

Joseph T. Mahoney

Economic Foundations of Strategy



Foundations for
Organizational
Science
A Sage Publications Series



Economic Foundations of Strategy

To John Joseph

An excellent scholar.

Good year,

Joe Mahaney

11/15/2024

FOUNDATIONS FOR ORGANIZATIONAL SCIENCE

A Sage Publications Series

Series Editor

David Whetten, *Brigham Young University*
Editors

Anne S. Huff, *University of Colorado and Cranfield University* (UK)
Benjamin Schneider, *University of Maryland*
M. Susan Taylor, *University of Maryland*

The FOUNDATIONS FOR ORGANIZATIONAL SCIENCE series supports the development of students, faculty, and prospective organizational science professionals through the publication of texts authored by leading organizational scientists. Each volume provides a highly personal, hands-on introduction to a core topic or theory and challenges the reader to explore promising avenues for future theory development and empirical application.

Books in This Series

PUBLISHING IN THE ORGANIZATIONAL SCIENCES, 2nd Edition

Edited by L. L. Cummings and Peter J. Frost
SENSEMAKING IN ORGANIZATIONS

Karl E. Weick
INSTITUTIONS AND ORGANIZATIONS

W. Richard Scott

RHYTHMS OF ACADEMIC LIFE
Peter J. Frost and M. Susan Taylor

RESEARCHERS HOOKED ON TEACHING:

Noted Scholars Discuss the Synergies of Teaching and Research
Rae André and Peter J. Frost

THE PSYCHOLOGY OF DECISION MAKING: People in Organizations
Lee Roy Beach

ORGANIZATIONAL JUSTICE AND HUMAN RESOURCE MANAGEMENT
Robert Folger and Russell Cropanzano

RECRUITING EMPLOYEES: Individual and Organizational Perspectives
Alison E. Barber

ATTITUDES IN AND AROUND ORGANIZATIONS
Arthur P. Brief

IDENTITY IN ORGANIZATIONS: Building Theory Through Conversations
Edited by David Whetten and Paul Godfrey

PERSONNEL SELECTION: A Theoretical Approach
Neal Schmitt and David Chan

BUILDING STRATEGY FROM THE MIDDLE: Reconceptualizing Strategy Process
Steven W. Floyd and Bill Woolridge

MISSING ORGANIZATIONAL LINKAGES: Tools for Cross-Level Research
Paul S. Goodman

THE CONTINGENCY THEORY OF ORGANIZATIONS
Lex Donaldson

ORGANIZATIONAL STRESS: A Review and Critique of Theory, Research, and Applications
Cary L. Cooper, Philip J. Dewe, and Michael P. O'Driscoll

INSTITUTIONS AND ORGANIZATIONS, Second Edition
W. Richard Scott

ORGANIZATIONAL CULTURE: Mapping the Terrain
Joanne Martin

PERSONALITY IN WORK ORGANIZATIONS
Lawrence R. James and Michelle D. Mazerolle

CAREERS IN AND OUT OF ORGANIZATIONS
Douglas T. Hall

ORGANIZATION CHANGE: Theory and Practice
W. Warner Burke

COMPENSATION: Theory, Evidence, and Strategic Implication
Barry Gerhart and Sara L. Rynes

SERVICE QUALITY: Research Perspectives
Benjamin Schneider and Susan S. White

ECONOMIC FOUNDATIONS OF STRATEGY
Joseph T. Mahoney

Joseph T. Mahoney
University of Illinois at Urbana-Champaign

Economic Foundations of Strategy

Foundations for
Organizational
Science
A Sage Publications Series



 SAGE Publications
Thousand Oaks ■ London ■ New Delhi

Copyright © 2005 by Sage Publications, Inc.

All rights reserved. No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the publisher.

For information:



Sage Publications, Inc.
2455 Teller Road
Thousand Oaks, California 91320
E-mail: order@sagepub.com

Sage Publications Ltd.
1 Oliver's Yard
55 City Road
London EC1Y 1SP
United Kingdom

Sage Publications India Pvt. Ltd.
B-42, Panchsheel Enclave
Post Box 4109
New Delhi 110 017 India

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

Mahoney, Joseph T.

Economic foundations of strategy / by Joseph T. Mahoney.

p. cm.

Includes bibliographical references and index.

ISBN 978-1-4129-0542-8 (cloth : acid-free paper) — ISBN 978-1-4129-0543-5
(pbk. : acid-free paper)

1. Managerial economics. 2. Managerial economics—

Case studies. I. Title.

HD30.22.M335 2005

338.6'01—dc22

2004005692

This book is printed on acid-free paper.

12 13 14 15 16 17 10 9 8 7 6 5 4 3 2

Acquisitions Editor: Al Bruckner

Editorial Assistant: MaryAnn Vail

Production Editor: Julia Parnell

Copy Editor: Cheryl Duksta

Typesetter: C&M Digitals (P) Ltd.

Indexer: Kay M. Dusheck

Cover Designer: Edgar Abarca



Contents

Editor's Introduction	vii
Foreword	ix
Acknowledgments	xi
Overview	xiii
1. Behavioral Theory of the Firm	1
The Functions of the Executive (<i>Barnard, 1938</i>)	3
Administrative Behavior (<i>Simon, 1947</i>)	6
Organizations (<i>March & Simon, 1958</i>)	22
A Behavioral Theory of the Firm (<i>Cyert & March, 1963</i>)	33
Models of Bounded Rationality: Behavioral Economics and Business Organization (<i>Simon, 1982</i>)	40
2. Transaction Costs Theory	55
The Limits of Organization (<i>Arrow, 1974</i>)	57
The Firm, the Market, and the Law (<i>Coase, 1988</i>)	68
Markets and Hierarchies: Analysis and Antitrust Implications (<i>Williamson, 1975</i>)	72
The Economic Institutions of Capitalism (<i>Williamson, 1985</i>)	87
The Mechanisms of Governance (<i>Williamson, 1996</i>)	98
Cooperative Game and Mutual (Sunk-Cost) Commitment	105
3. Property Rights Theory	109
Contracting for Property Rights (<i>Libecap, 1989</i>)	111
Institutions, Institutional Change, and Economic Performance (<i>North, 1990</i>)	118
An Economic Analysis of Property Rights (<i>Barzel, 1989</i>)	123
Economic Behavior and Institutions (<i>Eggertsson, 1990</i>)	125
Firms, Contracts, and Financial Structure (<i>Hart, 1995</i>)	128

4.	Agency Theory	139
	The Modern Corporation (<i>Berle & Means, 1932</i>)	143
	“Principals and Agents: An Overview” (<i>Pratt & Zeckhauser, 1985</i>)	145
	“The Economics of Agency” (<i>Arrow, 1985</i>)	149
	“A Survey of Agency Models of Organizations” (<i>Levinthal, 1988</i>)	155
	“Theory of the Firm: Managerial Behavior, Agency Costs, and Capital Structure” (<i>Jensen & Meckling, 1976</i>)	162
5.	Resource-Based Theory, Dynamic Capabilities, and Real Options	167
	The Theory of the Growth of the Firm (<i>Penrose, 1959</i>)	169
	Scale and Scope: The Dynamics of Capitalism (<i>Chandler, 1990</i>)	178
	Mobilizing Invisible Assets (<i>Itami & Roehl, 1987</i>)	183
	An Evolutionary Theory of Economic Change (<i>Nelson & Winter, 1982</i>)	185
	Real Options: Managerial Flexibility and Strategy in Resource Allocation (<i>Trigeorgis, 1996</i>)	208
6.	The Theoretic Building Blocks of Organizational Economics	219
	References	225
	Index	239
	About the Author	253



Editor's Introduction

One sign of increasing maturity in the strategy field is an increasing attention to cumulative, rather than individual, theory development. Our research journals are requesting clearer theoretical foundations and demanding more significant theoretical contributions. We are systematizing and taking more seriously our sources in economics, sociology, psychology, and other fields. As a result, we are clarifying the nature of the academic discipline. We are better prepared to contribute to practice. And I believe we are in a position to begin giving something back to the base disciplines from which we draw.

The *Foundations of Strategy* series contributes to this important development by reviewing the foundations of the strategy field in economics, the behavioral sciences, and psychology. One objective is to specify the most important sources in each tradition for the use of students and for those who are more distant from the field but who have an interest in our subject matter. But a more pressing purpose is to provide the tools for active involvement in theory development. It is our hope, and belief, that the sophistication and standards of the best work from the past will inspire and support the best work for the future.

In this volume, Joe Mahoney identifies five theories of particular importance to the strategy field and outlines the contributions of particularly influential works in each area. As Oliver Williamson notes in his foreword, these authors provide alternative theories of the firm. To some extent they are complementary accounts, but they also propose alternatives with somewhat different purposes in mind. This book can be used to understand the key issues for strategy raised by theorists taking an economics perspective and, ideally, to further advance answers to five important questions:

1. How can organizations operate efficiently (the basic question of Chapter 1, “Behavioral Theory of the Firm”)?
2. How can firms minimize costs (the basic question of Chapter 2, “Transaction Costs Theory”)?

3. How can firms create and allocate wealth (the basic question of Chapter 3, “Property Rights Theory”)?
4. How can firms align individual self-interest (“the basic question of Chapter 4, “Agency Theory”)?
5. How can resources be acquired, developed, and deployed to improve the likelihood of survival and profitable growth (the basic question of Chapter 5, “Resource-Based Theory, Dynamic Capabilities, and Real Options”)?

These are big questions, still only partially answered, and the theory of the firm is of contemporary interest as an organized way to answer them. The questions asked also articulate some of the most basic concerns of managers. As Joe notes in his concluding chapter, the works cited here provide important building blocks for new theoretical developments in the field. Thus, our *Foundations* title applies.

Anne S. Huff



Foreword

An economics of organization has been taking shape over the past 30 years, the origins of which go back at least to the 1930s. As the name suggests, and as described by Joseph Mahoney, much of the economics of organization is interdisciplinary—drawing, as it does, on contributions from economics, organization theory (broadly conceived to include sociology, political science, and social psychology), cognitive psychology, and aspects of the law, especially property and contract law.

In large measure, the economics of organization arose in response to perceived limitations in the neoclassical theory of the firm, which described a firm as a production function that transformed inputs into outputs according to the laws of technology. Useful as this construction was (and is) for the study of prices and output (and, more generally, of resource allocation), this theory was less useful for understanding the modern corporation and for the making of public policy. The seeds of discontent were planted by Adolph Berle and Gardiner Means (1932), in relation to the separation of ownership from control; by Ronald Coase (1937), with reference to the puzzle of vertical integration and the theory of the firm; by Chester Barnard (1938), who featured adaptation of a cooperative kind accomplished through hierarchy (as opposed to adaptation through the market in response to changes in relative prices); by the marginalist controversy of the 1940s; and by Armen Alchian's (1950) appeal to evolutionary arguments to explain observed economic regularities.

The short of it is that too much was being asked of the neoclassical theory of the firm. As Harold Demsetz (1988) put it, it is a “mistake to confuse the firm of [orthodox] economic theory with its real world namesake. The chief mission of neoclassical economics is to understand how the price system coordinates the use of resources, not the inner workings of real firms” (1988, p. 189). Efforts to use the neoclassical theory of the firm in an all-purpose way nevertheless persisted. Strange and convoluted interpretations of nonstandard and unfamiliar

contracting practices and organizational structures often resulted, and interpretations were sometimes responsible for public policy error.

A new theory of the firm was evidently needed. But as Mahoney makes clear, what we have is not a new theory (singular) but new theories (plural). The five new theories of the firm (more generally, of economic organization, in that the firm is not a stand-alone entity but is to be examined in relation to the market and other modes of organizing economic activity) described by Mahoney are the behavioral theory of the firm, transaction costs theory, property rights theory, agency theory, and resource-based/dynamic capabilities. In many respects these are complementary in that they illuminate different issues. But sometimes they collide by providing rival explanations for the same phenomena. Ways have nevertheless been found to join parts of these, and more of this is in prospect. As matters stand presently, the study of economic organization is at a preunified state of development.

Readers of this book will want to ask themselves, as they work their way through successive chapters, what are the strengths and weaknesses of each proposed perspective? What phenomena does each illuminate? What predictions? What is the implied empirical research agenda? Does the theory scale up from the simple (often two-actor) model to complex (many actor) economic organization? What public policy ramifications accrue?

The new theories of the economics of organization that have taken shape over the past 30 years have vastly expanded the research agenda in both economic theory and organization theory as well as in the applied fields to which Mahoney makes reference: strategic management, agricultural economics, industrial organization, international business, and marketing. Indeed, applications to business and economic history, economic development, transition economics, positive political theory, and the law have also been made and are in progress. The economics of organization is an idea whose time has come. Lucky are the students of economic organization who seize upon the opportunity to run with this proliferation of good ideas.

*Oliver E. Williamson
January 2003*



Acknowledgments

I have many people to thank on my journey to writing this book. I begin with my mother, who advised me to “be kind.” My father advised me that “no matter what happens to you in life, no one can take your education from you.” My younger sister, Rose, taught me courage, and my younger brother, Jim, whom in many ways I looked up to during our childhood days, taught me caring and thoughtful reflection.

In terms of education, I first thank all of my grade school and high school teachers who showed dedication to their chosen profession. Studying undergraduate economics in the Faculty of Arts and Sciences at the University of Pennsylvania provided a wonderful foundation that has influenced my thinking, and I especially thank Professors Jacques Cremer, Robert Inman, Irving Kravis, Laurence Seidman, and Sidney Weintraub. In particular, Sidney Weintraub’s passion for ideas has had a lasting influence on my life.

In graduate school, in the study of business economics at the Wharton School of Business of the University of Pennsylvania, I thank Professors Robert Pollak and Harbir Singh for teaching me microeconomic theory and corporate strategy, respectively. I thank Professors Claudia Goldin, Bruce Kogut, and Gordon Walker for serving on my doctoral committee. My dissertation advisor, Professor Almarin Phillips, provided encouragement and guidance. Professor Ned Bowman not only served on my dissertation committee but also was a source of inspiration for me both personally and professionally. Support from the Reginald Jones Center, under Ned’s guidance, is gratefully acknowledged. I thank Professor Richard Marston, who gave me wise counsel to work for Ned Bowman at Wharton.

At the University of Illinois at Urbana-Champaign I thank my numerous colleagues over the past 15 years. I especially thank Irene Duhaime, Anne Huff, and Ravi Madhavan for their friendship and support. I also thank the excellent doctoral students I have worked with over the years, such as Bill Bogner, Jongwook Kim, Yasemin Kor,

Sung Min Kim, Chamu Sundaramurthy, and Danchi Tan, among others.

Finally, I owe so much to my wife, Jeanne Marie Connell, who has been there in good times and in bad. She has been a guiding force and a source of constant love and compassion. To her I dedicate this book. Thank you, Jeanne.



Overview

While writing this research book on the economic foundations of strategy, I kept two issues in mind. Suppose readers (at various levels of previous training) wanted to learn about organizational economics, especially from the perspective of strategic management: What significant themes best capture the directional tendencies of organizational economics today? How did current positions evolve? Suppose readers wanted to develop their own capacity to theorize and carry out research within the tradition of the economics of organization: What unresolved issues in this field are especially relevant to modern strategy research? What concepts are key building blocks? Where is additional empirical evidence needed?

Organizational economics has been informed by many great minds in social science research, including Nobel Prize recipients in economics, such as Kenneth Arrow, Ronald Coase, Douglass North, and Herbert Simon, and potential future recipients, such as Oliver Hart (for property rights theory) and Oliver Williamson (for transaction costs theory). In addition, research contributions by business school professors such as Alfred Chandler, James March, and Sidney Winter provide a wealth of insights based on business experiences. It is an important field taking a central place in the study of strategy, though of course it has also been useful for policy and other purposes, such as agricultural economics, international business studies, management information systems, marketing, organization theory, and so on.

In this book, organizational economics includes the following five interrelated theories: (1) a behavioral theory of the firm, (2) transaction costs theory, (3) property rights theory, (4) agency theory, and (5) (evolutionary) resource-based theory. In each area I have summarized my view of the critical observations of a few authors who have shaped the theory. Their work is presented in the present tense—even though a number of contributors unfortunately are no longer alive—because the work itself is very much alive. These books are worthy of

careful attention. Not only are these books widely cited, but they also exemplify the way theories are constructed by individual authors and the way the conversation among authors develops over time to create complex and compelling answers to important questions.

In the conclusion, I discuss complementarities among the theories outlined in this book, while recognizing their distinctive features. Most important, I outline some key questions for the ongoing economic study of organization that draws on these resources. I have written this book because the seminal works cited here are the foundations for so much current research. Key contributions of each author, in my view, are emphasized in *italics*. It may surprise some readers to find how current these observations sound, and some links to other research literatures are discussed here. But that discussion is limited to accommodate the wide variety of directions these foundational works can support. I hope that many different readers will be as inspired as I am by this classic literature. By understanding and emulating the best of our predecessors we can make more significant contributions today.

*Joseph T. Mahoney
November 2003*



1

Behavioral Theory of the Firm

The chapter begins with Barnard's (1938) *The Functions of the Executive* and is followed by four books from the Carnegie School: Simon's (1947) *Administrative Behavior*, March and Simon's (1958) *Organizations*, Cyert and March's (1963) *A Behavioral Theory of the Firm*, and Simon's (1982) *Models of Bounded Rationality: Behavioral Economics and Business Organization*. These books contain some of the best scholarly writings that the research literature has to offer on the behavioral theory of the firm. These research books are worth studying in detail because they continue to be widely cited today and because their clarity and relevance have not yet been surpassed.

The decision to classify the behavioral theory of the firm as part of an organizational economics approach to strategic management has its precedents, notably in the work of Barney and Ouchi (1986). Nonetheless, given that my book is part of a series, it might seem that this topic should be reserved for a research book on the behavioral foundations of strategy. I am sure it will reappear in that context. However, the behavioral theory of the firm also is part of organizational economics. Organizational economics is a multidisciplinary endeavor that draws on the broader field of economics and also gives attention to contributions from organization theory, law, and other areas. As an important example, Herbert Simon, whose 1947 and 1982 books are reviewed in this chapter, was awarded the Nobel Prize in the discipline of economics for work that included his contributions to the behavioral theory of the firm.

Furthermore, the behavioral theory of the firm serves as an important building block in transaction costs theory (Williamson, 1975). This theory is the subject of Chapter 2 and a central topic in organizational economics. Behavioral theory is also an important building block in dynamic capabilities theory and evolutionary economics (Nelson & Winter, 1982). This research is the subject of Chapter 5.

In terms of the five books chosen, Barnard (1938) combines the two cultures of science and art, and it is the aesthetic reading of Barnard that explains the intensity of students' responses to this work. Barnard offers an intense, structured, and coherent art form that depends on students' use of their capacities and their readiness to apprehend the aesthetic experience of management based on the author's intimate, habitual, interested experience (Mahoney, 2002).

Simon (1947) proposes a theory of human choice and decision making that aims to accommodate both those rational aspects of choice that have been the principal concern of economists and those properties and limitations of the human decision-making mechanisms that have attracted the attention of psychologists and practical decision makers. Simon focuses primarily on the decision-making processes that are internal to the organization and describes how organizations influence the decisions of their members, bring about consistency among those decisions, and guarantee that the decisions will be compatible with the overall organizational goals.

March and Simon (1958) persuasively argue that an adequate study of human behavior in organizations must take into account the motivational, attitudinal, and rational aspects of human behavior. Thus, both the works of economists on the planning process and the works of psychologists on organizational communication and problem-solving capabilities contribute to the evolving science of organization.

Cyert and March (1963) emphasize the actual process of making business decisions and provide detailed observations of the ways in which organizations make these decisions. Cyert and March develop an empirically relevant, process-oriented general theory of economic decision making by a business firm that, in my judgment, has stood the test of time. Cyert and March present the rudiments of a behavioral theory of the firm that have proven to be relevant both to economic theory and to the theory of complex organizations.

Simon's (1982) *Models of Bounded Rationality* takes up where *Administrative Behavior* (Simon, 1947) left off—attempting to understand decision making in its most general sense and, in particular, to show that economics and psychology could contribute to illuminating organizational decision-making processes. More specifically, Simon (1982) is concerned with explaining why there has been so little mutual influence of economics and psychology on each other, why a deeper dialogue needs to be developed between these two disciplines, and what the subject matter of their discourse could be.

In the process, Simon (1982) reveals a deep belief in and commitment to the interdependencies and complementarity of the several social sciences.

