decades. I survey a host of performance factors – share of dye production, patent frequencies, number of firms participating in the industry, and so forth – to establish beyond doubt that Germany overtook Britain and came to dominate the synthetic dye industry in the decades before World War I.

Because industries are embedded in a larger social environment (Zukin and DiMaggio, 1990), understanding an industry's particular path requires knowledge of the specific opportunities and constraints would-be entrepreneurs faced in the three national contexts. In the next section, I paint a rough sketch of the differences in three countries before 1856 to give the reader sufficient knowledge about the histories of the three countries to fully understand the subsequent institutional analyses. I then analyze what features of the national environment caused the performance differences in the synthetic dye industry, beginning with the educational and training systems and then moving to the development of professional and trade associations, the role of the organic chemistry knowledge network, the social organization of production at the shop floor, and finally intellectual property right regimes.

I highlight the fact that the German educational and training system gave German firms a large advantage, particularly after science became a more precise tool in developing dyes. The formation of professional and trade organizations that supported the dye industry proceeded at different speeds in the three countries. In Britain, the infrastructure supporting the trade in natural dyes was much larger than in Germany, creating more obstacles and inertia in the process of developing supporting institutions specific to the synthetic dye industry. Earlier in this introduction, I touched on the critical role of synthetic organic chemical knowledge in creating successful enterprises in the synthetic dye industry. The educational system was a place where not only knowledge was created but also strong ties between professors and students. Many students came to work for industrial firms and thereby provided the link back to the university laboratories, where important new knowledge and discoveries were made.

Following Tilly's (1998) relational analysis of social processes, I dissect in some detail the academic–industrial knowledge network that developed alongside the industry. The analysis shows that the central players in this network were professors such as Hofmann and Baeyer.[27] Because the international network had its centers in Germany, German firms were, on average, closer in social (and geographic) distance to the sources of new knowledge and inventions than were British and U.S. firms.[28] Being located more on the periphery of the knowledge network, Britain and the United States possessed an inherent competitive disadvantage in recruiting talent and discovering new technological threats and opportunities.

Inventing new products, however, is not sufficient for commercial success. Firms also have to manufacture the product efficiently. German firms apparently were more frequently able to make the transition from having a foreman control the shop floor to letting chemical scientists and engineers make the calls about how to organize and manage production. Bringing scientific methods (i.e., rational and systematic analyses) to the shop floor allowed German firms to push the forefront of production efficiency. All firms who entered the dye industry in the very early days had to experiment to produce dyes on a large scale in a factory rather than on a small scale in a laboratory. Shop floor experimentation appears to have been more extensive in Germany because differences in patent regimes in 1857 across the three countries led to a larger number of plants there than in Britain or the United States. The patent laws and associated practices acted as important selection mechanisms on the populations of firms that

[27] Baeyer was not related to the founding family of the firm Bayer.

[28] For an application of network ideas and the concept of embeddedness to the study of economic action, see Uzzi (1997).

would come to exist in the three countries. Because Britain offered patent protection on dyes in 1857, firms with patents on particular dyes, such as Perkin & Sons, were shielded from fierce competition and realized large profits based on their monopoly. By contrast, the absence of effective patent protection in Germany until 1877 led to a very different selection regime. Firms could enter freely, and the forces of competition would eliminate firms that could not keep up with the efficiency gains of the best producers.

The passage of the all-German patent law in 1877 was fortuitous in its timing because it came after the industry had already developed strong firms and science was providing the tools to do systematic R&D on new dyes. German firms could pay for their R&D efforts after 1877 by getting fifteen-year monopolies on new dye-process inventions. The experience of the German and British dye industries shows that having a patent system is not necessarily as advantageous for the long-term development of an industry as is now often assumed. It also strongly suggests that the timing of the German patent law was crucial for its later beneficial effects. Had the German patent law arrived in 1858, it is doubtful that as many German firms would have developed into such strong competitors. Fewer firms would have entered the industry, and inefficient firms would have been more likely to survive, as was the case in Britain. The most important institution in the early success of the German dye industry was the university system, but patent laws were a second key factor that allowed the German firms to capture a dominant position.

By no means was every German firm a success. Quite to the contrary, 91 (or 185, if we count distinct legal entities as opposed to continuing economic units) of the 116 (200) entrants before 1914 went out of business. The figures for Britain are 36 (60) out of 47 (70) firms and for the United States 25 (27) out of 35 (40) firms. Chapter 3 examines in detail why some firms became more successful than others by tracing the development of six firms from their beginning until they went out of existence or until 1914, the end of the period studied in this book. A matched comparison of a winner and a loser in each of the three countries (Bayer and Jäger in Germany, Levinstein and Brooke, Simpson & Spiller [BS&S] in Britain, Schoellkopf and American Aniline in the United States) allows us to identify strategies that led to success.