Simon borrows not only from economics but also from operations research, artificial intelligence, and cognitive psychology for the purpose of building a theory of procedural rationality (i.e., a theory of the processes of decision making) in complex, dynamic circumstances.

Though these arguments are a sufficient introduction to the chapter, I would make a final observation related to teaching. Those in strategic management who teach managers and managers-to-be will know that our students appreciate receiving not only theories for predicting but also theories that provide explanation. In other words, practitioners appreciate know-how but are deeply seeking advances in know-why. It has been my experience in teaching executives that a behavioral theory of the firm resonates with these managers and proves instructive for them. It makes beginning with the work of a practicing manager highly appropriate.

The Functions of the Executive (*Barnard, 1938*)

In my judgment, this book is the most high-powered intellectual contribution to organization or economic theory ever written by a practicing manager. Barnard's (1938) purpose is to provide a comprehensive theory of cooperative behavior in formal organizations.¹ Barnard observes that formal organization involves conscious, deliberate, and purposeful cooperation among people. One of the indispensable functions of an organization is to promote communication among these individuals. Another function is to maintain cohesiveness by regulating the willingness of various stakeholders to serve the organization and by maintaining the stability of authority. A third function is to maintain a feeling of personal integrity, self-respect, and independent choice.

But Barnard (1938) maintains that successful cooperation in or by formal organizations is the abnormal, not the normal, condition. We observe from day to day the successful survivors among innumerable organizational failures. Failure to cooperate, failure of cooperation, failure of organization, disorganization, dis-integration, destruction of organization—and reorganization—are the characteristic facts of human history.

The executive is critical. Executives inculcate belief in a common purpose. More concretely, executives synthesize the actions of contradictory

¹In addition to Barnard's (1938) classic, Barnard (1948) provides a collection of his selected papers. For modern assessments of Barnard (1938), see Mahoney (2002); Mahoney, Huff, and Huff (1994a, 1994b); Scott (1987); and Williamson (1995). Mahoney (2002) summarizes the major elements of Barnard's (1938, 1948) theory with special attention to Barnard's concepts of leadership and responsibility that are essential for distinctive competence (Selznick, 1957).

forces and reconcile conflicting instincts, interests, conditions, positions, and ideals.

Informal Organization. While Barnard (1938) defines the formal organization as a system of consciously coordinated activities or forces of two or more persons, this book also emphasizes the important role of informal organization within formal organizations. Crucially, Barnard regards informal organization as a means of maintaining the personality of the individual against certain effects of formal organizations that tend to disintegrate the personality. In fact, Barnard concludes that *expansion of cooperation and the development of the individual are mutually dependent realities and that a due proportion or balance between them is a necessary condition of human welfare.*

Incentives. Barnard (1938) observes that incentives are fundamental in formal organization. Inadequate incentives mean dissolution, unwarranted changes of organization purpose, or failure of cooperation. Hence, in all sorts of organizations, affording adequate incentives becomes essential. The specific means available include (a) material inducements, not just money but other things; (b) personal, non-material inducements, including distinction, prestige, and personal power; (c) desirable physical conditions; and (d) ideal benefactions, by which Barnard means the capacity of organizations to satisfy personal ideals.

The remarks about personal ideals and interests are very much in line with more recent discussions about identity and identification. Barnard (1938) is also contemporary in recognizing the incentives associated with (e) social attractiveness, or the social compatibilities people feel in their work environment; (f) conditions of habitual methods and attitudes; (g) the opportunity for enlarged participation; and (h) the condition of communion, or the feeling of solidarity or comradeship. None of this solidarity happens without effort. In addition to incentives, the book discusses persuasion and the inculcation of motives as important aspects of the organization.

Authority. Authority is the character of a communication (or order) in a formal organization by virtue of which a contributor accepts such an order. Barnard (1938) suggests that a person can and will accept a communication as authoritative only when four conditions simultaneously are met:

- The person can and does understand the communication.
- At the time of the person's decision, the person believes that the order is not inconsistent with the purpose of the organization.

- At the time of the person's decision, the person believes the order to be compatible with his or her personal interest as a whole.
- The person is able mentally and physically to comply with the order.

Perhaps the most well-known idea in the book is found in this discussion. Barnard (1938) argues that there exists a zone of indifference in each individual within which orders are acceptable without conscious questioning of their authority. Barnard further maintains that since the efficiency of organization is affected by the degree to which individuals assent to others, denying the authority of an organization communication is a threat to the interests of all individuals who derive a net advantage from their connection with the organization, unless the orders are unacceptable to them also. Thus, nothing is more real than authority.

An interesting corollary can be found in the assertion that the fine art of executive decision making includes not deciding questions that are not now pertinent, not deciding prematurely, not making a decision that cannot be made effective, and not making decisions that others should make. These are interesting, and rather unique, observations; Barnard (1938) argues the proper use of authority preserves morale, develops competence, and maintains authority. However, the natural reluctance of some people to decide, their persistent disposition to avoid responsibility, and their fear of criticism typically overwhelm executives. Executives thus must learn to protect themselves from the excessive burdens of decision making, if they are not already protected by a well-regulated and habitual distribution of responsibilities.

Another contemporary feature of this book is that the executive process Barnard (1938) describes transcends intellectual methods. *Feeling, judgment, sense, proportion, balance, appropriateness*, and other words are used to describe what executives should aspire to become. Leadership is more a matter of art than a matter of science. The processes used are more aesthetic than logical, derived chiefly from intimate, habitual, interested experience. For Barnard, coordination is a creative act.

Executive responsibility is also emphasized. Whatever morality exists in an individual becomes effective in his or her conduct, and the organization as a collective of cooperating individuals endures in proportion to the breadth of the morality by which it is governed. This assertion is only to say that foresight, long purposes, and high ideals are the basis for the persistence of cooperation (e.g., "old men and old women plant trees").

Although emphasizing instincts and morality, Barnard (1938) believes that a science of organization is also possible. Barnard recommends that treatises on management be written from various perspectives, including social anthropology, sociology, social psychology, and institutional economics but warns that we should not deceive ourselves by thinking that a science of cooperation and organization will alone promote greater integration of social forces. Inspiration is necessary to inculcate the sense of unity and to create economic ideals. Emotional rather than intellectual acceptance is required.

Barnard (1938) presents a systems view of the organization that contains a psychological theory of motivation and behavior, a sociological theory of cooperation and complex interdependencies, and an ideology based on a meritocracy. These insights greatly influenced Simon (1947), to whose early and influential book we now turn.

Administrative Behavior (*Simon, 1947*)

Indeed, Barnard wrote the foreword to Simon's (1947) *Administrative Behavior*. Barnard writes, “[Simon's book] has the right ‘feel.’ This means that I find Professor Simon’s apprehension of the structure of organized action consonant with my own experience. It therefore appeals to me as sound” (p. xiii). From Simon’s classic book concerning decision-making processes, readers should be able to discern principles of general organization that apply to administrative organization of great variety. Simon provides us with a self-conscious attempt to develop adequate linguistic and conceptual tools for realistically and significantly describing organizations. Simon’s primary thesis is that decision making is the heart of organization and that the vocabulary of organization theory must be derived from the logic and psychology of human choice.

Simon (1947) provides a brilliant synthesis of the practical teachings of Barnard (1938) and the evolving positive science of organization theory. As already noted, Simon’s *Administrative Behavior* is a landmark in organization theory as well as the economics of organization. Indeed, the organization theorist William Scott (1987, p. 45) classifies Simon within the paradigm of organizations as rational systems. From the perspective of the rational systems view, the behavior of organizations is considered as actions performed by purposeful and coordinated agents. In this sense, Simon is consistent with the logic of economics and uses the familiar language of *information, efficiency, implementation, and design*. Unlike neoclassical economics, however, Simon also insists

on coming to terms with cognitive limitations, which are discussed in terms of *constraints*, *authority*, *routines*, and *bounded rationality*. These terms imply that the rationality of organization behavior takes place within clearly specified limits. In short, this landmark book provides *an attention-based theory of the firm* of interest to both economic and organizational theorists.

Bounded Rationality. Simon (1947) observes that a person does not live for years in a particular position in an organization, exposed to some streams of communication, shielded from other streams of communication, without profound effects on what the person knows, believes, hopes, emphasizes, fears, and proposes. Researchers can understand neither the input nor the output of executives without understanding the organization in which executives work.

The term *organization*, for Simon (1947), refers to a complex pattern of human communications and relationships. This pattern of relationships provides each member of an organization or group within an organization much of the information and many of the assumptions, goals, and attitudes that enter into decisions. The pattern of relationships provides a set of stable and comprehensible expectations as to what the other members of the group are doing and how other members are likely to react to what is said and done. Every executive makes decisions and takes actions with one eye on the matter itself and one eye on the effects of this decision on the future pattern of relationships—that is to say, on its organizational consequences.

In summary, organizations are important because they provide much of the input that develops an executive's personal qualities and habits. Organizations also provide those in responsible positions the means for exercising authority and influence over others, a topic discussed in some detail in the following pages. Third, the organization influences the environments of information in which decisions are carried out.

When executives give attention to these indirect consequences, they concern themselves with organization. Sales managers react like sales managers because they occupy particular organizational positions, receive particular kinds of communications, are responsible for particular subgoals, and experience particular kinds of (economic) pressures. Executives can modify beliefs and attitudes by changing the flows of communications and thus modify decisions being made.

Decisions are also influenced by the authority relationship. On the one hand, classical organization theory emphasizes formal lines of authority in a hierarchical organization, implying (as Barnard [1938] observed) that legitimate commands are typically carried out. On the

other hand, the human relations school emphasizes the value of broad participation in decision making, demonstrates the importance of informal organization and the consequent limits on formal authority, and raises difficult questions about the human costs of excessively authoritarian environments.

These are not only different, more social, concepts of authority but also different concepts of rationality. Simon (1947) argues that the social sciences suffer from acute schizophrenia in their treatment of rationality. At one extreme, neoclassical economists attribute to "economic man" an omniscient rationality. Economic man has a complete and consistent system of preferences that allows him to choose correctly among the entire set of alternatives available to him. He is completely aware of what these alternatives are, there are no limits on the complexity of the computations he can perform to determine what alternatives are best, and he correctly makes all probability calculations. Tendencies uncovered by social psychology, traceable to Freud, which tend to reduce all cognition to affect, are at the other extreme. This alternative point of view notes that coins look larger to poor children than to rich children, observes that the pressures of a social group can persuade a person that he or she sees spots that are not there, shows that the process of group problem solving involves accumulating and discharging tensions, and so on.

Simon's (1947) major contribution to the economics of organization, as well as to organization theory, is the argument that it is precisely in the realm where human behavior is intendedly rational, but only limitedly so, that there is room for a genuine theory of organization. Simon maintains that *organizational behavior is the theory of intended and bounded rationality—it is about the behavior of humans who satisfice because they do not have the abilities to maximize*. Whereas *neoclassical economic man maximizes—selects the best alternative from among all those available to him—organizational man satisfices—looks for a course of action that is satisfactory or good enough. Economic man deals with the real world in all of its complexity, whereas organization man perceives a drastically simplified model of the real world*.

What is the significance of these two characteristics of satisficing and bounded rationality for organizational man? First, because he satisfices rather than maximizes, organizational man can make choices without first examining all possible behavior alternatives and without ascertaining that alternatives considered are in fact all those available. Second, because he ignores the interrelatedness of all things (so stupefying

to thought and action), organizational man can make decisions with relatively simple rules of thumb that do not make impossible demands on the capacity for thought. These critical theoretic observations have many interesting consequences.

Decision Making and Administrative Organization. Simon (1947) argues that it is *the process of choice that leads to action*. Although any practical activity involves both deciding and doing, it was not commonly recognized until this important book that a theory of organization should be concerned with the processes of decision as well as with the processes of action.

Simon (1947) notes that all behavior involves conscious or unconscious selection of particular actions out of all those actions that are physically possible to the actor and to those persons over whom the actor exercises influence and authority. The term *selection* is used without any implication of a conscious or deliberate process. Selection refers to the fact that if the individual follows one particular course of action, there are other courses of action that the individual thereby forgoes.

Simon (1947) employs a definition of authority substantially equivalent to that put forth by Barnard (1938). Subordinates accept authority whenever these subordinates permit their behaviors to be guided by the decision of a superior, without independently examining the merits of that decision. When exercising authority, the superior does not seek to convince the subordinates but only seeks to obtain their acquiescence. In actual managerial practice, of course, authority is usually mixed with suggestion and persuasion. If a superior attempts to carry authority beyond a certain point, which may be described as the subordinate's zone of acceptance, disobedience will typically follow.

The magnitude of the zone of acceptance depends on the various sanctions that authority has available to enforce its commands. The structure of formal authority in an organization typically is related to the appointment, disciplining, and dismissal of personnel. Informal authority relations in the tactical (day-to-day) work of the organization commonly supplement these formal lines of authority. The formal hierarchy is largely reserved for hearing and settling disputes.

Problems of Organizational Theory. Simon (1947) maintains that the authority relationship enables an organization to bring about specialization in the work of making decisions so that each decision is made at the point in the organization where the decision can be made more expertly for achieving purpose. *Purpose* is defined as the objective for which an activity is carried on, and *process* as a means of accomplishing purpose.

Simplistically, the concept of purpose involves a hierarchy of decisions—each step downward in the hierarchy consisting of an implementation of the goals set forth in the step immediately prior. Behavior is purposive in so far as it is guided by general goals; it is rational in so far as it selects alternatives that are conducive to the achievement of the previously selected goals. More realistically, the achievement of purpose often requires attention along multiple dimensions in the organization. Providing a useful analogy to make this particular point, Simon (1947) states that closet space is an important item in the design of a successful house, yet a house designed entirely with a view to securing a maximum of closet space—all other considerations being forgotten—would be considered somewhat unbalanced.

Similarly, unity of command, specialization by function, and decentralization are items to be considered in the design of an efficient organization. No single item is of sufficient importance to suffice as a guiding principle for the organizational analyst. In the design of organizations, as in their operation, overall efficiency is a guiding criterion. Mutually incompatible advantages are balanced against each other, just as an architect weighs the advantages of additional closet space against the advantages of a larger living room. A valid approach to the science of organization requires that the relevant diagnostic criteria be identified, that each organizational situation be analyzed in terms of the relevant set of criteria, and that research be instituted to determine how weights can be assigned to the several criteria when they are mutually incompatible.

Simon (1947) also observes that before a science of organization can develop theoretical principles, it must possess concepts. To be scientifically useful, the concepts chosen must be operational; that is, their conceptual meanings must correspond to empirically observable facts. As an analogy, before a law of gravitation could be formulated, it was necessary to have the concepts of acceleration and weight, and there had to be commonly accepted measures of these terms.

The theory of organization, in Simon's (1947) view, is concerned with how an enterprise should be constructed and operated to accomplish its work efficiently. A fundamental principle of organization, which follows almost immediately from the rational character of "good" organization, is that among several alternatives involving the same expenditure the one should be selected that leads to the greatest accomplishment of organizational objectives, and, among several alternatives that lead to the same accomplishment, the one should be selected that involves the

least expenditure. Since this principle of efficiency is characteristic of any activity that attempts rationally to maximize the attainment of certain ends with the use of scarce means, it is as characteristic of economic theory as it is of organizational theory. In this sense, the organization man takes his place alongside the neoclassical economic man. However, as already noted, one of Simon's (1947) most important contributions to science is the argument that individuals are limited by those skills, habits, and reflexes that are no longer in the realm of the conscious.

Limits to Rationality. At the most simplistic level, performance may be limited by manual dexterity or reaction time, and decision-making processes may be limited by the speed of mental processes. Individuals also are limited by their values and those conceptions of purpose that influence them in making their decisions, and these tend to be shaped by their organizational experience. If their loyalty to the organization is high, their decisions may evidence sincere acceptance of the objectives set for the organization; if loyalty is lacking, personal motives may interfere with organizational efficiency. Finally, individuals are limited by their knowledge of factors relevant to their job. This limitation applies both to basic knowledge required in decision making—bridge designers must know the fundamentals of mechanics—and to the information that is required to make decisions appropriate to a given situation.

In discussing means and ends, as well as facts and values, Simon (1947) is forthcoming concerning his own (logical positivist) philosophical perspective. Simon maintains that every decision involves elements of two kinds, which are called factual and value elements, respectively. This distinction is of primary importance for organization. Simon holds as a fundamental premise the idea that ethical terms are not completely reducible to factual terms. There is therefore no way in which the correctness of ethical propositions can be empirically tested. From this positivist perspective, if a sentence declares that some particular state of affairs "ought to be," or that it is "preferable" or "desirable," then the sentence performs an imperative function and is neither true nor false.

In contrast, a statement concerning the observable world is factual if, in principle, its truth or falsity may be tested. Simon (1947) hastens to add that in practice, the separation between ethical and factual elements in judgment can usually be carried only over a short distance. Further, the values involved in organizational decisions are seldom the final values in any psychological or philosophical sense.

Rationality in Organizational Behavior. Rationality in the world of experience is a complex concept. Simon (1947) provides a scenario of

two soldiers who sit in a trench opposite a machine gun. One soldier stays under cover. The other soldier, at the cost of his life, destroys the machine gun with a grenade. Which action is rational? Simon suggests that perhaps the only way to clarify these complexities is to use the term rational in conjunction with appropriate adverbs. Action may be called objectively rational if, in fact, it is the correct behavior for maximizing given values in a given situation. An action is subjectively rational if it maximizes attainment relative to the actual knowledge of the subject. An action is consciously rational to the degree that the adjustment of means to ends is a conscious process. An action is deliberately rational to the degree that the adjustment of means to ends has been deliberately brought about (by the individual or by the organization). An action is organizationally rational if it is oriented to the organization's goals; an action is personally rational if it is oriented to the individual's goals.

The Psychology of Organizational Decisions. From a rational point of view, choice is the process by which an alternative for each moment's behavior is selected. The task of choice involves three steps: (1) the listing of all alternatives, (2) the determination of all the consequences that follow each of these alternatives, and (3) the comparative evaluation of these sets of consequences. Each individual, to determine uniquely the consequences of actions, must know what will be the actions of others. This knowledge is of fundamental importance for the whole process of decision making.

Things are not so simple from an organizational point of view. Simon (1947) agrees with Barnard (1938) that organizations are systems of cooperative behavior. From the logical positivist perspective, rationality concerns the selection of preferred behavior alternatives in terms of some system of values whereby the consequences of behavior can be evaluated. But Simon argues that it is impossible for the behavior of a single, isolated individual to reach a high degree of rationality. The number of alternatives the individual must explore is so great, the information the individual would need to evaluate so vast, that even an approximation to objective rationality is hard to conceive. Individual choice takes place in an environment of givens—premises that are accepted by the individual as bases for choice—and behavior is adaptive only within the limits set by these givens.

Objective rationality would imply that the behaving individual molds behavior into an integrated pattern by (a) viewing the behavior alternatives prior to choice in panoramic fashion, (b) considering the whole complex of consequences that would follow from each choice,

and (c) with the system of values as criterion singling out one from the whole set of alternatives. Observed behavior, even that which is ordinarily considered rational, possesses many elements of disconnectedness not present in this idealized picture. However, one function the organization performs is to place members in a psychological environment that helps adapt their choices to organizational objectives. It also provides them with information needed to make these choices.

Even so, if individual or organization behavior is examined over time, it exhibits a mosaic character. Each piece of the pattern may tend to be integrated with others by their orientation to a common purpose, but such purpose shifts from time to time with shifts in knowledge and attention and is held together in only slight measure by any conception of an overall criterion of choice. Actual behavior falls short, in at least two ways, from objective rationality. First, rationality requires both complete knowledge and total anticipation of the consequences that will follow each choice. In fact, knowledge of consequences is always fragmentary. Second, rationality requires a choice among all possible alternative behaviors. In actual behavior, only a few of all these possible alternatives come to mind. Complete rationality is limited by lack of knowledge.

Humans striving for rationality but restricted within the limits of their knowledge develop working procedures that partially overcome these difficulties. These procedures assume they can isolate from the rest of the world a closed system containing only a limited number of variables and a limited range of consequences. Simon (1947) notes that the problem of discovering which factors are (and which factors are not) important in any given situation is as essential to choice as knowledge of the empirical laws governing those factors that are finally selected as relevant. *Rational choice is feasible to the extent that the limited set of factors on which a decision is based corresponds, in nature, to a closed system of variables, that is, to the extent that significant indirect effects are absent.* Only in the cases of some important decisions is it possible to bring to bear sufficient resources to unravel an involved chain of effects.

Humans do have some important abilities, including the capacity to observe regularities in nature and to communicate with others. Both help to shorten materially the learning process. The first capability means that previous experiences with other choices (of the same sort) may enable decision makers to infer something about the character of the particular choice that they face. Then, communication provides a

tremendous advantage in learning. For example, engineers designing a paved area do not have to base their attempts entirely on experimentation but can use reference sources that describe the conclusions of others. In effect, a relatively small amount of experience can serve as the basis for a wide range of decisions.

Memory. Memory as described in this book may be either natural or artificial—information may be stored in the mind, or it may be recorded in such a way as to be accessible. The artificial kind of memory that exists in libraries, files, and records is the most important in organizations. For either natural or artificial memory to be useful, there must be mechanisms that enable the memory to be drawn on when needed. Hence, human rationality relies heavily on the psychological and artificial associational and indexing devices that make the store of memory accessible when needed for the making of decisions.

An equally important mechanism that assists in the preservation of useful behavior patterns is habit. Habit, like memory, has an artificial organization counterpart that can be termed *organizational routine*. This idea has become important in the more recent literature, notably Nelson and Winter's (1982) work reviewed in Chapter 5.

Attention. Memory affects attention, where *attention* refers to the set of elements that enter into consciousness at any given time. To a considerable extent, the limits of rationality are the result of the limits of attention. It is important that both attention and behavior, once initiated in a particular direction, tend to persist in that direction for a considerable time interval. This persistence of attention holds even when the original choice of activity was a matter of relative indifference. Activity often results in psychological sunk costs that make persistence of attention in the same direction advantageous. A second reason for persistence is that the activity itself creates stimuli that focus attention toward its continuance and completion.

Simon (1947) maintains that the process of the integration of behavior involves three principal steps:

1. Individuals (or organizations) make broad decisions regarding the values to which they are going to direct their activities for substantive planning.
2. Individuals design and establish mechanisms that will direct their attention and channel information and knowledge in such a way as to cause the specific day-to-day decisions to conform to the substantive plan. This decisional activity is called *procedural planning*.
3. Individuals execute the plan.

But there are at least two intervening organizational and institutional influences on individual behavior: First, organizations and institutions permit (indeed encourage) stable expectations, and, second, organizations and institutions provide the general stimuli and attention directors that channel the behaviors of members of the group and provide those members with the intermediate goals that stimulate action.

Simon (1947) suggests the following mechanisms of organization influence:

- The organization divides work among its members. By giving each worker a particular task to accomplish, it directs and limits attention to that task.
- The organization establishes standard operating procedures.
- The organization transmits decisions by establishing systems of authority and influence.
- The organization provides (formal and informal) channels of communication running in all directions through which information flows.
- The organization trains and inculcates its organizational members.

The Equilibrium of the Organization. Simon (1947) maintains that individuals are willing to accept organization membership when their activity in the organization contributes, directly or indirectly, to their personal goals. The phrase *personal goals* should be understood in a broad sense. It is not restricted to egoistic goals, much less to economic goals. In a discussion reminiscent of Barnard (1938), the members of an organization contribute to the organization in return for inducements that the organization offers them. If the sum of *the contributions* is sufficient to supply the necessary inducements, the organization survives and grows; otherwise, the organization shrinks and ultimately disappears. In return for their inducements, members typically offer the organization not a specific service but their undifferentiated time and effort. Organizational members place this time and effort at the disposal of those directing the organization, to be used as those directing see fit. Thus, both the customer relation (in the commercial organization) and the employee relation originate in contract, but in contracts of different kinds. The employment contract results in the creation of an authority relationship between the organization and the employee.

How can this be? Why does the employee sign a blank check, so to speak, when entering employment? First, from the perspective of the organization, nothing would be gained by offering inducements to employees unless the employees' behaviors could be brought into

a system of organization behavior through their acceptance of its authority. Second, from the perspective of the employees, the precise activities with which their time of employment is occupied may, within certain limits, be a matter of relative indifference to them. In addition to the salary that employees receive, employees may value the status and prestige that their positions in the organization give them, and employees may value their relations with the working group of which they are part.

Organizational Goals. Three bricklayers were asked what they were doing. "Laying bricks," "Building a wall," "Helping to build a great cathedral" were their respective answers. This story conveys Simon's (1947) idea that in the world of experience the line of demarcation between personal and professional interests is not a sharp one because personal satisfactions may arise from the competent performance of a professional role and because both personal satisfactions and dissatisfactions may arise from innumerable conditions that surround the employment relationship.

Particular professional training may provide individuals with specific techniques and knowledge for solving problems (e.g., accounting techniques, legal techniques), which are then drawn on as part of the program evoked by their roles. In this way, a chief executive with an accounting background may find different problem solutions from a chief executive, in the same position, with a legal background. Individuals may incorporate in their role not only a professional style but also a personal style. Individuals may bring to the role, for example, habits and beliefs that provide them with crucial premises for their handling of interpersonal relationships.

An interesting question is why most commercial organizations tend to maintain fairly stable goals. Simon's (1947) answer is, first, that there are both economic and psychological sunk costs that make rapid adjustment unprofitable. Second, the organization requires know-how in a particular field—which becomes a sunk asset and part of the influencing organizational environment. Third, the organization acquires goodwill, which is also a sunk asset (i.e., the asset is not easily redeployed) and, thus, is not readily transferable to another area of activity.

The Role of Authority. Authority is defined as the power to make decisions that guide the actions of another. It is a relationship between two individuals, one superior and the other subordinate. The superior transmits decisions with the expectation that the subordinate will accept these decisions. The subordinate expects such decisions, and

these decisions influence the conduct of the subordinate. The relationship of authority is defined, therefore, in behavioral terms. Authority involves behaviors on the part of both superior and subordinate. When, and only when, these behaviors occur does an authority relationship exist between the superior and subordinate.

Individuals who do not have recognized status, or who are not recognized by their associates as expert with respect to a certain kind of knowledge, will have a more difficult time convincing their listeners that a recommendation is sound than those who possess expert credentials. Recommendations are judged partly on their merits and partly on the expertise of the persons making the recommendations. This pattern of judgments holds both because the individuals acting on the recommendations often do not have the expertise needed to judge them and because time pressures require these individuals to accept the recommendations of those whom they trust.

Furthermore, it is not implied that this resistance to irregular suggestions is entirely a weakness of organization. The specialization of decision-making functions and the fixing of responsibility for particular kinds of expertise on particular individuals are important sources of organizational efficiency that need to be balanced against the potential loss of independent ideas that results. When there is a disagreement between two persons, and when the disagreement is not resolved by discussion, persuasion, or other means of conviction, then the disagreement must be decided by the authority of one or the other participant. It is this right to the last word that is usually meant in speaking of lines of authority in an organization.

Simon (1947) proposes that the degree of obedience expected will vary with the social situation. The American workers of his day, for example, probably had a somewhat wider zone of acceptance, so far as the employer's instructions are concerned, than workers today. In part, this difference in the degree of authority may be due to the worker's weaker bargaining position back then or, conversely, the stronger sanctions of the employer, but there is probably a more fundamental change in social attitudes as to what is proper for an employer to ask an employee to do. These changed attitudes are reflected in social legislation limiting the terms of the employment contract. Professional workers and skilled workers are apt to have relatively narrow zones of acceptance, particularly in the areas of their own professional competencies.

The field of organizational behavior has stressed purpose as a sanction of primary importance. Subordinates are willing to obey commands

because subordinates realize that the coordination secured thereby is useful to the attainment of the joint purpose. Several conditions must be satisfied if purpose is to be an effective sanction of authority. Subordinates must have confidence that the command is issued in furtherance of a purpose with which they are in sympathy. Second, subordinates must have confidence that the command will be effective in achieving this purpose. This confidence may be based less on their knowledge of the correctness of the command than on their faith in the ability of those who issued the command; their recognition that those in authority have information they do not have; and their realization that their own efforts will not be effective in reaching the desired objective without some coordination from above.

Within limits, subordinates will even accept commands they know to be incorrect because they do not wish to challenge or unsettle a system of authority that they believe to be beneficial to their aims in the long run. There are, however, restrictions in the authority relationship. In a very real sense, the leader is merely a bus driver whose passengers will leave their leader unless their leader takes them in the direction they wish to go. Thus, subordinates give their leader only minor discretion as to the road to be followed.

Three functions of authority deserve special notice:

1. Authority enforces the responsibility of the individual.
2. Authority secures expertise in decision making.
3. Authority permits coordination of activities.

Simon (1947) notes that the core of many of the more important social institutions consists of a system of authority and a set of sanctions for enforcing the authority relationship. National government is the primary example, but the law of property, the church, and even families are included in this category (see Commons, 1934). Authority refers to the acceptance by subordinates of the decisions of the leader and not the power of the leader to apply sanctions in the case of noncompliance.

Communication. Both Barnard (1938) and Simon (1947) see communication as central to a theory of organization. Simon argues that without communication there can be no organization because there is no possibility then for the group to influence the behavior of the individual. Organization members sometimes use informal communication to advance their personal goals. From this informal behavior

the phenomenon of cliques arises—groups that build an informal network of communications and use this informal network as a means of securing power in the organization. Rivalry among cliques, in turn, may lead to general tensions in social relationships and defeat the purpose of the informal communications system.

Simon (1947) conjectures that weakness of the formal system of communications and failure to secure adequate coordination through that system probably encourage the development of cliques. A great deal of communication is categorized as gossip. In many organizations the grapevine probably plays, on the whole, a constructive role. Its chief disadvantages are, first, that it discourages frankness because confidential remarks may be spread about, and, second, that the information transmitted by the grapevine is often (deliberately or inadvertently) inaccurate. On the other hand, the grapevine is valuable as a barometer of public opinion in the organization.

It is also important that information does not automatically transmit itself from its point of origin to the rest of the organization; the individuals who first obtain the information must transmit this information. In transmitting the information, organizational members will naturally be aware of the consequences its transmission may have for them. When organizational members believe that the boss is going to be angered by the news, the news is likely to be suppressed. Hence, information tends to be transmitted upward in the organization only if its transmission will not have unpleasant consequences for the transmitters, the superior will hear of it anyway from other channels (and it is better to tell the superior first), or it is information that the superior needs in dealings with corporate leaders, and the superior will be displeased if caught without the information.

In addition, there is often failure to transmit information upward simply because subordinates cannot visualize accurately what information their superior needs. A major communication problem, then, of the higher levels of the organization hierarchy is that much of the information relevant to the decisions at this level originates at lower levels and may not reach the higher levels unless the executive is extraordinarily alert. Simon (1947) also states that there is a converse problem that arises when a superior withholds information from subordinates. This omission, again, may be accidental—the superior does not realize that subordinates need the information. On the other hand, superiors may use their exclusive possession of information as a means of maintaining authority over subordinates.

In an argument picked up by authors reviewed later in this book, notably Nelson and Winter (1982), Simon (1947) maintains that organizations, to a far greater extent than individuals, need artificial memories. Organizational routines that would become habitual in the case of the individual must be recorded in manuals for the instruction of new organization members. Among the repositories that organizations use are records systems, files, libraries, and follow-up systems. Simon also observes the importance of motivation: Every effective teacher recognizes that motivation is key to the learning process. Furthermore, personal motives may lead organization members to try to divert the communication system to their own uses and may lead organization members to withhold information from superiors and colleagues.

The Criterion of Efficiency. Simon (1947) notes that the simplicity of the efficiency criterion in commercial organizations is due, in large part, to the fact that money provides a common denominator for the measurement of both output and income and permits commercial organizations to be directly compared. Underlying all organizational decisions is a limitation—a scarcity—of available resources. This scarcity is the fundamental reason why time and money are costs. Because they are limited in quantity, their application to one organization purpose prevents the realization of alternative possibilities. The criterion of efficiency dictates the choice of alternatives that produce the largest economic result for the given application of resources. Simon argues that the concept of perfect efficiency is not required. Actual problems, as they present themselves to the decision maker, are always concerned with relative efficiencies, and no measure of absolute efficiency is needed. Furthermore, Simon does not assert that the criterion of efficiency dominates executives' decisions.

The Anatomy of Organization. If there were no limits to human rationality, organizational theory would be barren in Simon's (1947) view. Organization theory would consist of a single precept: Always select the alternative, among those available, that leads to the most complete achievement of desired goals. The need for an organizational theory resides in the fact that there are practical limits to human rationality and that these limits are not static but depend on the organizational environment in which the individuals' decisions take place. The task of organization is to design this environment so that individuals approach as close as practicable to rationality (judged in terms of the organization's goals) in their decisions.

In certain situations, it is possible to reorient individuals from identification with a subgoal of the organization to identification with a broader and more inclusive goal. When a particular item of knowledge is needed repeatedly in decision, the organization can anticipate this need and, by providing individuals with this knowledge prior to decision, can extend their area of rationality. This knowledge flow is the basic task of organization—to provide each operative employee with an environment of decision of such a kind that behavior that is rational from the standpoint of this environment is also rational from the standpoint of group values and the group situation. Simon (1947) concludes that the assumption so often made in organizational studies, that an arrangement is effective because it exists, is a circular argument of the worst sort. The only procedure of evaluation that can possibly be valid is the comparison of alternative organization arrangements in terms of their objective results. This procedure is a pragmatic test of what works in practice.

Information Processing. Information need not be processed just because the information is there. Nor should individuals believe that getting more information will always help solve their problems. In some cases, seeking more information indicates a touching faith in more water as an antidote to drowning. Simon (1947) counsels that (social) science does not advance by piling up information—science organizes information and compresses it. In scientific inquiry, *knowing* refers to knowing parsimoniously.

Any division of labor among decisional subsystems creates externalities, which arise out of the interdependencies among the subsystems that are ignored. What is required for the efficiency of the overall system is a factorization that minimizes these externalities and consequently permits a maximum degree of decentralization of final decisions to the subsystems and a maximum use of relatively simple and cheap coordinating devices, like the market mechanism, to relate each of the decisional subsystems with the other subsystems. Simon (1947) argues that the information-processing systems of modern civilization swim in an exceedingly rich soup of information. In a world of this kind, the scarce resource is not information; it is processing capacity to attend to information. *Attention is the chief bottleneck in organizational activity, and the bottleneck becomes narrower as we move to the tops of organizations, where parallel processing capacity becomes less easy to provide without damaging the coordinating function that is a prime responsibility of these levels.* Thus, the inherent capacity limits of

information-processing systems impose at least two requirements on organizational design: that the totality of decision problems be factored in such a way as to minimize the interdependence of the components and that the entire system be structured to conserve the scarce resource of attention.

Organizations (*March & Simon, 1958*)

This is a good point to turn to further theoretical developments found in March and Simon (1958)—a book that moves from a “closed rational system model” to an “open rational system model” (Scott, 1987, p. 100) of the organization. The organization is viewed as evolving toward both increased order and increased complexity. It is a work that provides new insight into the coping mechanisms of the organization.

March and Simon (1958) focus on the history of formal organizations. Taking the perspective of social psychologists, March and Simon are interested in influences that impinge on individuals from their environment and how individuals respond to such influences. March and Simon argue that roles in organizations tend to be highly elaborated, relatively stable, and defined to a considerable extent in explicit and even written terms. It is this predictability that enables organizations to deal in a coordinated way with their environments.

March and Simon (1958) take the viewpoint that a decision maker can be usefully regarded as an information-processor. March and Simon provide a picture of a choosing, decision-making, problem-solving individual who can do only one or a few things at a time and who can attend to only a small part of the information recorded in memory and presented by the environment.

March and Simon (1958) note that task allocations are efficient to the extent that such task allocations are based on similarities in activities that are recognized as yielding important complementarities in task performance. The key idea is to search for complementarities or, in modern terminology, economies of scope (Baumol, Panzar, & Willig, 1982; Teece, 1980). Beyond this point, solution of the task assignment problem requires empirical knowledge of the specific empirical complementarities that exist.

Behavior in the organization is not determined in advance and once and for all by a detailed blueprint and schedule. Even if it is highly routinized, the routine has the character of a dynamic capability rather than a fixed program. March and Simon (1958, p. 48) provide their own general model

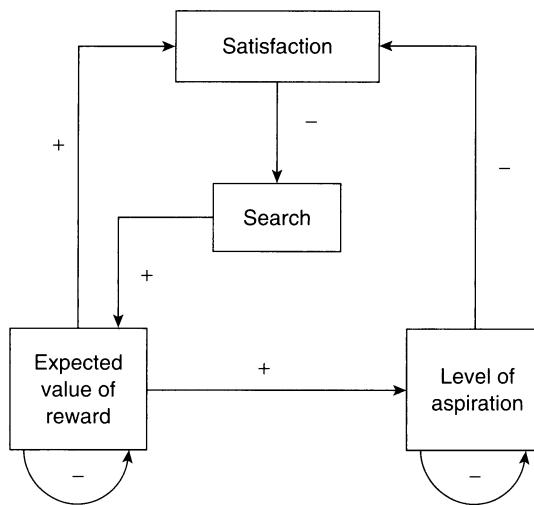


Figure 1.1 General Model of Adaptive-Motivated Behavior

SOURCE: March and Simon (1958, p. 49)

of intraorganizational decisions. The essential steps of March and Simon's (1958) behavioral model (see Figure 1.1) are as follows:

1. The lower the satisfaction of the individual, the more search for alternative programs the individual will undertake.
2. The more search, the higher the expected value of reward.
3. The higher the expected value of the reward, the higher the satisfaction.
4. The higher the expected value of the reward, the higher the level of aspiration of the individual.
5. The higher the level of aspiration, the lower the satisfaction.

Organizational Rewards. March and Simon (1958) note that many organizational models historically tended to relegate the reward schemes of management to the background. However, March and Simon insist that a model of a decision maker that does not give a prominent place to economic incentives is, for most humans, a poor predictive model. Further, March and Simon argue that an organization with a promotional scheme that essentially rewards seniority will be less productive than one that relates promotion to some index of productivity.

March and Simon (1958) emphasize that since employees are often cynical regarding announced performance criteria, the factors affecting the subjective operationality of performance standards are important. March and Simon also observe that, in general, the introduction of an incentive wage scheme results in increased production over a straight hourly or day rate, and the introduction of a flat-rate payment in place of a former incentive system depresses production. An employment contract based on a flat rate typically is regarded as controlling the type of activities performed but not the rate at which activities are performed.

March and Simon (1958) note that the greater the vertical mobility within an organization, the stronger the identification of individuals with the organization. Expectations of vertical mobility create expectations of interactions as well as felt similarities between subordinates and superiors.

Motivational Constraints: The Decision to Participate. The decision to participate is at the core of the theory of what Barnard (1938) and Simon (1947) call *organizational equilibrium*: the conditions of survival of an organization. Equilibrium reflects the organization's success in arranging compensations to its members that are adequate to motivate their continued participation. The Barnard-Simon theory of organizational equilibrium is essentially a theory of motivation—that is, a statement of the conditions under which an organization can induce its members to continue their participation and hence increase the likelihood of organizational survival. March and Simon (1958) describe the chief participants of most business organizations and, generally, focus on the following five major stakeholders: employees, investors, suppliers, distributors, and consumers. Most obvious in any catalogue of organizational participants are the employees, including the management.

In at least one respect, an employee's relationship to the organization is quite different from that of other stakeholders. In joining the organization employees accept an authority relationship. Employees agree that within certain limits (defined both explicitly and implicitly by the terms of employment contracts) they will accept as the premises of their behavior instructions supplied to them by the organization.

On the assumption that employees act in a subjectively rational manner, March and Simon (1958) predict the scope of the authority relationship from a knowledge of the inducements and contributions of the employees and other organization members. Employees are willing to enter into employment contracts only if it does not matter to them

very much which activities (within the zone of acceptance agreed to in the contracts) the organization will instruct them to perform or if employees are compensated in some way for the possibility that the organization will impose unpleasant activities on them. It is advantageous to subject employees to the organization's authority in those aspects that are of relatively great interest to the employer, comparatively unimportant to the employees, and about which the employer cannot make accurate predictions much in advance of performance.

The problems of both defining and enforcing the employment contract are matters of concern, and potential conflict, for all organizational participants. Whether dissatisfaction with the organization leads to withdrawal from the organization depends on whether the participants perceive the employment contract as given or as subject to change. Where the contract is viewed as unchangeable, the only options are accept or reject. Where the contract can be changed, participation by no means precludes internal conflicts and bargaining.