The winners in all three countries shared one thing in common: In contrast to the losers, they had strong ties to the centers of organic chemistry knowledge. The German firm Bayer, which was closer in social space to the central nodes of the university knowledge network, was able to establish the strongest ties to leading researchers in organic chemistry and thereby assured itself more timely access to new chemical knowledge and talent that could run the firm's operations.

Confirming Chandler's views (1990), Bayer, unlike Jäger and the other four foreign firms, became a world leader because it invested its profits into building new organizational capabilities in marketing, production, R&D, and administration. Furthermore, Bayer, in contrast to the other five firms, also made a transition from a family firm to a joint stock company controlled by a team of professional managers, providing it with more growth potential than any owner-managed firm. Last but not least, in the first twenty years as synthetic dye firm, Bayer had an entrepreneur in its leadership ranks who took considerable risks. Bayer once came close to going out of business, but the gamble on a new dye factory (the alizarin plant) paid off handsomely and the firm returned to a profitable path. Jäger never took such risks and became relegated to the role of a niche player. The entrepreneurial spirit that pervaded Bayer, even when it was later run by professional managers, clearly mattered for becoming a world leader that could renew its product portfolio through systematic R&D.

Chapter 4 expands the Chandlerian view of the rise of the managerial firms to include political processes stressed in the writings of Fligstein (1990) and Roy (1997). The preceding chapters work with the assumption that national institutions have beneficial or harmful effects on local firms, but that firms themselves have no significant effect on the structure of these institutions. In Chapter 4, I abandon this working assumption and investigate the extent to which firms were collectively able to influence key institutional features of their national environment. This chapter represents the core of my empirical analysis because, although many scholars have long recognized the need to examine how firms shape their environment, few have done so systematically. The chapter also integrates Chandler's (1990) views with those of Fligstein (1990) and Roy (1997). We see that German firms were more successful than British and American firms in their lobbying efforts to upgrade educational institutions and to change patent laws and practices in their favor. I show, by conducting a careful study of the political dynamics that led to changes in the German patent laws and the failed attempts to change laws by British firms, that German firms' collective organization allowed them to shape German patent laws and practices to their competitive advantage.

The final chapter (Chapter 5) assesses the adequacy of existing theories in accounting for Germany's long dominance of the synthetic dye industry; in it I develop a coevolutionary model to explain how Germany moved from a laggard to an uncontested leader in this industry. I argue that although academic disciplines from economics to strategic management have provided a variety of theories to account for industrial success and failure, no one theory can adequately explain how and why Germany dominated the synthetic dye industry for so long and why firms within Germany differed so dramatically in their fortunes. Whereas international economics can explain quite well the success of the German industry as a whole, it cannot deal with the vast differences in performance of firms within the same national environment. In contrast, a theory such as the resource-based view of firms, developed by management scholars, has difficulty explaining why most of the successful players cluster in a particular national environment, rather than being spread out evenly across countries. Both kinds of theories, however, have difficulty explaining the dynamics of how competitive advantages change over time. The shortcomings in current theoretical arguments provide the starting point for the second part of Chapter 5, in which I develop an institutional theory of competitive advantage that deals with the national industry and firm level at the same time. To make coevolutionary arguments persuasive and not a catchall label for the analysis of organizational and environmental change, we need to know the specific mechanisms that characterize coevolution. After showing that the industry, technology, and institutions evolved, I articulate a more detailed coevolutionary model of industry development that I sketched earlier in this chapter. Besides highlighting the three abstract causal mechanisms of evolutionary explanations (variation, selection, and retention), I identify the exchange of personnel, the formation of commercial ties, and lobbying on behalf of the other social sphere as the more specific causal mechanisms that connected the evolution of national firm populations with the evolution of national populations of universities.

In this final chapter I also discuss the implications of the present study for evolutionary economics, corporate strategy, and business history. Because scholars from these different domains are all interested in the study of industrial change, the last section of the book proposes future research that should be carried out in the fertile triangle of evolutionary economics, management, and business history. I hope this overview has stirred your interest to read a detailed analysis of the evolutionary dynamics in the synthetic dye industry. Let us return to the beginning of the industry.