Conflict in Organizations by a Bargaining Outcome. March and Simon (1958) argue that game theory, in its original form, was no more satisfactory than neoclassical economic theory in providing an exact prediction of the outcome of a bargaining situation. What game theory offered was a specification of a set of feasible outcomes—the solution of the game. For example, in the case of highly specialized executives bargaining with their organization over salary, the salary paid will be somewhere between the economic value of the best alternative available to the executives elsewhere (i.e., what the executives can guarantee to themselves without cooperation) and the cost of the organization of hiring and training replacements (i.e., what the organization can guarantee to itself without cooperation). Since this feasible range may be quite wide, the theory is not overly helpful for providing reasonably precise economic predictions. March and Simon also provide the critical comment that, with rare exceptions, bargaining theory has operated in an empirical vacuum. The assumptions about human motivations and behaviors have usually been made on the basis of introspection, inspection of special cases, and mathematical tractability.

Cognitive Limits on Rationality. How does the rationality of organizational man compare with that of neoclassical economic man or with the rational man of modern statistical decision theory? The rational decision makers of economics and statistical decision theory make optimal choices in a highly specified and clearly defined environment:

1. When we first encounter them in a decision-making situation, rational decision makers already have laid out before them the whole set of alternatives from which they will choose their actions. This set of alternatives is simply given; the theory does not tell them how this set of alternatives is obtained.
2. To each alternative is attached a set of consequences—the events that will ensue if that particular alternative is chosen. Existing theories related to consequences fall into three categories:
 - a. *Certainty*: theories that assume the decision maker has complete and precise knowledge of the consequences that will follow on each alternative
 - b. *Risk*: theories that assume accurate knowledge of a probability distribution of the consequences of each alternative
 - c. *Uncertainty*: theories that assume that the consequences of each alternative belong to some subset of all possible consequences but that the decision maker cannot assign definite probabilities to the occurrence of particular consequences (see Knight, 1921)
3. At the outset, the decision maker has a utility function or a preference ordering that ranks all sets of consequences from the most preferred to the least preferred.
4. The decision maker selects the alternative leading to the preferred set of consequences. In the case of certainty, the choice is unambiguous. In the case of risk, rationality is usually defined as the choice of that alternative for which the expected utility is greatest. Expected utility is defined here as the average, weighted by the probabilities of occurrence, of the utilities attached to all possible consequences. In the case of uncertainty, the definition of rationality becomes problematic.

Some Difficulties in the Neoclassical Theory. There are difficulties with this neoclassical model of rational man. In the first place, only in the case of certainty does the neoclassical model agree well with commonsense notions of rationality. In the case of uncertainty, especially, there is little agreement, even among exponents of statistical decision theory, as to the correct definition of rationality or whether, indeed, the term *correct* has any meaning here.

A second difficulty with existing models of rational man is that these models make three exceedingly high demands on the choice-making mechanism. These models assume that (1) all the alternatives of choice are given, (2) all of the consequences attached to each alternative are known, and (3) the rational man has a complete utility ordering for all possible sets of consequences.

Routinized and Problem-Solving Responses. As a challenger to the neoclassical theory of rational choice, the theory of rational choice put forth by March and Simon (1958) incorporates two fundamental characteristics: (1) Choice is always exercised with respect to a limited,

approximate, simplified model of the real situation and (2) the elements of the definition of the situation are not given but are themselves the outcome of psychological and sociological processes, including the choosers' own activities and the activities of others in the choosers' environments.

Activity (individual or organizational) can usually be traced back to environmental stimuli of some sort (e.g., customer orders). The responses to stimuli are of various kinds. At one extreme, a stimulus evokes a response that was developed and learned at some previous time as an appropriate response for a stimulus of this classification. This response is the routinized end of the continuum, where a stimulus calls forth a standard operating procedure almost instantaneously. At the other extreme, a stimulus evokes problem-solving activities directed toward finding performance activities with which to complete the response. Problem-solving activities can generally be identified by the extent to which these activities involve search: search aimed at discovering alternatives of action or consequences of action. Discovering alternatives may involve inventing and elaborating whole performance programs where these programs are not already available in the repertory of the problem solver.

Search is partly random, but in effective problem solving search is not blind. The design of the search process is itself often an object of rational decision. Finding the optimal alternative is a different problem from finding a satisfactory alternative. An alternative is optimal if there exists a set of criteria that permits all alternatives to be compared and the alternative in question is preferred by these criteria to all other alternatives. An alternative is satisfactory if there exists a set of criteria that describes minimally satisfactory alternatives and the alternative in question meets, or exceeds, all these criteria.

Most human decision making, whether individual or organizational, is concerned with the discovery and selection of satisfactory alternatives; only in exceptional cases is it concerned with the discovery and selection of optimal alternatives. To optimize often requires processes several orders of magnitude more complex than processes required to satisfice. An example is the difference between searching a haystack to find the sharpest needle and searching the haystack to find a needle sharp enough to sew with adequately.

An optimizing rule would be to set the standard at the level where the marginal improvement in alternatives obtainable by raising the standard would be just balanced by the marginal cost of searching for

alternatives meeting the higher standard. Of course, in practice, the marginal improvement and the marginal cost are seldom measured in comparable units or with much accuracy. Thus, the optimizing rule is more a reconstructed logic used by a researcher to make predictions of behavior, rather than characterizing the logic-in-use (Kaplan, 1964) of the manager in the actual decision-making process.

Performance Programs. March and Simon (1958) argue that under certain circumstances the search and choice processes are abridged. At the limit, an environmental stimulus may evoke immediately from the organization a highly complex and organized set of responses. Such a set of responses is called a *performance program*. Situations in which a relatively simple stimulus sets off an elaborate program of activity without any apparent interval of search, problem solving, or choice are not rare. Knowledge of the performance program of an organization permits one to predict in considerable detail the behavior of members of the organization.

March and Simon (1958) first argue that organizations attempt to influence employees by specifying standard operating procedures and attaching organizational rewards and penalties to them. Second, performance programs are important parts of the coordination system in the organization. These performance programs help fulfill the needs for interdepartmental predictability. Insofar as performance programs are to function as controls, the programs must be linked to variables that are observable and measurable.

March and Simon (1958) expect performance program content to be a function of the *ease of observing job activities, the ease of observing job output, and the ease of relating activities to output*. Discretion available to the organizational participants is a function of their performance programs and in particular the extent to which the programs specify activities (means) and the extent to which these programs specify products or outcomes (ends).

March and Simon (1958) observe that in organizations there generally is a considerable degree of parallelism between the hierarchical relationships between members of the organization and the hierarchical relations between program elements. That is to say, the programs of members of higher levels of the organization have as their main output the modification or initiation of programs for individuals at lower levels. An important objective of standardization is to widen, as far as possible, the range of situations that can be handled by combination and recombination of a relatively small number of elementary programs.

March and Simon (1958) contend that rational behavior involves substituting for complex reality a model of reality that is sufficiently simple to be handled by problem-solving processes. In organizations where various aspects of the whole complex problem are being handled by different individuals and different groups of individuals, a fundamental technique for simplifying the problem is to factor the problem into a number of nearly independent parts so that each organizational unit handles one of these parts and can omit the others from its definition of the situation.

March and Simon (1958) note that the tendency of members of an organizational unit to evaluate action only in terms of subgoals—even when these subgoals are in conflict with the goals of the larger organization—is reinforced by at least three cognitive mechanisms. The first cognitive mechanism is located within the individual decision maker; the second mechanism, within the organizational unit; and the third mechanism, in the environment of the organizational unit.

In the individual, there is cognitive reinforcement through selective perception and rationalization. The propensity of individuals to see things that are consistent with their established frame of reference is well established in organizational psychology. Perceptions that are discordant with the frame of reference are filtered out before they reach consciousness or are reinterpreted or rationalized to remove the discrepancy. The frame of reference serves just as much to validate perceptions as the perceptions do to validate the frame of reference.

Within the organization unit, content of in-group communication provides cognitive reinforcement. Such communication affects the focus of information and thereby increases subgoal persistence. The vast bulk of our knowledge of fact is not gained through direct perception but through the secondhand, thirdhand, and *n*th-hand reports of the perceptions of others, transmitted through the channels of social communication. Two principal types of in-groups are of significance in filtering: in-groups with members in a particular organizational unit and in-groups with members in a common profession.

Finally, there is reinforcement through selective exposure to environmental stimuli. The division of labor in the organization affects the information that various members receive. This differentiation of information contributes to the differentiation of subgoals. For example, sales personnel live in an environment of customers, company treasurers live in an environment of bankers, and each sees a quite distinct part of the business world.

March and Simon (1958) observe that weatherpersons make observations of temperature, humidity, and barometric pressure but may communicate only their conclusions in the form of weather forecasts. In organizational communication, evidence is replaced with conclusions drawn from that evidence, and these conclusions then become the "facts" on which the rest of the organization acts.

When a means of testing actions is perceived to relate to a particular goal or criterion with possible courses of action, the criterion will be called *operational*. Otherwise, the criterion will be called *nonoperational*. For some purposes, we need the further distinction between cases where means-end relations can be evaluated prior to action and cases where means-end relations can be evaluated only after the fact. March and Simon (1958) call operational goals in the former case *operational ex ante* and in the latter case, *operational ex post*.

The goal of promoting the general welfare is frequently a part of the definition of the situation in governmental policy making. It is a non-operational goal because this goal does not provide (either *ex ante* or *ex post*) a measuring rod for comparing alternative policies. Strictly speaking, whether a goal is operational or nonoperational is not a yes-no question. There are all degrees of operability. It will often be convenient, however, to refer simply to the two ends of the continuum.

Important circumstances causing the substitution of subgoals for more general goals as the criteria for decision making occur when the subgoals are perceived as operational and the goals are perceived as nonoperational. For example, a business firm may understand to some degree how its specific actions affect its market share but may understand less surely how its actions affect long-term profitability. In such circumstances, the subgoal of maintaining a particular market share may become the effective criterion of action—the operational goal. When a number of individuals are participating in a decision-making process, and these individuals have the same operational goals, differences in opinion about the course of action will typically be resolved by predominantly analytic processes (i.e., by the analysis of the expected consequences of courses of action for realization of the shared goals). When either of the postulated conditions is absent from the situation (when the goals are not shared or when the shared goals are not operational and the operational subgoals are not shared), the decision will typically be reached predominately by bargaining processes.

Interdependence does not by itself cause difficulty if the pattern of interdependence is stable and fixed because, in this case, each subprogram

can be designed to take account of all the other subprograms with which it interacts. Difficulties arise only if program execution rests on contingencies that cannot be predicted in advance. In this case, coordinating activity is required to secure agreement about the estimates that will be used as the basis for action or to provide information to each subprogram unit about the relevant activities of the others. Hence, March and Simon (1958) arrive at the research proposition that the more repetitive and predictable the situation, the greater the *tolerance for interdependence*.

Communication and Coordination. An important method for increasing the organization's tolerance for interdependence is to increase the efficiency of communication by making it possible to communicate large amounts of information with relatively few symbols. An obvious example is the blueprint, which provides a common plan stated in detail. A blueprint uses a carefully defined, highly developed language or set of symbolic and verbal conventions. Because of this standardized language, a blueprint can convey large quantities of information. The same attention to standardization of language is seen in accounting systems and other reporting systems that use numerical data. Accounting definitions and blueprint conventions are examples of a still more general phenomenon: technical languages, whose symbols have definite and common meanings to the members of an organization.

March and Simon (1958) observe that the world tends to be perceived by the organization members in terms of the particular concepts that are reflected in the organization's vocabulary. The particular categories and schemes of classification that the organization employs are reified and become, for members of the organization, attributes of the world rather than mere conventions.

Organization Structure and the Boundaries of Rationality. March and Simon (1958) maintain that because of the limits of human intellectual capacities in comparison with the complexities of the problems that individuals and organizations typically face, rational behavior calls for simplified models that capture the main features of a problem without capturing all of a problem's complexities. The simplifications have a number of characteristic features:

- Optimizing is replaced by satisficing.
- Alternatives of action and consequences of action are discovered sequentially through search processes.
- Repertoires of action programs are developed by organizations and individuals, and these repertoires serve as the alternatives of choice in recurrent situations.

- Each specific action program deals with a restricted range of situations and a restricted range of consequences.
- Each action program is capable of being implemented in semi-independence of the others—these action programs are only loosely coupled together.

This one-thing-at-a-time, or *ceteris paribus*, approach to adaptive behavior is fundamental to organization structure. Organization structure consists of those aspects of the pattern of behavior in the organization that are relatively stable and that change only slowly. If behavior in organizations is intendedly rational, then March and Simon (1958) expect aspects of relatively stable behavior that either represent adaptations to relatively stable elements in the environment or are the learning programs that govern the process of adaptation.

March and Simon (1958) maintain that a great deal of the inertia of going concerns can be explained on the basis of (economic and psychological) sunk costs. A simple example is whether to move to a new location with high moving costs. In addition to tangible sunk costs, persistence comes about primarily because the individual or organization does not search for, or consider, alternatives to the present course of action unless that present course is in some sense unsatisfactory. March and Simon (1958) also suggest a Gresham's law of planning: Daily routine drives out planning. Stated less cryptically, March and Simon predict that when an individual is faced both with highly programmed and highly unprogrammed tasks, the highly programmed tasks tend to take precedence over the highly unprogrammed tasks, even in the absence of strong, overall time pressure. Although left unstated by March and Simon, we may anticipate that problems of persistence can be greater for long-term strategy formulation and implementation.

Concluding Comments. Scott (1987) notes that there are important differences between Simon (1947) and March and Simon (1958). Although there is still a concern with the cognitive limits of individual decision makers and with how organizational structures can help to support improved decision making, March and Simon (1958) place a greater emphasis on the variable nature of challenges posed by tasks and environments. The organization is viewed as more open to its environment. Whereas some performance programs can be routinized, other performance programs must be problem-solving responses, requiring the decision maker to exercise more discretion in the face of greater uncertainty. Moreover, it is recognized that some organizations face such volatile environments that these organizations institutionalize

innovation, devising programs for routinely changing existing programs, often rapidly.

In conclusion, my understanding of the evolution of management theory in the 1947–1958 period is that in comparison with Simon (1947), March and Simon (1958) provide a stronger sense that organizations face environments of varying complexity. Furthermore, organizations must adjust their internal decision-making procedures to take these variations into account, and some environments impose levels of complexity that organizations cannot manage unless these organizations impose simplifying restrictions on the information processed.

A Behavioral Theory of the Firm (*Cyert & March, 1963*)

We turn from March and Simon (1958) to the next landmark in the behavioral theory of the firm by Cyert and March (1963). Cyert and March focus on a small number of key economic decisions made by the firm and develop process-oriented models of the firm.

Cyert and March (1963) are concerned with the business firm and the way the business firm makes economic decisions, and the authors make detailed observations of the processes and procedures by which firms make decisions, using these observations as a basis for a theory of decision making in business organizations. Cyert and March (1963) argue that one way to understand modern organizational decision making is to supplement the microeconomic study of strategic factor markets with an examination of the internal operation of the business firm—to study the effects of organizational structure and conventional practices on the development of goals, the formation of expectations, and the implementation of choices.

Cyert and March (1963) make four major research commitments:

- To focus on the small number of key economic decisions made by the firm
- To develop process-oriented models of the firm
- To link models of the firm as closely as possible to empirical observations
- To develop a theory with generality beyond the specific firms studied

Cyert and March's (1963) conception of the theory-building task is that of constructing a theory that takes the firm as its basic unit; considers the prediction of firm behavior with respect to such decisions as price, output, and resource allocation as its goal; and *emphasizes the*

actual process of organizational decision making. To build the behavioral theory of the firm, Cyert and March develop four major subtheories concerning the following:

- Organizational goals
- Organizational expectations
- Organizational choice
- Organizational control

Organizational Goals. A theory of organizational goals considers how goals arise in an organization, how goals change over time, and how the organization attends to these goals. Cyert and March (1963) view an organization as a coalition of stakeholders, with some of these stakeholders organized into subcoalitions. In a business organization the coalition members include managers, workers, stockholders, suppliers, customers, lawyers, tax collectors, regulatory agencies, and so on. Clearly then, organizational goals must deal successfully with the potential for internal goal conflicts inherent in a coalition of diverse individuals and groups.

Since the existence of unresolved conflicts among organizational stakeholders is a key feature of organizations, it is difficult to construct a useful descriptively accurate theory of the organizational decision-making process if we insist on internal goal consistency. Cyert and March (1963) do not insist then that such a decision-making process necessarily produces consistent organizational goals.

Because individuals have limited capacities, and limited time, to devote to any particular aspect of the organizational system, such limitations constrain the bargaining process. As an adaptive response, coalition members are motivated to develop mutual control systems, such as the budget and the allocation of tasks by the division of labor and specialization. A budget becomes a precedent for future budgets; an allocation of tasks becomes a precedent for future task allocations. Thus, coalition agreements are institutionalized into semipermanent arrangements.

Another important mechanism for dealing with stakeholder conflicts is the sequential attention to conflicting goals. A consequence of this mechanism is that organizations ignore many conditions that outside observers see as direct contradictions. The decentralization of decision making (and goal attention), the sequential attention to goals, and the adjustment in organizational slack that acts as a cushion in down times permit the business firm to make decisions with inconsistent goals under many (and perhaps most) conditions.

Organizational Expectations. A theory of organizational expectations considers how and when an organization searches for information or new alternatives and how information is processed through the organization. Expectations are by no means independent of hopes, wishes, and the internal bargaining needs of subunits in the organization. Information about the consequences of specific courses of action in a business organization is frequently hard to obtain and of uncertain reliability. As a result, both conscious and unconscious biases in expectations are introduced. Thus, local priorities and perceptions obtain. In addition, there is some evidence of more conscious manipulation of expectations. Communication in a complex organization includes considerable biasing and influence activities—and considerable bias correction as well. In addition, organizations often protect themselves from the worst effects of influence activities by focusing on verified data in lieu of uncertain estimates and by using easily checked feedback information.

Organizational Choice. A theory of organizational choice needs to characterize the process by which the alternatives available to the organization are ordered and selected. Organizational decisions depend on information estimates and expectations that ordinarily differ appreciably from reality. These organizational perceptions are influenced by some characteristics of the organization and its procedures. Second, organizations consider only a limited number of decision alternatives. Finally, organizations vary with respect to the amount of resources that such organizations devote to their organizational goals on the one hand and suborganizational and individual goals on the other hand. The firm is considered to be an adaptively rational system in which the firm learns from experience. General choice procedures are summarized in terms of three basic principles:

- *Avoid uncertainty.* The firm looks for procedures that minimize the need for predicting uncertain future events. One method uses short-run feedback as a trigger to achieve action; another accepts (and enforces) standardized decision rules.
- *Maintain the rules.* Once the firm has determined a feasible set of decision procedures, the organization abandons them only under duress.
- *Simplify the rules.* The firm relies on individual judgment to provide flexibility around simple rules.

Organizational Control. A theory of organizational control specifies the difference between executive choice in an organization and the decisions actually implemented. Organizational control within an organization depends on the elaboration of standard operating procedures. It is hard to see how a theory of the firm can ignore the effect of such organizational

procedures on decision-making behavior within the organization. The effects fall into at least four major categories: (a) effects on individual goals within the organization, (b) effects on individual perceptions of the environment, (c) effects on the range of alternatives considered, and (d) effects on the managerial decision rules used. Cyert and March's (1963) basic theory of organizational control assumes the following:

- *Multiple, changing, acceptable-level goals.* The criterion of choice is that the alternative selected meets the demands (goals) of the coalition.
- *An approximate sequential consideration of alternatives.* The first satisfactory alternative evoked is accepted. When failure occurs, search is intensified.
- *Uncertainty avoidance.* The organization seeks to avoid uncertainty by following standard operating procedures and a policy of reacting to feedback rather than forecasting the environment.

Summary of Cyert and March (1963). Cyert and March propose two major organizing devices: a set of variable concepts and a set of relational concepts. The variable concepts discussed previously are organizational goals, organizational expectations, organizational choice, and organizational control. There are also four major relational concepts (see Figure 1.2):

- Quasi-resolution of conflict
- Uncertainty avoidance
- Problemistic search
- Organizational learning

Quasi-Resolution of Conflict. In keeping with numerous theories of organizations, Cyert and March (1963) assume that the coalition in an organization is a coalition of members having different personal goals. Members require some procedure for resolving conflicts, such as acceptable-level decision rules, sequential attention to goals, or both.

Uncertainty Avoidance. Cyert and March (1963) submit that organizations typically try to avoid uncertainty. First, organizations avoid the requirement that they correctly anticipate events in the distant future by using decision rules emphasizing short-run reactions to short-run feedback, rather than anticipation of long-run uncertain events. Second, organizations avoid the requirement that they anticipate future reactions of other parts of their environment by arranging a negotiated environment. Organizations impose plans, standard operating procedures, industry tradition, and uncertainty-absorbing contracts on that environment.

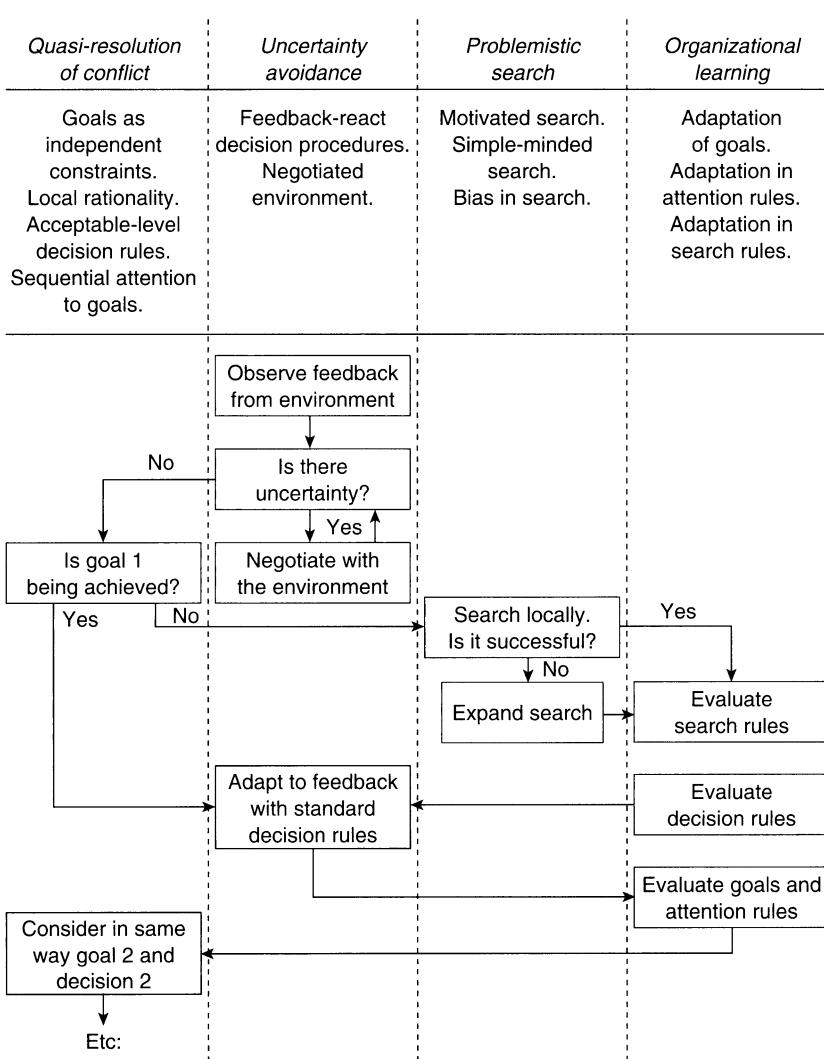


Figure 1.2 Organizational Decision Process in Abstract Form

SOURCE: Cyert and March (1963, p. 126)

Problemistic Search. Cyert and March's (1963) behavioral models assume that search, like decision making, is problem directed. *Problemistic search* means search that is stimulated by a problem (usually a rather specific one) and is directed toward finding a solution to that problem. Such organizational search is assumed to be motivated, simple-minded,

and biased. This bias may reflect training or experience of various parts of the organization. This bias may reflect the interaction of hopes and expectations, and communication biases are expected to reflect unresolved conflicts within the organization.

Organizational Learning. To assume that organizations go through exactly the same processes as individuals go through seems unnecessarily naive, but organizations exhibit (as do other social institutions) adaptive behavior over time. Cyert and March (1963) focus on adaptation with respect to three different phases of the decision process: adaptation of goals, adaptation in attention rules, and adaptation in search rules. Cyert and March submit that organizations change their goals, shift their attention, and revise their procedures for search as a function of their experience.

Scott (1987) notes that Cyert and March's (1963) concept of *coalitions* offers the following features:

- The problem of reification is avoided; individuals and groups have interests, and the processes by which these preferences come to be imposed on the organization are specified.
- It is recognized that although individuals and groups are allowed to specify the goals of the organization, there is no presumption that they do so on an equal footing, nor is it assumed that they hold common objectives.
- It is recognized that although individuals and groups impose goals on the organization, in most cases no single individual or group is powerful enough to determine completely the organization's goals; hence, the organization's goals are typically distinct from those of any of its participants.
- Allowance is made for differences in interests among participants. Some, but not all, of these differences may be resolved by negotiation, so at any time conflicting goals may be present.
- It is recognized that the size and composition of the dominant coalition may vary from one organization to another and within the organization from time to time.

In my judgment, Cyert and March's (1963) work provides a more conceptually refined and systematic outline of the behavioral theory of the firm that improves on March and Simon (1958). Three statements can summarize Cyert and March's arguments:

- The business firm is a relevant unit of investigation.
- It is possible to construct a theory of decision-making behavior within such a unit.
- Such a theory must focus explicitly on actual organizational decision processes.

To this purpose, Cyert and March (1963) show how to construct behavioral models of firm-level decision making and indicate the basic

theoretical framework within which such models are embedded. Cyert and March's (1963) behavioral theory of the firm can be applied to price and output decisions, internal resource allocations, innovations, competitive dynamics, and predictions of other organizations' behavior.

I hold the hope that current students studying the economics of organization may build on Cyert and March (1963) and connect this research agenda with other branches of organizational economics, which we consider in subsequent chapters. It is my view that building a science of organization that suppresses issues of bounded rationality and limited information processing by organizational members would lead the strategic management field up a blind alley.

Now that we have examined the classic work of Barnard (1938) and of the Carnegie School of Simon (1947), March and Simon (1958), and Cyert and March (1963), we apply the Carnegie framework to the Cuban Missile Crisis, where for 13 days the United States and the Soviet Union paused at the nuclear precipice (Allison, 1971).

Application: Explaining the Cuban Missile Crisis

The Cuban missile crisis was a seminal event in the history of the United States. On the days between October 16 and October 28, 1963, the United States and the Soviet Union came dangerously close to nuclear war. Using the Carnegie School framework for explaining the crisis, we consider the following question: Why did the Soviet Union place strategic offensive missiles in Cuba?

Why did the Soviet Union place strategic offensive missiles in Cuba? From the Carnegie School framework, explanation for this action requires an identification of the relevant Soviet organizations and displays the patterns of organizational behavior from which the actions emerged. An explanation, from this perspective, must identify trends that reflect established organizations and their somewhat rigid operating procedures and programs. That is, governmental behavior is understood less as deliberate choices (as if from a unitary actor) and more as outputs of large organizations functioning according to standard patterns of behavior.

While the final decision to put missiles in Cuba must have been made in the presidium, the details of this operation—that is, the

(Continued)

(Continued)

path from the general decision to the actual appearance of operation missiles in Cuba—were probably delegated to appropriate Soviet organizations, such as the GRU (Soviet military intelligence), the KGB (the Communist party security agency), the SAM (the Soviet Air Defense Command), and a quite separate Soviet military service, the Strategic Rocket Forces. Standard Soviet operations, particularly when nuclear weapons were involved, imposed a very high level of secrecy. Thus, each organization's tendency to follow standard operating procedure was reinforced by a lack of information about the activity of other organizations and the impossibility of an overview of the whole operation. Allison (1971) discusses several instances where contradictory behaviors and anomalies (from the perspective of a unitary actor model) are explained from the Carnegie School framework. Many crucial details of implementation followed from organizational routines rather than from central choice.

The lesson, as Allison (1971) suggests, is that nuclear crises between machines as large as the United States and the Soviet Union have elements of genuine uncertainty. The information and estimates available to leaders about the situation will reflect organizational goals and routines as well as the facts. The alternatives presented to leaders will be much narrower than the menu of options that would typically be more desirable. The implementation of choices will exhibit unavoidable rigidities of organization's standard operating procedures. In a crisis, the overwhelming problem will be that of control and coordination of large organizations.

SOURCE: Adapted from Allison (1971)

Models of Bounded Rationality: Behavioral Economics and Business Organization (*Simon, 1982*)

Finally, I conclude this first chapter with a summary of Simon's (1982) work in the research area of behavioral economics, which is worthy of careful attention by students studying the economics of organization. Simon, leading by example, shows how fruitful social science research can be for those who are not intimidated by disciplinary boundaries and that anything that can improve our understanding of complex organizations should be valued.

Simon (1982) argues that organization theory, economics (especially the theory of the firm), and cognitive psychology are all basically concerned with the same phenomena. All three are theories of human decision-making and problem-solving processes, yet each of the three domains has developed in relative isolation from the other two domains. Simon is concerned with both the causes for this isolation and its remedies.

The Business Firm as an Organization. Simon (1982) notes that the firm of neoclassical economic theory is little more than an entrepreneur who has attached a cost curve or a production function. Since profit maximization and internal efficiency are assumed, there is little room in the neoclassical theory for the familiar institutional characteristics of real business firms—for example, that one of a business firm's principal inputs typically is labor, a commodity that is contracted for on quite a different basis from other commodities, and that decisions are reached within a hierarchy of authority relationships among the employees.

A Formal Theory of the Employment Relationship. Simon (1982) observes that neoclassical economic theory abstracts away the distinctive characteristics of the employment contract, and neoclassical economic theory ignores the most significant features of the organizational process (i.e., the process of actually managing the factors of production, including the input of labor). Simon sets forth a theory of the employment relationship that reintroduces some of the more important of these empirical realities into the economic model. Perhaps in this way a bridge can be constructed in the discipline of strategic management between economists, with their theories of the firm and of factor allocations, and organization theorists, with their theories of organization—a bridge wide enough to permit some free trade of ideas between two intellectual domains that have been isolated from each other.

The authority relationship that exists between an employer and an employee, an economic relationship created by the employment contract, plays a central role in Simon's (1982) theory. Let employer B (for boss) hire employee W (for worker). We say that B exercises authority over W if W permits B to select behavior x . That is, W accepts authority when W 's behavior is determined by B 's decision. In general, W will accept authority if the decision is within W 's zone of acceptance.

We say that W enters into an employment contract with B when the former agrees to accept the authority of the latter and the latter agrees to pay the former a stated wage (w). This contract differs fundamentally from a sales contract—the kind of contract that is assumed in typical formulations of neoclassical price theory. In the sales contract, each

contractual party promises a specific consideration in return for the consideration promised by the other contractual party. The buyer (like B) promises to pay a stated sum of money, but the seller (unlike W) promises in return a specified quantity of a completely specified commodity. W will be willing to enter an employment contract with B only if it does not matter to W very much which x (within the agreed-on area of acceptance) B will choose or if W is compensated in some way for the possibility that B will choose an x that is not desired by W (e.g., that B will ask W to perform an unpleasant task).

It will be advantageous to B to offer W added compensation for entering into an employment contract if B is unable to predict with certainty, at the time the contract is made, which x will be the optimal one, from W 's standpoint. That is, B will pay for the privilege of postponing, until some time after the contract is made, the selection of x . This option to wait has real options value, which is explained more fully and illustrated by a mathematical example in Chapter 5.

Simon's (1982) model deals with a particular problem of planning under uncertainty. It analyzes a business situation in which it may be advantageous to postpone a decision (selection of x) to gain from information obtained subsequently. The postponement of choice may be regarded as a kind of liquidity preference, where the liquid resource is the employees' time, instead of money.

A Comparison of Organization Theories. According to Simon (1982), the economic theory of the firm and the organization theory of the firm are both concerned with the behavior of a person, or people, trying to achieve certain goals by the manipulation of (strategic) variables at their disposal. The problem of optimal, rational, or efficient behavior with respect to these goals can be formulated as a problem of finding the maximum (with respect to the strategic variables) of some function that is taken as a measure of success in attaining these goals (e.g., in the theory of the firm, finding the output that maximizes economic profit). Theories of organization, to a greater extent than the economic theory of the firm, have been concerned not only with optimal solutions but also with the whole set of viable solutions—that is, solutions that permit the survival of the organization (e.g., in the theory of the firm, outputs that yield at least a competitive rate of economic return).

In the neoclassical economic theory of the firm, a single participant, the entrepreneur, is explicitly treated as a rational individual. The other participants—employees, customers, and suppliers—enter into the neoclassical theory only implicitly as conditions to which the entrepreneur

adjusts in finding an economic solution that is optimal to the entrepreneur. One such condition is the price of the input factor labor; another is the demand schedule, which describes the behaviors of customers.

In the organizational theory, participants are generally treated in a more symmetrical way (Barnard, 1938; Simon, 1947). Participants are offered inducements for their participation in the organization. Through their participation, organizational members make contributions to the organization. The organization transforms its members' contributions into inducements that the organization, in turn, distributes to these members.

As a simple example, consider an organization with an entrepreneur, one employee, and one customer. The system of inducements and contributions may then be represented thus:

Participant	Inducements	Contributions
Entrepreneur	Revenue from sales	Cost of production
Employee	Wage	Labor
Customer	Goods	Purchase prices

Organization theory has generally been concerned not so much with optimality as with the conditions necessary for organizational survival, that is, the conditions under which the participants will continue to participate. The picture of the firm that is emerging from this research is that of a searching, information-processing, satisficing, allocating mechanism. Human thinking is an important—indeed, the most important—economic resource. The bulk of the productive wealth of our economy is not embodied in factories and machines but is found in the knowledge and skills stored in individuals' minds.

The Scarcity of Attention. A rabbit-rich world is a lettuce-poor world, and vice versa. Similarly, in an information-rich world, an abundance of information means a dearth of something else: a scarcity of whatever information consumes. Information consumes the attention of its recipients. In an information-rich world, most of the economic costs of information are the costs incurred by the recipients. It is not enough to know how much it costs to produce and transmit information; we must also know how much it costs, in terms of scarce attention, to receive information. As I later show, in Chapter 5, this idea informs Penrose's (1959) resource-based theory of managerial attention as the scarce resource, which is the binding constraint on the rate of the growth of the firm.

Many proposals for eliminating information overload (another phrase to describe life in an information-rich world) call for a new information system. An information-processing subsystem (a computer or new organization unit) will reduce the net demand on the rest of the organization's attention only if this subsystem absorbs more information previously received by others than it produces—that is, if this subsystem listens and thinks more than it speaks. The proper aim of a management information system is not to bring managers all the information they need but to reorganize the managers' environment of information to reduce the amount of time they must devote to receiving information. Restating the organization problem this way leads to a very different system design.

Simon (1982) notes that if a library has only one copy of each book, the library still has a high degree of informational overlap. Simon conjectures that if half of the titles in the Library of Congress were destroyed at random, little of the world's knowledge would be lost. The most important form of redundancy derives from the world's being highly lawful. Facts are lawful if they can be predicted from other facts. We need store only the fraction needed to predict these other facts. This process is exactly what science does: the process of replacing large amounts of unordered facts with tidy statements of orderly relations from which these facts can be inferred.

Simon (1982) observes that the dream of thinking everything out before we act, of making certain we have all the facts and know all the consequences, is a sick Hamlet's dream. It is a dream of someone with no analytical appreciation of the seamless web of causation, the limits of human thinking, or the scarcity of human attention.

Theories of Decision Making in Economics and Behavioral Science. Simon (1982) notes that the neoclassical economic theory of markets with perfect competition and rational agents is deductive theory that requires almost no contact with empirical data once the assumptions are accepted. Undoubtedly, there is an area of human behavior that fits these assumptions to a reasonable approximation, where the neoclassical microeconomic theory with its assumptions of rationality is a powerful and useful tool. Without denying the existence of this area, or its importance, Simon observes that neoclassical microeconomic theory fails to include some of the central problems of conflict and dynamics with which organization theory and strategic management have become increasingly concerned. Simon provides a metaphor to help show the reason for this failure:

Suppose we were pouring some viscous liquid—molasses—into a bowl of very irregular shape. What would we need to make a theory of the form the molasses would take in the bowl? How much would we have to know about the properties of molasses to predict its behavior under the circumstances? If the bowl were held motionless, and if we wanted only to predict behavior in equilibrium, we would have to know little, indeed, about molasses. The single essential assumption would be that the molasses, under the forces of gravity, would minimize the height of its center of gravity. With this assumption, which would apply as well to any other liquid, and a complete knowledge of the environment—in this case the shape of the bowl—the equilibrium is completely determined. Just so, the equilibrium behavior of a perfectly adapting organism depends only on its goal and its environment; it is otherwise completely independent of the internal properties of the organism. If the bowl into which we were pouring the molasses were jiggled rapidly, or if we wanted to know about the behavior before equilibrium was reached, prediction would require much more information. It would require, in particular, more information about the properties of molasses: its viscosity, the rapidity with which it adapted itself to the containing vessel and moved toward its goal of lowering its center of gravity. Likewise, to predict the short-run behavior of an adaptive organism, or its behavior in a complex and rapidly changing environment, it is not enough to know its goals. We must know also a great deal about its internal structure and particularly its mechanisms of adaptation. (p. 255)

Simon (1982) argues that broadening the definition of rationality to encompass goal conflicts and uncertainty made it difficult to ignore the distinction between the objective environment in which economic actors really live and the subjective environment that they perceive and to which they respond. When this distinction is made, we can no longer predict their behavior—even if they behave rationally—from the characteristics of the objective environment. *We also need to know something about their perceptual and cognitive processes.* Simon maintains that models of satisficing behavior are richer than models of maximizing behavior because models of satisficing behavior consider not only equilibrium but also the method of reaching equilibrium. Neoclassical economic theory is a theory of an individual choosing among fixed and known alternatives, to each of which the known consequences are attached. But when perception and cognition intervene between the decision maker and an objective environment, neoclassical economics no longer proves adequate. We need a description of the choice process that recognizes that alternatives are not given but must be sought and a description that takes into account the arduous task of determining which consequences will follow from each alternative.

Decision makers' information about their environment is actually much less than an approximation to the real environment. The term

approximation implies that the subjective world of decision makers resembles the external environment closely but lacks, perhaps, some fineness of detail. Actually, the perceived world is quite different from the "real" world. The differences involve both omissions and distortions and arise in both perception and inference. The sins of omission in perception are arguably more important than the sins of commission. Decision makers' mental models of the world encompass only a minute fraction of all the relevant characteristics of the real environment, and these inferences extract only a small fraction of all the information that is present.

Perception is sometimes referred to as a *filter*. This term is as misleading as *approximation* and for the same reason: Perception implies that what comes through to the central nervous system is really quite a bit like what is out there. In fact, the filtering is not merely a passive selection of some part of a presented whole but is an active process involving attention to a very small part of the whole and exclusion, from the outset, of almost all that is not within the scope of attention.

Simon (1982) argues that every human lives in an environment that generates millions of bits of new information each second, but the bottleneck of the perceptual apparatus certainly does not admit more than 1,000 bits per second and probably much less. Equally significant omissions occur in the processing that takes place when information reaches the brain. There are hosts of inferences that might be drawn from the information stored in the brain that are not in fact drawn. The consequences implied by information in the memory become known only through active information processing and hence through active selection of particular problem-solving paths from the myriad problem-solving paths that might have been followed.

Theories of Bounded Rationality. Simon (1982) argues that rationality, as is typically defined in the social sciences, denotes behavior that is appropriate to the achievement of given goals, within the limits imposed by given constraints. Those theories that postulate important constraints arising from the limitations of the actors themselves as information processors may be called theories of bounded rationality.

In some sense, chess is a trivial game: If the complete decision tree of possible games was fully known, there would be nothing of interest left to play. Unfortunately, the triviality of chess, as viewed from this high level of abstraction, offers no practical guide to a player in actually choosing a move. The proof that guarantees the validity of one (and only one) of three alternatives that the game must have a value of win,

lose, or draw for white, gives no practically usable method to determine the true outcome. This relative, human difficulty necessitates the use of those incomplete, heuristic methods of playing, which constitute "good" chess, and without this human difficulty there would be no element of struggle and surprise in this game. Simon (1982) emphasizes that the chess player's difficulty in behaving rationally has nothing to do with uncertainty—whether of consequences or alternatives—but it is a matter of complexity. There is no risk or uncertainty, in the sense in which those concepts are used in economics or in statistical decision theory, in the game of chess. It is a game of perfect information. No probabilities of future events need enter the calculations, and no contingencies, in a statistical sense, arise.

What we refer to as uncertainty in chess is uncertainty introduced into a perfectly certain environment by inability—computational inability—to ascertain the structure of that environment. But the result of the uncertainty, whatever its source, is the same; approximation must replace exactness in reaching a decision.

A satisficing decision procedure can often be turned into a procedure for optimizing by introducing a rule for optimal amount of search or, what amounts to the same thing, a rule for fixing the aspiration level optimally. Thus, the aspiration level in chess might be adjusted, dynamically, to such a level that the expected improvement in the move chosen, per minute of additional search, would just balance the incremental cost of the search.

Although such a reconstructed logic (Kaplan, 1964) is formally possible, to carry it out in practice requires additional information and assumptions beyond those needed for satisficing. First, the values of alternatives must be measured in units comparable with the units for measuring search costs to permit comparison at the margins. Second, the marginal productivity of search—the expected increase in the value per unit of search time—must be estimated on some basis. If one were designing a chess-playing program, it is doubtful whether effort spent in attempting to place the program in such a dynamic, optimizing framework would be nearly as worthwhile as an equivalent effort given to improving the selectivity of the program's move-generating and move-evaluating heuristics. Research on satisficing procedures has focused primarily on the efficiency of search—on the nature of the heuristic methods.

Simon (1982) observes that most of the formal techniques that constitute the technical backbone of management science and operations

research are procedures for finding the best of a set of alternatives in terms of some criterion. Linear programming and dynamic programming are among the more powerful of these techniques. The dominant approach to problems in this sphere has been to simplify the real-world problems to the point where the formal optimizing models can be used as approximations.

Perhaps the technique most widely used in management science to deal with situations too complex for the application of known optimization methods is simulation. In simulation, the trial and error is supplied by the human investigators rather than by the technique of analysis itself. The satisficing approach has been most often employed in models where heuristic or trial-and-error methods are used to aid the search for plausible alternatives. These computational tools make substantially more tractable the task of matching bounded capabilities with the difficulty of the problems.

From Substantive to Procedural Rationality. Simon (1982) uses the term *substantive rationality* to refer to the concept of rationality developed within economics and the term *procedural rationality* to refer to the concept developed within psychology. Behavior is substantively rational when such behavior is appropriate to the achievement of given goals within the limits imposed by given constraints. Notice that, by this definition, the rationality of behavior depends on the actors in only a single respect—their goals. Given these goals, rational behavior is determined entirely by the characteristics of the environment in which such behavior takes place.

Neoclassical economic analysis rests on at least two fundamental assumptions. The first assumption is that economic actors have particular goals, for example, utility maximization or profit maximization. The second assumption is that economic actors are substantively rational. Given these two assumptions, and given a description of a particular economic environment, economic analysis (descriptive or normative) could usually be carried out using such standard tools as differential calculus, linear programming, or dynamic programming.

Thus, the assumption of utility or profit maximization on the one hand, and the assumption of substantive rationality on the other hand, freed economics from any dependence on psychology. As long as these assumptions went unchallenged, there was no reason why economists should acquaint themselves with the psychological literature on human cognitive processes or human choice. There was absolutely no point at which the empirical findings of psychological research could be injected

into the process of economic analysis. The irrelevance of psychology to neoclassical economics was complete. Behavior is procedurally rational when such behavior is the outcome of appropriate deliberation. Its procedural rationality depends on the process that generated it. Historically, there have been three main categories of psychological research on cognitive processes: learning, problem solving, and concept attainment.

The search for computational efficiency is a search for procedural rationality, and computational mathematics is a normative theory of procedural rationality. In this normative theory, there is no point in prescribing a particular substantively rational solution if there exists no procedure for finding that solution with an acceptable amount of computing effort. So, for example, although there exist optimal (substantively rational) solutions for combinatorial problems of the traveling-salesman type, and although these solutions can be discovered by a finite enumeration of alternatives, actual computation of the optimum is infeasible for problems of any size and complexity. The combinatorial explosion of such problems simply outraces the capacities of computers, present and prospective.

Hence, a theory of rationality for problems like the traveling-salesman problem is not a theory of best solutions—of substantive rationality—but a theory of efficient computational procedures to find good solutions—a theory of procedural rationality. Notice that this change in perspective involves not only a shift from the substantive to the procedural but also a shift from concern for optimal solutions to a concern for good solutions (e.g., good decision rules for inventory and workforce smoothing).

Simon (1982) argues that the demands of computability led to two kinds of deviation from neoclassical optimization: simplification of the model to make computation of an optimum feasible or, alternatively, searching for satisfactory rather than optimal choices. Simon regards both of these solutions as instances of satisficing behavior rather than optimization. To be sure, using reconstructed logic we can formally view these as optimizing procedures by introducing, for example, a cost of computation and a marginal return from computation and using these quantities to compute the optimal stopping point for the computation. But the important difference between the satisficing procedures and the optimizing procedures remains. The problem has been shifted from one of characterizing the substantively optimal solution to one of devising practicable computation procedures for making reasonable choices.

Ignorance of the future prevents decision makers from behaving in a substantively rational manner; decision makers can only adopt a rational choice procedure, including a rational procedure for forecasting or otherwise adapting to the future. Once we become interested in the procedures—the rational processes—that economic actors use to cope with uncertainty, we must broaden our horizons further. Uncertainty calls forth a whole range of actions. These actions are at least of four kinds:

- Intelligence actions to improve the data on which forecasts are based, to obtain new data, and to improve the forecasting models
- Actions to buffer the effects of forecasting errors—for example, holding inventories, insuring, and hedging
- Actions to reduce the sensitivity of outcomes to the behavior of competitors—for example, steps to increase product and market differentiation
- Actions to enlarge the range of alternatives whenever the perceived alternatives involve high risk

As organizational economics and strategic management become more concerned with procedural rationality, they will necessarily have to borrow from psychology or build for themselves a far more complete theory of human cognitive processes. Even if our research interest in strategic management is in normative rather than descriptive behavior, we will need such a theory. There are still many areas of decision—particularly those that are ill-structured—where human cognitive processes are more effective than the best available optimization techniques or artificial intelligence methods. A great deal can still be learned about effective decision procedures by studying how humans make choices.

We can expect substantive rationality only in those situations that are sufficiently simple as to be transparent to the decision maker's mind. In all other situations, we must expect that the decision maker's mind will use such imperfect information as it has, will simplify and represent the situation as it can, and will make such calculations as are within its powers (Duhaime & Schwenk, 1985). We cannot expect to predict what the decision maker's mind will do in such situations unless we know what information it has, what forms of representations it prefers, and what algorithms are available to it.

In my judgment, there seems to be no escape from psychology. If organizational economics and strategic management are to deal with uncertainty, they will have to understand how humans in fact behave

in the face of uncertainty and by what limits of information and computability humans are bound. Bobby Fischer, in 1972, played chess differently from Paul Morphy, who played in 1861. Much of that difference was the result of the knowledge of the game that had accumulated over the century through the collective experience of the whole society of professional chess players. Organizational economics and strategic management are, like chess, inevitably culture bound and history bound. A business firm equipped with the tools of operations research does not make the same decisions, for example, concerning inventory management, as it did before it possessed such tools.

Simon (1982) maintains that (organizational) economics is one of the sciences of the artificial. Organizational economics is a description and explanation of human institutions, whose theory is no more likely to remain invariant over time than the theory of bridge design. Decision processes, like all other aspects of economic institutions, exist inside human heads. Decision processes are subject to change with every change in what humans know and with every change in their means of calculation. Simon submits that for this reason the attempt to predict, and prescribe, human economic behavior by deductive inference from a small set of unchallengeable premises must fail and has failed.

Simon (1982) suggests that organizational economics will progress as we deepen our understanding of human thought processes and will change as human individuals and human societies use progressively sharpened tools of thought in making their decisions and designing their institutions. A body of theory for procedural rationality is consistent with a business world in which humans continue to think and continue to invent; a theory of substantive rationality is not.

Simon (1982) notes that the shift from theories of substantive rationality to theories of procedural rationality requires a basic shift in scientific style, from an emphasis on deductive reasoning within a tight system of axioms to an emphasis on detailed empirical exploration of complex algorithms of thought. As organizational economics becomes more involved in the study of uncertainty, and more concerned with the complexity of business decision making, the shift in research program becomes inevitable. Wider areas of organizational economics and strategic management will replace the oversimplified assumptions of situationally constrained omniscient decision makers, with a realistic (and psychological) characterization of the limits on decision makers' rationality, and the consequences of those limits for their economic and managerial behavior.

Simon (1982) argues that complexity is deep in the nature of things, and discovering tolerable approximation procedures and heuristics that permit huge spaces to be searched selectively is at the heart of intelligence, whether human or artificial. A theory of rationality that does not give an account of problem solving in the face of complexity is sadly incomplete. It is worse than incomplete; such theory can be seriously misleading by providing “solutions” to organizational economic questions that are without operational significance. The theory of heuristic search, cultivated in artificial intelligence and information-processing psychology, is concerned with devising or identifying search procedures that will permit systems of limited computational capacity to make complex decisions and to solve difficult problems. As Franco Modigliani was fond of saying, “If businessmen are not now maximizers, after enough of them have graduated from business school, they will be.” So we might even expect that a positive theory of organizational economic behavior and strategic management will have to include as a subtheory the way in which business schools produce, and diffuse, decision-making techniques. Procedural rationality is the rationality of a person for whom the time and effort required for computation are scarce human resources.

Simon (1982) concludes by noting that there is a saying in politics that “you can’t beat something with nothing.” You can’t defeat a measure, or a candidate, simply by pointing to defects and inadequacies. You must offer an alternative. What then is the status of the neoclassical economic theory of the firm? There can no longer be any doubt that the micro-analytic assumptions of neoclassical economic theory—the assumptions of perfect rationality—are contrary to fact. It is not a question of approximation; the assumptions of perfect rationality do not even remotely describe the processes that humans use for making decisions in complex business situations.

Moreover, there is an alternative. If anything, there is an embarrassing richness of alternatives. Today, we have a large mass of descriptive data from both laboratory and field, which show how human problem solving and decision making take place in a wide variety of situations. A number of theories incorporate the replacement of optimization by targets and satisficing goals and mechanisms of learning and adaptation. If our research interest is in descriptive decision theory (or even normative decision theory), it is now clear that the neoclassical economic theory of the firm has been challenged by a superior alternative that provides researchers with a much closer approximation to what

is actually going on here.² Now that we have studied the behavioral theory of the firm from the Carnegie School, I turn next to Chapter 2 concerning transaction costs theory. Oliver Williamson was a doctoral student at Carnegie in the 1960s, and we shall see the influence of Richard Cyert, James March, and especially Herbert Simon on Williamson's (1975, 1985, 1996) transaction costs theory. Indeed, transaction costs theory combines their works with economics and aspects of the law in an effort to reconceptualize the problem of economic organization. Organization theory supports transaction costs theory in terms of insisting that workably realistic behavioral assumptions are an alternative to the assumptions of economics, which are typically chosen for analytical convenience; the autonomous adaptation of the market is joined with cooperative adaptation by organizations; and the embeddedness (e.g., the institutional environment) of a transaction matters (Granovetter, 1985; North, 1990). Before analyzing Williamson (1975), however, we begin the next chapter with Arrow's (1974) *The Limits of Organization*, which is followed by Coase (1988).

²For further theory development and applications of the Carnegie School, see Allison (1971); Earl (2001); Gimeno, Folta, Cooper, and Woo (1997); Mahoney (1992c); March (1988, 1999); Scott (1987); Simon (1957, 1996); and Thompson (1967). Several consequences of bounded rationality have been observed, including (1) selective perception of information; (2) an adaptive, sequential manner of information processing; (3) mental effort that is reduced by heuristic procedures; and (4) a process of active reconstruction for memory. Systematic biases result with insensitivity to prior probability of outcomes, insensitivity to sample size, misconceptions of chance, failure to recognize regression toward the mean, the retrievability of instances, illusory correlation, insufficient adjustment and anchoring, and biases in the evaluation of conjunctive and disjunctive events (Kahneman, Slovic, & Tversky, 1982). Given the limitations and (systematic) biases of the individual, those operating from a behavioral perspective tend to view the organization as a more efficient information processor than any given individual. The firm is considered to be an institutional response to uncertainty and bounded rationality—a theme that I discuss later in transaction costs theory.





2

Transaction Costs Theory

The origin of transaction costs theory is Coase's (1937) classic journal article on the nature of the firm. However, it took until the mid-1970s for transaction costs theory to become influential in both research and public policy following the works of Arrow (1974) and especially Williamson (1971, 1979). This chapter covers Arrow (1974), Coase (1988), and Williamson's three transaction costs books (1975, 1985, 1996).

Arrow's (1974) book, *The Limits of Organization*, was originally given as the Fels Lecture for 1970–1971 to the Fels Center of Government at the University of Pennsylvania. This lucidly written book focuses on at least four major themes: (1) the concept of rationality (individual and social), (2) information economics, (3) the agenda of organizations, and (4) the concepts of authority and responsibility. Throughout this work, Arrow considers the (often conflicting) demands of society and the needs of the individual and insists that some sense of balance is required.

Coase's (1988) book, *The Firm, the Market and the Law*, is a collection of his journal articles and economic writings. Coase argues that if we move from a regime of zero transaction costs to one of positive transaction costs, what becomes immediately clear is the crucial importance of the legal system. Coase maintains that it makes little sense for economists to discuss the process of exchange without specifying the institutional setting within which the trading takes place, since the institutional setting affects the economic incentives to produce and the economic costs of transacting.

Williamson's (1975) *Markets and Hierarchies* is a work of scholarship written for posterity. This research book combines ideas from Commons (1934), Coase (1937), Barnard (1938), Simon (1947), March and Simon (1958), Chandler (1962), and Arrow (1974), among others. In particular, Williamson emphasizes the fundamental importance of information (Arrow, 1974), organizational innovation (Chandler, 1962), transaction costs (Coase, 1937), and behavioral assumptions

(Simon, 1947). Simultaneously, Williamson provides original ideas concerning vertical integration and the theory of the firm that have proven to be fruitful for the evolving science of organization.

Students of strategic management are well advised to study closely transaction costs theory. Transaction costs theory combines logical rigor with practical relevance to help describe, explain, and predict governance based on comparative efficiency criteria. In addition, the habits required of good scholarship can be learned from Williamson (1975):

- Describe others' works fairly and accurately.
- Search extensively to find and develop a plausible theory to help understand the phenomena at hand.
- Have an active mind.
- Work from a disciplinary base.
- Have the courage to be interdisciplinary, if so inclined.

The organizational economics principles found in transaction costs theory are durable. Indeed, we are currently witnessing greater vertical deintegration (e.g., strategic outsourcing), arguably as a result of fundamental transaction costs changes—impacting input and output measurement costs and asset specificity (Mahoney, 1992b)—that are due to dramatic changes in the development and diffusion of information technology (Shapiro & Varian, 1999). I believe this research area to be especially fruitful for students in the next generation of strategic management research.

Williamson (1985) documents well that empirical research testing transaction costs theory has been largely corroborative. Nevertheless, like all empirical research studies, students today continue to be challenged by potential specification problems, measurement problems, and identification problems in econometric testing. Given the current state of the empirical research literature testing transaction costs theory (and even more so in other research areas of organizational economics), students in the next generation can arguably do better.

Williamson (1996) continues the agenda of joining law, economics, strategic management, and organization theory. A systems perspective is employed to look at (incomplete) contracting in its entirety. Credible commitments are viewed as ways of safeguarding against contractual hazards inherent in incomplete contracts.

Like Barnard (1938) and Simon (1947), Williamson's (1975, 1985, 1996) transaction costs trilogy provides a conceptual framework simple

enough to be used and complex enough to accommodate continuing insights into the workings of organization. Students of strategic management taking an organizational economics approach should take heart in achieving and pursuing the science of organization.

The Limits of Organization (*Arrow, 1974*)

This chapter begins with Arrow (1974), who, along with Williamson (1971, 1975), was among the earliest economists to build on Coase (1937, 1960).

Rationality: Individual and Social. Arrow (1974) states that the intricacies and paradoxes in relations between individuals and their actions in social contexts have been put well by the sage Rabbi Hillel: "If I am not for myself, then who is for me? And if I am not for others, then who am I? And if not now, when?" *Here we have, in three successive questions, the essence of a tension that most feel between the claims of individual self-fulfillment and the claims of social conscience and action.* Tensions between society and the individual are inevitable. Their claims compete within the individual conscience, as well as in the area of social conflict.

Arrow (1974) discusses the relationship between society and the individual in the rational spirit of the economist and notes that economists by training think of themselves as guardians of rationality, ascribers of rationality to others, and the prescribers of rationality to the social world. It is this social science role that Arrow chooses to play.

Arrow (1974) observes that a truly rational discussion of collective action in general or in specific contexts is necessarily complex, and such a discussion is necessarily incomplete and unresolved. Arrow takes the (logical positivist) position of Simon (1947) and maintains that rationality has to do with means and ends and their relation. The concept of rationality does not specify what the ends are.

The role of an economist is sometimes unpleasant. It's probably not entirely accidental, though a little unfair, that Thomas Carlyle referred to economists as the practitioners of the dismal science. Economists frequently have to point out the limits of our opportunities. Economists have to say, "This or that, not both. You can't do both." Moreover, economists have to point out frequently that the economic system is complex in its nature. It can easily happen that a step that on its face is an obvious way of achieving certain desired values may, in fact, frequently lead to the opposite. Arrow (1974) cites, for example, many proposals for high increases in minimum wages. Surely, most would want to redistribute

income to the lower end of the wage scale. The most obvious thing to do is to raise the wages. Economists realize that the situation is not that simple, that the system can react to that policy; the system does not passively accept such a change. The end result may be a substitution of capital for labor and consequently an increase in the volume of unemployment, an economic outcome that is arguably worse than low wages.

The basic resources of society—its natural resources, its human resources, and its technological resources—are limited in supply, and the realization of alternative values or the search to find alternative activities for meeting those values implies a competition for these scarce resources. If we do things one way, we cannot do them another way. So we need to have a system that mediates this competition for the services of scarce resources, whether it is a market or an authoritative allocation system, as in the military or in the socialist state. Further, interpersonal organization is needed to secure the economic gains that can accrue from cooperation. The essential considerations are two: Individuals are different and in particular have different talents, and individuals' efficiency in the performance of social tasks usually improves with specialization. We need cooperation to achieve specialization of function. This cooperation and specialization involve all the elements of trade and the division of labor (Stigler, 1968).

How do we evaluate alternative social organizations? Again, for a commonplace of economic thought, we use the concept of efficiency or optimality that is associated with the name Vilfredo Pareto. Whatever else we mean by better or efficient, we certainly mean the following: One situation, one system, or one allocation is better than another if every individual feels it is better according to his or her own individual values. Under specific assumptions, it is shown that efficiency can be achieved through a particular kind of social system, the price system. Although not mentioned by Arrow (1974), a state in which some people are starving and suffering from acute deprivation while others are tasting the good life can still be Pareto optimal if the poor cannot be made better off without cutting into the pleasures of the rich—no matter by how small an amount. Pareto optimality is faint praise indeed.

Arrow (1974) notes that in a price system, individuals have a sense of freedom. Individuals are free to act within the system; there is no direct order telling individuals what to do. Individuals have an income, and they can spend it. Needless to say, this freedom of action is, from a certain point of view, somewhat illusory. This freedom of action can be very small indeed if their income is very low. One's income is not

determined by justice but by a complicated system of interactions where the ethical meaning is difficult to define. *The idealization of freedom through the market completely ignores that this freedom can be, to a large number of people, very limited in scope.*

The price system can also be challenged on the grounds that it harnesses self-serving motives that our ethical systems frequently condemn. The price system makes a virtue of selfishness. We should, on the other hand, not ignore the enormous economic gains in efficiency that can be achieved through the price system, as compared with most conceivable alternatives.

Nevertheless, there are profound difficulties with the price system, even within its own logic, and these difficulties strengthen the view that, *valuable though the price system is in certain realms, the price system cannot be made the complete arbiter of social life. The price system does not, in any way, prescribe a just distribution of income.*

There are other drawbacks, which can be discussed in a somewhat more precise way. In fact, in a strictly technical sense, the price system does not always work. You simply cannot price certain things. A classic example of considerable importance is the pollution of water and air.

Arrow (1974) observes that trust has an important pragmatic value. Trust is an important lubricant of a social system. Trust is extremely efficient; it saves a lot of trouble to have a fair degree of reliance on other people's word. Unfortunately, trust is not a commodity that can be purchased very easily. If you have to buy trust, you already have some doubts about what you have purchased. Trust and similar values, loyalty, or truth telling are examples of what economists call (positive) externalities. These positive externalities have real, practical, economic value: They increase the efficiency of the system, and they enable people to produce more goods or more of whatever values are held in high esteem. But trust and similar values are not commodities for which trade on the open market is technologically possible or even meaningful.

From the perspective of efficiency, as well as from the perspective of distributive justice, something more than the market is called for. Other modes of governing the allocation of resources are needed. Most conspicuous among these governance modes is the government at all its levels. Government influences the allocation of resources by means that operate within the price system, but also otherwise. Government has its host of laws and regulations, coercive and certainly nonmarket methods of controlling and directing the economy and indeed society in general. Government's role in internalizing externalities is then straightforward in principle, which does not imply that it is easy in practice.

With regard to distributive justice, the root facts here are the incommensurability and incomplete communicability of human wants and human values. George Bernard Shaw long ago observed, "Do not do unto others as you would have them do unto you. They may have different tastes."

Conflict, to be sure, is mitigated by the essential human feeling of sympathy, a sense of feeling oneself to be in the other person's place. This motive operates with some (though doubtless inadequate) strength, and this motive can operate better in an institution such as the government, designed to give some scope to expressing altruistic interests.

A firm, especially a large corporation, provides another major area within which price relations are held in partial abeyance. The internal organization is hierarchical and bureaucratic. Prices no doubt have powerful influences from the outside, and in many firms there are concerted attempts to simulate the operations of a market, perhaps even to do better than the sluggish and imperfectly informed markets of economic reality. But internally, and especially at lower levels, the relationships among the employees of a firm are different from arm's-length bargaining. As Simon (1947) observed, an employment contract is different in many ways from a commodity contract; an employee is willing to obey authority. The employee is free to leave, but because the transaction costs of leaving are always present and frequently nontrivial, the employment relationship creates an expectation of continued participation.

There is still another set of institutions that Arrow (1974) calls to our attention. These are invisible institutions: the principles of ethics and morality. Societies in their evolution have developed implicit agreements to certain kinds of regard for others, agreements that are essential to the survival of the society or at least contribute greatly to the efficiency of its working. It has been observed, for example, that among the properties of many societies whose economic development is backward is a lack of mutual trust. Collective undertakings of any kind, not merely governmental, become difficult or impossible. And it is clear that this lack of social consciousness is a distinct economic loss in a very concrete sense, as well as a loss in the possible well running of a political system.

That we cannot mediate all our responsibilities through prices, through paying for them, makes it essential in the running of society that we have what might be called *conscience*, a feeling of responsibility for the effects of one's actions on others. Unfortunately, this philosophical position cannot be pushed too far. We cannot know all the effects of our actions on all other people. When you take these obligations to others

seriously you are forced into a difficult (existential) position where you take actions whose consequences you cannot really know and yet you feel responsible for these actions.

Social demands may be expressed through formal rules and authorities, or social demands may be expressed through internalized demands of conscience. Looked at collectively, these demands may be compromises that are needed to increase the efficacy of all. At any moment these demands are apt to be felt by the individual as a set of shackles. And, unfortunately, there are still further problems. A commitment to a war or to a revolution or to a religion is typically one that is difficult to reverse, even if conditions have changed from the time when the commitment started. Even if experience has shown the unexpectedly undesirable consequences of a commitment, the past may continue to rule the present. It is such thinking, Arrow (1974) argues, that gives rise to the greatest tragedies of history, this sense of commitment to a past purpose that reinforces the original agreement precisely at a time when experience has shown that such a commitment must be reversed. Rationality and foresight are indeed capable of creating delay and doubt; so, too, are conscience, respect for others, and the sense of vague respect to distant and unanticipated consequences that we may worry about. There are no simple answers here and Arrow does not give any. There are moments in history when we simply must act, fully knowing our ignorance of possible consequences, but to retain our full rationality we must sustain the burden of action without certitude, and we must always keep open the possibility of recognizing past errors and changing course.

Organization and Information. Arrow (1974) submits that organizations are a means of achieving the benefits of collective action in situations in which the price system fails. The term *organization* is interpreted quite broadly. Formal organizations, firms, labor unions, universities, and government are not the only kind. Ethical codes and the market system itself are interpreted as organizations. The purpose of organizations is to use the fact that many (virtually all) decisions require the participation of many individuals for their effectiveness. Arrow stresses one particular cause for the failure of the price system to allocate resources perfectly, one that is central to the understanding of organizations—the presence of uncertainty. Uncertainty may be about conditions of production or tastes that, if known, would affect individuals' desires to trade. Then, instead of contracts to buy and sell fixed amounts of goods, it would be better to have conditional contracts or contracts in contingent commodities. But the range of contingencies

for which conditional contracts are available is much more limited than would be ideally desirable in theory. The taking of desirable economic risks is inhibited by the inability to insure against business failure, for example.

There is more than one reason for the failure of the theoretically desirable contingent prices to exist. One reason is the sheer complexity of the price schedule. An insurance policy would have to specify an enormous number of contingencies with, in general, different payments for each possibility. Drawing up such contracts would be expensive. Moreover, the courts of law, on the basis of long experience, have shown little faith in the ability of the average individual to understand complicated contracts.

Another major reason for limitation of the price system for allocating risk bearing is the difficulty of distinguishing between genuine risks and failures to optimize, a difficulty known as *moral hazard*. For example, the outbreak of a fire may be due to a combination of exogenous circumstances and individual choice, such as carelessness or, in the extreme case, arson. Hence, a fire insurance policy creates an economic incentive for individuals to change their behavior and ceases to be pure insurance against an uncontrollable event.

Consider also the problem known in the insurance literature as *adverse selection* (Arrow, 1971). The insured may know their risks better than the insurer, for example, in life insurance. The insurer may start by choosing rates on some actuarial basis. But then, the high-risk groups will buy more of the insurance than the average, while the low-risk groups will buy less. Hence, the experience of the insurer, as weighted by dollars, will be less favorable than the actuarial. The rates will have to be raised, but this increase in rates drives still more of the low-risk groups out. A situation will be created in which there are many whose risks are inadequately covered because it is not known how low those risks really are. The essential cause of market imperfections in this case is asymmetric information between the two contractual parties. Akerlof (1970) provides a mathematical model where adverse selection can lead to such market failure in the market for lemons for used cars.

Another illustration of asymmetric information among economic agents is the relation between patient and physician. It is of the essence of this relationship or other relations between principal and agent that they differ in their information about the world. But this information asymmetry means that there can really be no contract that ensures against the agent's failure to do business properly. Arrow (1974) argues

in the context of medical economics that one might regard professional ethics as an example of an institution that fills, in some measure, the gap created by the corresponding failure of the price system.

Consider the organization as a processor of information. The scarcity of information-handling capability is an essential feature for the understanding of both individual and organizational behavior. The transformation of probabilities due to signals is precisely what constitutes the acquisition of information.

This definition of information is qualitative, so it will remain for Arrow's (1974) purposes. The quantitative definition that appears in information theory is probably of only limited value for economic analysis, for reasons pointed out by Marschak (1968); different bits of information, equal from the viewpoint of information theory, will usually have different economic benefits or costs. Thus, let statements *A* and *B* be any two statements about the world, for neither of which is its truth or falsity known *a priori*. Then a signal that *A* is true conveys exactly as much information as a statement that *B* is true. But the economic value of knowing whether or not *A* is true may be vastly greater than the economic value of knowing *B*'s truth value, or it may be that the resources needed to ascertain the truth value of *A* are much greater than those for *B*. In either case, the information-theoretic equivalence of the two possible signals conceals their vast economic difference. There is little that one can say systematically about the economic benefits for information in general. The main remark that can be ventured is that there are increasing returns to the uses of information.

Let us now turn to the economic costs of information. First, and most important, individuals themselves are inputs. Immediately or ultimately, the information must enter their brains through their sensory organs, and brains and senses are limited in capacity. Information may be accumulated in files, but this information must be retrieved to be useful in decision making. The psychological literature has many empirical research studies of the limits on the sensory perception abilities of humans and some on their limits as information processors. Individuals' limited capacity for acquiring and using information is a fixed factor in information processing, and one may expect a sort of diminishing returns to increases in other information resources. Organization theorists have long recognized limits of this kind under the heading "span of control."

A second key characteristic of information costs is that information costs are in part capital costs, and more specifically they typically represent

an irreversible investment. Arrow (1974) does not place much weight on the physical aspects of communication, telephone lines, and the like. Rather, Arrow emphasizes the need for having made an adequate investment of time and effort to be able to distinguish one signal from another. Learning a foreign language is an obvious example.

Now, by its very nature, the economic value of an (irreversible investment) information channel is uncertain, so we have an economic problem that resembles the demand for inventories under conditions of uncertainty. We may venture some possible generalizations. One is that the demand for investment in information is less than the demand would be if the economic value of the information were more certain. The second generalization is that the random accidents of history (initial conditions) will play a bigger role in the final equilibrium. Once the investment has been made and an information channel acquired, it will be cheaper to keep using this information channel than to invest in new channels. Thus, it will be difficult to reverse an initial commitment in the direction in which information is gathered.

A third basic characteristic of information costs is that information costs are not uniform in different directions. At any given moment, individuals are bundles of abilities and accumulated information. Individuals may easily find it cheaper to open certain information channels rather than others in ways connected with these abilities and this knowledge (exhibiting path dependencies). It is also easier to communicate with other individuals with whom one has a common approach or a common language.

The Agenda of Organizations. Arrow (1974) notes that in neoclassical microeconomic (maximizing) theory it is implicit that the values of all relevant variables are at all moments under consideration. All variables are therefore agenda of the organization; that is, their values always have to be chosen. On the other hand, in everyday observation and in research studies of organization, difficulty in arranging for a potential decision variable to be placed on the agenda—once it has been recognized as worthy of time and attention—may be greater than choosing a value for the variable.

Arrow (1974) maintains that the combination of uncertainty, indivisibility, and (idiosyncratic) capital intensity associated with information channels and their use imply the following: that the actual structure and behavior of an organization may depend heavily on history and that the very pursuit of efficiency may lead to core rigidity and unresponsiveness to further change.

Decisions are necessarily a function of information. Hence, if it is decided to collect no information relevant to a certain class of decisions, those decisions are nonagenda. A decision area may be active, monitored, or passive. An active area is one in which experiments are performed, signals received from them, and terminal acts chosen as a function of the signals. A monitored area is one in which some experiments are being performed; the signals received convey too little information to take terminal acts, but if appropriate signals are received, it is optimal to make further experiments, which in turn will yield enough information to bring the terminal acts onto the agenda. Finally, a passive area is one in which no experiments are being conducted.

Experience may place an item on the agenda. According to William James, a coercive fact may be more persuasive than any speculation about the potential benefits from change. The sinking of the *Titanic* led to iceberg patrols. No doubt the changes in payoffs may be changes in perceptions rather than in actuality.

In general, the information received by a member of the organization can be transformed into a much smaller volume for retransmission without losing value for choice of terminal acts. The theory of sufficient statistics is an example of this reduction of information without loss of value. It is this reduction in retransmission that explains the utility of an organization for information handling. Since information is costly, it is better, in general, to reduce the internal transmission still further. The efficiency of a channel can be increased by a suitable choice of code. The teaching and learning of codes by individuals are acts of irreversible investments for them. It is therefore also irreversible capital accumulations for organizations. It follows that organizations, once created, have distinct identities, because the costs of changing the code are those of unanticipated obsolescence.

History matters. The code is determined in accordance with the best expectations at the time of the firm's creation. Since the code is part of the firm's organizational capital, the code of a given firm will be modified only slowly over time. Hence, the codes of firms starting at different times will in general be different even if they are competitive firms. The need for codes mutually understandable within the firm imposes a uniformity requirement on the behavior of the participants. They are specialized in the information capable of being transmitted by the codes so that they learn more in the direction of their activity and become less efficient in acquiring and transmitting information not easily fitted into the code. Hence, the firm itself serves to mold the behavior of its members.

If we think of education as the primary source of new information, then the youngest and newest members of the organization introduce new information into the organization. Thus, we have the possibility of changes in organizational agenda induced by generational changes. More generally, the prime need in organizational design is increasing capacity to handle a large agenda. To the extent that information and its handling are accumulations of personal capital, what is needed is the circulation of elites, the turnover of decision makers. More generally, what is needed is a circulation of information and decision rules. Short-run efficiency within a narrow framework of alternatives may be less important in the long run than a wide compass of potential activities that have real options value.

Authority and Responsibility. Arrow (1974) notes that among the most widespread characteristics of organizations is the prevalence of the authority relationship. Virtually universally, in organizations of any size, decisions are made by some individuals and carried out by others. The fields in which an authority is valid may be limited, and the recipients of orders at one level may have their own field of authority. But within these limits, the giving and taking of orders, having someone tell someone else what to do, is an essential part of the mechanism by which organizations function.

The role of authority does vary among organizations. The military is the extreme case, in which authority is all pervasive and essential. The state also exemplifies authoritative behavior in relation to its citizens, particularly with respect to police and legal control. The state shares with firms a more limited kind of authoritative control over employees. Indeed, as Simon (1947) emphasized, an employment contract is precisely a contract on the part of the employee to accept authority. An employment contract differs, therefore, from a contract to purchase a commodity; what is bought and sold is not a definite objective thing but rather a personal relationship. Within the scope of the contract, the relation between the employer and employee is no longer a market relation but an authority relationship. Of course, the scope of this authority will usually be limited by the terms of the contract, and, more fundamentally, the scope of this authority is limited by the freedom with which an employee can leave the job. But since there are normally transaction costs to exercise this freedom, the scope of this authority is not trivial.

When either interests or information differ among the members of the organization, the costs of achieving consensus rise and hence the value of consensus as a mode of organizational decision making declines relative

to that of authority. Despite the vast research literature in this area, including theoretical development of the theory of games, we are far from a good understanding. It is certainly clear that the process of bargaining can itself be a costly one, especially when the successive offers and threats take place not in the play world of recontract but in the real world of economic ruin and the savage destruction of human lives in war.

The aim in designing institutions for making decisions should be to facilitate the flow of information to the greatest extent possible. This design involves the reduction of the volume of information while preserving as much of its value as possible. To the extent that reduction of volume is accomplished by reduction in the number of information channels, we are led back to the superior efficiency of authority.

Organization is a means of handling social functions when the price system fails. Within the firm, the sanctions that authority can use are basically those of hiring and firing. The state employs the sanction of the criminal law. At one level of analysis, this is a suitable answer. These sanctions do operate, and decisions by authority are obeyed, in part, because of the punishments that might otherwise ensue.

Ultimately, it seems to Arrow (1974), *authority is viable to the extent that it is the focus of convergent expectations*. Individuals obey authority because they expect that others will obey authority. Traffic laws, and in particular signal lights, may be obeyed because it is clearly worthwhile to have a system in which everybody obeys them.

Thus, it is important to make authority visible so that it serves as a coordinating signal. This need for signaling is why external symbols surround authority. The emphasis on convergent expectations as the source of authority implies its fragility.

The efficiency loss due to informational overload is increased by the tendency to filter information in accordance with one's preconceptions. It is easier to understand and accept information congruent with previous beliefs than to overcome cognitive dissonance. Political and especially military history from Pearl Harbor to Vietnam is filled with dismal and disastrous examples. To go to an earlier period, when the *Titanic* began to broadcast for help, the captain of a nearby ship decided that the message must be a hoax; it was well-known that the *Titanic* was unsinkable. For another important perspective on authority, Fromm (1941) emphasizes man's psychological craving for authority.

It is difficult to imagine an organization in which some element of responsibility does not exist, at least in the long run. First, every real organization is of limited scope. Hence, as Hirschman (1970) stressed, exit

from an organization is always possible, though possibly at considerable cost. Ultimately, an authority can be held to account for the exit of its organizational members. Disobedience to orders, organized or unorganized, frequently sets limits to authority, and, like many other sanctions, the fear of such disobedience constitutes an internalization of responsibility.

In my view, Arrow (1974) provides an exemplary contribution to the economics of organization. The book combines microeconomic logic, transaction cost economics, and behavioral economics. Now that we have studied Arrow's (1974) classic, we turn next to Coase (1988) and *The Firm, the Market and the Law*. This book contains some of Coase's seminal articles (e.g., Coase, 1937, 1960).

The Firm, the Market, and the Law (*Coase, 1988*)

I have argued elsewhere (Mahoney, 1992c) that there is an isomorphism between the Coase theorem (1960) that in the absence of transaction costs, liability (property rights) rules do not matter for achieving efficient economic outcomes and that in the absence of transaction costs, organizational form (e.g., vertical contracting vs. vertical financial ownership) does not matter for achieving efficient economic outcomes (Coase, 1937). In one article, Coase (1960) shows that if transaction costs were zero, the law would have no purpose in serving economic efficiency, and in the other article, Coase (1937) shows that if transaction costs were zero, the firm would have no purpose in serving economic efficiency. These two articles then may be regarded as stepping stones on the way to an economic analysis of studying the law and organizations with positive transaction costs. Such an approach is well underway, as we shall see from the works of Williamson (1975, 1985, 1996).

Coase (1988) argues that the firm and the market together make up the institutional structure of the economic system. The concept of transaction costs helps explain why the firm exists and which activities the firm will undertake. Transaction costs include search and information costs, bargaining and decision costs, and policing and enforcement costs. Coase posits that firms emerge to organize what would otherwise be market transactions whenever their costs are less than the costs of carrying out transactions through the market.

The Market. Markets are institutions that exist to facilitate exchange; that is, they exist to reduce the costs of carrying out exchange transactions. In an economic theory, which assumes that transaction costs are nonexistent, markets have no function to perform. In practice, all

exchanges regulate in great detail the activities of those who trade in these markets. For anything approaching perfect competition to exist, an intricate system of rules and regulations would normally be needed. Some economists observing the regulations of the exchanges often assume that they represent an attempt to exercise monopoly power and aim to restrain competition. Coase (1988) submits that they ignore or, at any rate, fail to emphasize an alternative explanation for these regulations: that institutions exist to reduce transaction costs and therefore to increase the volume of trade and economic value creation.

The Way Ahead. Coase (1988) maintains that without some knowledge of what would be achieved with alternative institutional arrangements, it is impossible to choose sensibly among them. We therefore need a theoretical system capable of analyzing the effects of changes in these institutional arrangements. To do this analysis, it is not necessary to abandon standard economic theory, but it does mean incorporating transaction costs into the analysis because so much that happens in the economic system is designed to reduce transaction costs. Thus, not to include transaction costs impoverishes the theory.

The Nature of the Firm. Coase (1988) argues that it can be assumed that the distinguishing mark of the firm is the supersession of the price mechanism. Coase points out that while economists treat the price mechanism as a coordinating instrument, they also admit the coordinating function of the entrepreneur, and it is surely important to inquire why coordination is the work of the price system in one case and of the entrepreneur in another case. We have to explain the basis on which, in practice, the organizational governance choice between alternatives is influenced.

Coase (1988) argues that *the main reason why it is profitable to establish a firm would seem to be that there is a cost of using the price mechanism.* The most obvious transaction cost of organizing production through the price system is that of discovering what the relevant prices are. This transaction cost may be reduced but will not be eliminated by the emergence of specialists who will sell this information. The costs of negotiating and concluding a separate contract for each exchange transaction, which takes place on a market, must also be taken into account. In effect, a series of contracts is replaced with a substituted one. The contract is one whereby the employee, for certain remunerations (which may be fixed or fluctuating), agrees to obey the direction of an entrepreneur within certain limits. The essence of the contract is that it should only state the limits to the control rights of the entrepreneur. Within these limits, the entrepreneur can therefore direct the employees.

Coase (1988) then asks the following question: Why isn't all production carried on in one big firm? Coase states that there are at least two important reasons. First, as a firm gets larger, there may be decreasing returns to the entrepreneur function, that is, the costs of organizing additional transactions within the firm may rise. Second, as the transactions that are organized increase, the entrepreneur fails to place the factors of production in the use where their economic value is greatest, that is, fails to make the best use of the factors of production. These two reasons correspond to the phrase of "diminishing returns to management."

The Problem of Social Cost. Coase (1988) is concerned here with another aspect of transaction costs economics. Here, the focus is on those actions of business firms that have harmful effects on others. The standard example is that of a factory, the smoke from which has harmful effects on those occupying neighboring properties.

Coase (1988) insightfully explains that *we are dealing with a problem of a reciprocal nature. To avoid the harm to B would be to inflict harm on A. The real question that has to be answered is, should A be allowed to harm B, or should B be allowed to harm A?* Coase (1988) submits that with costless market transactions, the decision of the courts concerning liability of damage would be without effect on the allocation of resources in terms of economic efficiency.

When dealing with the problem of the rearrangement of legal rights through the market, Coase (1988) argues that a rearrangement would be made through the market whenever this change would lead to an increase in the economic value of production. But this argument assumes costless market transactions. Once the economic costs of carrying out market transactions are taken into account, it is clear that such a rearrangement of property rights will only be undertaken when the increase in the economic value of production consequent on the rearrangement is greater than the economic costs that would be involved in bringing such a change about. In these conditions, the initial delimitation of property rights does have an effect on the efficiency with which the economics system operates. One arrangement of rights may bring about a greater economic value of production than any other arrangement of rights. But unless this is the arrangement of rights established by the legal system, the costs of reaching the same result by altering and combining rights through the market may be so great that this optimal arrangement of rights, and the greater economic value of production that such an arrangement of rights would bring, may never be achieved.

In summary, the same approach that with zero transaction costs demonstrates that the allocation of resources remains the same whatever the

legal position also shows that, with positive transaction costs, the law plays a crucial role in determining how resources are used. As I later show in Chapter 3 on property rights, Coase's (1988) insights have been seminal contributions to not only transaction costs theory but also property rights theory. I now provide a brief application of the Coase theorem and then turn to Williamson's trilogy:

- *Market and Hierarchies* (1975)
- *The Economic Institutions of Capitalism* (1985)
- *The Mechanisms of Governance* (1996)

Application of the Coase Theorem: “The Fable of the Bees”

A prominently discussed case of externalities is the so-called fable of the bees. Beekeepers provide pollination services for the surrounding fruit growers, while the fruit growers, in turn, provide nectar for the bees. Many economists would regard this example to be a classic case of (positive) externalities. If beekeepers and growers do not receive compensation for the benefits that they bestow on the other parties, then the beekeepers and growers will underinvest in these activities (from a social welfare perspective).

The Coase theorem suggests the possibility that beekeepers and growers can privately negotiate with each other, provided the transaction costs are low, and can thus achieve a contractual solution to this (positive) externality problem. Indeed, Cheung (1973) found that this contractual solution is exactly what was done. Beekeepers and growers often enter into contractual relationships. Fruit growers hire hives of bees to provide pollination of those trees that give little suitable nectar, while the beekeeper pays the fruit growers for the privilege of using their bees on high nectar-producing trees. Given these contractual payments, beekeepers and growers have the economic incentives to consider the spillover effects on the other contractual party when they make their investment decisions. Through this market-based process, beekeepers and fruit growers can reach efficient levels of investment with no assistance from centralized (government) coordination.

SOURCE: Adapted from Cheung (1973)

Markets and Hierarchies: Analysis and Antitrust Implications (*Williamson, 1975*)

Williamson (1975) is concerned with the organization of economic activity within and between markets and hierarchies. Market transactions involve exchanges between autonomous economic parties. Hierarchical transactions are transactions for which some form of subordination prevails. Whereas received microeconomic theory generally regards the organization of economic activity between firms and markets as a datum, the study of markets and hierarchies assesses the comparative efficiency properties of alternative governance modes. This theoretical approach is usefully applied to product markets, labor markets, capital markets, and value-chain analysis (Porter, 1985). An organizational failures framework is proposed and employed in an attempt to assess the comparative efficacy of completing related sets of transactions across a market or within a firm.

Following Commons (1934), Williamson (1975) maintains that the transaction is the ultimate unit of microeconomic analysis. Williamson draws on an extensive body of market failure research literature that was in place by 1975 and approaches organizational boundary issues in an interdisciplinary way, where law, property rights theory, micro-economic theory, business history, and organization theory are usefully brought together to achieve a better understanding of the origins, evolution, and functions of various firm and market structures.

Williamson (1975) notes that hierarchy usually implies a superior-subordinate relationship. What is called an employment relationship is commonly associated with voluntary subordination. Williamson points out that the question is not merely whether internal organization can be substituted for the market with beneficial results but also which type of internal organization is to be employed. This second question poses organizational form issues. *Whereas simple hierarchy and vertical integration can be regarded as substitutions of internal organization for failures in the labor and intermediate product markets, respectively, conglomerate organization—of the appropriate multidivisional kind—constitutes the substitution of internal organization for failures in the capital market.*

Williamson (1975) identifies the following themes: (1) while the relation of technology to organization remains important, this relation is scarcely determinative of organizational form (transactional considerations, not technology, are typically decisive in determining which mode of organization will obtain in what circumstances and why); and (2) more self-conscious attention to rudimentary human attributes is

essential if we are to understand more adequately the problems of markets and hierarchies.

Important antecedents of Williamson (1975) are as follows:

- Commons (1924, 1934) made the transaction the ultimate unit of economic analysis and made transfers of legal control and the efficacy of contracting a primary research focus.
- Coase (1937) noted transaction costs are made the centerpiece of an economic analysis of the efficiency of completing transactions as between firms and markets.
- Hayek (1945, 1948) stated the adaptive property of competitive market systems under changing market circumstances is emphasized. Hayek (1945) highlights the marvel of the economic system, where prices serve as sufficient statistics, thereby economizing on bounded rationality. Williamson (1975) argues that, given bounded rationality, uncertainty, and idiosyncratic knowledge, prices often do not qualify as sufficient statistics and that a substitution of internal organization (hierarchy) for market-mediated exchange often occurs on this account.
- Market failures literature included analysis of the following:
 - a. Incomplete markets due to uncertainty
 - b. Insurance problems
 - c. Employment relations
 - d. Vertical integration
 - e. Capital markets
 - f. Increasing returns and sunk costs
 - g. Indivisibilities
 - h. Information asymmetries
 - i. Public goods
 - j. Lack of definition of property rights and externalities with positive transaction costs (see Coase, 1960)

Williamson (1975) notes that the transaction costs approach differs from this earlier economics literature because the transaction costs approach is interdisciplinary, combining economics, law, and organization theory. For example, Williamson combines contingent claims contracting ideas along with organization man (Simon, 1947) and strategic behavior (Schelling, 1960). Second, Williamson is more concerned with tracing the governance structure ramifications of bounded rationality. Third, Williamson introduces the concept of opportunism and the ways that opportunistic behavior is influenced by economic organization. Fourth, Williamson emphasizes that it is not uncertainty or small numbers, individually or together, that occasion market failure, but it is rather the joining of these factors with bounded rationality on the one hand and opportunism on the other hand that gives rise to exchange difficulties. Williamson provides then an organizational failures framework:

- Markets and firms are alternative governance modes for completing a related set of transactions.
- Whether a set of transactions ought to be executed across markets or within a firm depends on the relative efficiency of each governance mode.
- The costs of writing and executing contracts vary with objective properties of the market.
- Although the human and environmental factors that impede exchanges between firms (across a market) manifest themselves somewhat differently within the firm, the same set of factors apply to both. A symmetrical analysis of trading thus requires that we acknowledge the transactional limits of internal organization as well as the sources of market failure.
- The markets and hierarchies approach attempts to identify a set of environmental factors that together with a set of human factors explain the circumstances under which complex contingent claims contracts will be costly to execute and enforce. Faced with such contractual difficulties, and considering the risks that simple (or incomplete) contingent claims contracts pose, the firm may decide to bypass the market and resort to hierarchical modes of organization. Transactions that might otherwise be handled in the market are thus performed internally, governed instead by organizational processes. The environmental factors that lead to prospective market failure are uncertainty and small-numbers exchange relations. Unless joined, however, by a related set of human factors, such environmental conditions need not impede market exchanges. Pairing of uncertainty with bounded rationality and joining small numbers with opportunism are especially important.
- The principle of bounded rationality has been defined by Simon (1957) as follows: "The capacity of the human mind for formulating and solving complex problems is very small compared with the size of problems whose solution is required for objectively rational behavior in the real world" (p. 198). Bounded rationality refers to neurophysiological limits on one hand and language limits on the other. (If, in consideration of these contractual limits, it is very costly or impossible to identify future contingencies and specify *ex ante*, appropriate adaptations, long-term contracts may be replaced by internal organization. Internal organization permits adaptation to uncertainty to be accomplished by organizational processes in a sequential fashion. Rather than attempt to anticipate all possible contingencies from the outset, the future is permitted to unfold. Internal organization, in this way, economizes on the bounded rationality of decision makers in circumstances in which prices are not sufficient statistics and uncertainty is substantial.)
- Opportunism involves self-interest seeking with guile. Opportunistic inclinations pose little risk as long as there are competitive (large-numbers) exchange relations. Many transactions that at the outset involve a large number of qualified bidders are transformed in the process of contract execution so that a small-numbers supply condition obtains at contract renewal. Recurrent short-term contracting is costly and risky when there are opportunistic behaviors and small numbers.

Opportunism may include ex ante adverse selection (hidden information), ex post moral hazard (hidden action), and hold-up problems.

- Bounded rationality and uncertainty pose problems for long-term contracting. Opportunism and small numbers pose problems for short-term contracting. Thus, internal organization may arise because of its relatively greater efficiency.
- Internal organization allows for adaptive, sequential decision making (to economize on bounded rationality), and opportunism is attenuated because internal divisions do not have preemptive claims on profit streams and internal incentives and controls are much more refined than market exchanges. Thus, the firm is better able to take a long-term view for investment purposes (and hence is more prepared to put specialized plant and equipment in place) while simultaneously adjusting to changing market conditions in an adaptive, sequential manner.
- Whichever way the assignment of transactions to firm or market is made initially, the governance choice ought not be regarded as fixed. For example, the degree of uncertainty associated with the transactions in question may diminish, market growth may support large-numbers supply relations, and information asymmetries between the parties often shrink. Also, changes in information-processing technology may occur, which alter the degree to which bounded rationality limits apply, with the result that the previously selected assignment of activities between markets and hierarchies differs from the later, appropriate assignment. Thus, we ought periodically to reassess the efficiency of completing transactions by one governance structure rather than another governance structure.

Williamson (1975) notes that the key idea is a concern with comparative governance assessment. *Organizational failure* is a symmetrical term meant to apply to markets and hierarchies alike. Key concepts are discussed in the following paragraphs.

Bounded rationality refers to human behavior that is intendedly rational but only limitedly so (Simon, 1947). Bounded rationality involves neuro-physiological limits on the one hand and language limits on the other hand. The physical limits take the form of rate and storage limits on the abilities of individuals to receive, store, retrieve, and process information without error. Limited computational capacity prevents comprehensive contracting required for the standard economic welfare theorems on the existence and optimality of a competitive equilibrium (Arrow, 1974). To the extent that internal organization serves to economize on scarce computational capacity and does not experience offsetting disabilities, internal organization is favored. The advantages of internal organization are especially significant in relation to adaptive, sequential decision making.

Williamson (1975) notes that language limits refer to the inability of individuals to articulate their knowledge or feelings by the use of words, numbers, or graphics in ways that enable them to be understood by others.

Despite their best efforts, individuals may find that language fails them (possibly because they do not possess the necessary vocabulary, or the necessary vocabulary has not been devised). Demonstrations and learning by doing may be the only means of achieving understanding when such language difficulties arise. Bounds on rationality are pragmatically interesting only to the extent that the limits of rationality are reached—which is to say, under conditions of uncertainty, complexity, or both.

In the absence of either of these conditions, the appropriate set of contingent actions can be fully specified at the outset. An example is the two-dimensional tic-tac-toe game. The problem of *ex ante* specification of contingent responses for all moves in this game is, in relation to the computational abilities of most adults, relatively simple.

The corresponding chess problem, by contrast, is impossibly complex; *ex ante* specification of the full decision tree is infeasible. Chess, in this sense, is a trivial game: If the theory of chess (i.e., the complete tree of possible games) were fully known, there would be nothing interesting left to play. However, if the average length of a game is 40 moves, there are 10^{120} possibilities to consider. This example is a decision tree that is unimaginably vast. Given unbounded rationality, contingent claims contracting goes through, whatever the degree of complexity to be dealt with. Similarly, given a sufficiently simple environment, bounded rationality constraints are never reached, and comparative governance choices between firm and market are not posed—not in any pragmatically interesting way at least.

Williamson (1975) notes that when transactions are conducted under conditions of uncertainty or complexity, in which it is very costly—perhaps impossible—to describe the complete decision tree, the bounded rationality constraint is binding and an assessment of alternative governance modes, in efficiency respects, becomes necessary. The distinction between deterministic complexity and environmental uncertainty is pragmatically inessential. What may be referred to as uncertainty in chess is uncertainty introduced into a perfectly certain environment by inability—computational inability—to ascertain the structure of the environment. But the result of uncertainty, whatever its source, is the same: Approximation must replace exactness in reaching a decision. As long as uncertainty or complexity is present in requisite degree, the bounded rationality problem arises, and a pragmatically interesting comparative governance structure choice is often posed.

Williamson (1975) observes that internal organization often has attractive properties that permit the parties to deal with uncertainty

and complexity in an adaptive, sequential fashion without incurring the same types of opportunism hazards that market contracting would pose. Such adaptive, sequential decision processes greatly economize on bounded rationality. Rather than specifying the decision tree exhaustively in advance, and deriving the corresponding contingent prices, events are allowed to unfold and attention is restricted to only the actual rather than all possible outcomes. As a Bayesian, one considers the sequential process of successively revising a priori probabilities on the basis of new observations. Thus, you can cross your bridge as you come to it rather than phrase your detailed plan in advance, thereby crossing all possible bridges you might conceivably come to.

Williamson (1975) maintains that a further advantage of internal organization is that, as compared to recurrent market exchange, efficient codes are more apt to evolve and to be employed with confidence by the parties. Such coding also economizes on bounded rationality. Complex events are summarized in an idiosyncratic language. Communication systems become effective when these systems employ languages that carry large amounts of meaning with relatively fewer symbols. Organizations find such things as blueprints and occupational jargon helpful in increasing the efficiency of their communications. Although, in principle, the parties to recurrent market contracts could devise the same language, thereby realizing the same economies, such exchanges are more subject to contractual risks of opportunism—hence, they are less apt to be developed as fully. An additional advantage of internal organization is that internal organization promotes convergent expectations (Malmgren, 1961), serving to attenuate uncertainties that are generated when interdependent parties make independent decisions with respect to changing market circumstances.

Opportunism extends the conventional assumption that economic agents are guided by consideration of self-interest to make allowance for more complex behavior. Opportunism involves self-interest seeking with guile. Opportunistic behavior entails making false or empty threats and promises in the expectation that individual advantage will thereby be realized. Examples include selective or distorted information disclosure and self-disbelieved promises regarding future conduct. Since opportunistic individuals cannot be distinguished *ex ante* from sincere individuals, relying on promises exposes sales contracts, for example, to opportunism hazards during contract execution and at contract renewal. Internal organization may arise because internal organization permits economies to be realized in initial contracting, monitoring respects, or both.

Williamson (1975) cautions, however, that merely to harbor opportunistic inclinations does not imply that markets are flawed on this account. It is furthermore necessary that a small-numbers condition prevail. Absent a small-numbers condition, rivalry among a large number of bidders renders opportunistic inclinations ineffectual. Contractual parties who attempt to secure gains by opportunistic posturing will find, at the contract renewal interval, that such opportunistic behavior is nonviable. When, however, opportunism is joined with a small-numbers condition, the trading situation is transformed, entailing contractual problems of bilateral monopoly.

Williamson (1975) poses the following transactional dilemma: It is in the economic interest of contractual parties to seek terms most favorable to each of them, which encourages opportunistic representations and haggling. The interests of the system, by contrast, are promoted if the parties can be joined to avoid both the bargaining costs and the indirect costs (mainly maladaptation costs) that are generated in the process. What is of special interest to transaction costs analysis is that while frequently a large-numbers condition appears at the outset, this appearance may be illusory or may not continue into contract renewal stages. If parity among suppliers is changed by first-mover advantages, so that winners of original bids subsequently enjoy non-trivial cost advantages over nonwinners, the sales relationship eventually is effectively of the small-numbers variety. This transformation has relevance not only for decisions regarding internally made rather than purchased separable components but also in cases when exchange of workflow between successive individuals is under an employment rather than a sales relationship.

Williamson (1975) argues that internal organization enjoys economic advantages of at least three kinds over market modes of contracting in circumstances where opportunism and small-numbers conditions are joined:

- In relation to autonomous contractors, the parties to an internal exchange are less able to appropriate subgroup gains, at the expense of the overall organization (system), as a result of opportunistic representations. Preemptive claims on profits between separate firms are eliminated. Thus, the economic incentives to behave opportunistically are attenuated. Also, the management of the trading divisions is more susceptible to appeals for cooperation. Because the aggressive pursuit of individual interests is to the disadvantage of the system, and because the general office can vary present and prospective compensation (including

promotions) to reflect noncooperation, requests to adopt a cooperative mode are apt to be heeded. Altogether, more nearly joint profit maximizing attitudes and results are to be expected.

- Internal organization can be more effectively audited. The auditing advantage of internal organization is attributable to constitutional and incentive differences that operate in favor of the internal governance mode. External auditors are typically constrained to review written records and documents and in other respects to restrict the scope of their investigation to clearly pertinent matters. Internal auditors, by contrast, have greater freedom of action, both to include less formal evidence and to explore the byways into which their investigations lead. The differential improvement of auditing by merged railroad firms relative to auditing by railroad cartels is illustrative (Chandler, 1977).
- When differences do arise, internal organization has an advantage over market-mediated exchange in dispute-settling respects. The firm can settle many disputes by fiat and quasi-judicial functions.

Williamson (1975) notes that contractual problems are often posed by information asymmetries and opportunism (e.g., adverse selection and moral hazard). Williamson emphasizes that information problems can develop even when contractual parties have identical information under conditions where third-party enforcers will have trouble obtaining truthful information. This agency problem is avoided by one of two ways: self-enforcing agreements (e.g., Klein & Leffler, 1981; Telser, 1980) or courts' perfect enforcement of contracts. Relative to the agency research literature, transaction costs theory is more self-conscious about imperfect enforcement of contracts.

In summary, Williamson (1975) argues that the advantages of internal organization in relation to markets are as follows (see Figure 2.1 below):

1. In circumstances where complex, contingent claims contracts are infeasible and sequential spot markets are hazardous, internal organization facilitates adaptive, sequential decision making that economizes on bounded rationality.
2. Faced with present or prospective small-numbers exchange relations, internal organization attenuates opportunism.
3. Convergent expectations that reduce uncertainty are promoted.
4. Conditions of asymmetric information (information impactedness) are more easily overcome and, when conditions of asymmetric information appear, are less likely to give rise to opportunistic behavior.
5. A more satisfying trading atmosphere sometimes develops internally.

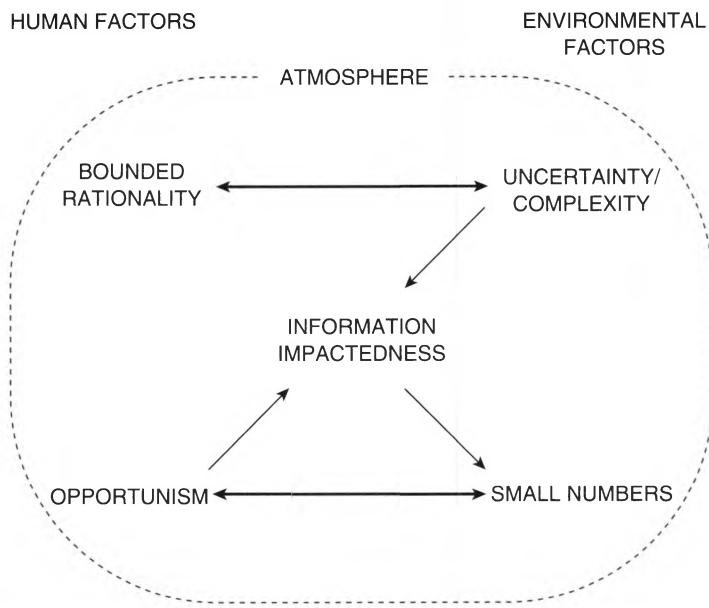


Figure 2.1 The Organizational Failures Framework

SOURCE: Williamson (1975)

Simple Hierarchy. Williamson (1975) notes that Alchian and Demsetz (1972) rely on technological nonseparabilities as the leading reason to explain the origin of firms. Alchian and Demsetz use the manual freight-loading example to illustrate worker nonseparabilities. Because of interaction effects between workers, the marginal productivity of each worker cannot be determined by observing the total weight loaded during the day. In team production, market mediation is replaced by internal organization where a boss monitors the performance of the team and allocates rewards among workers on the basis of observed input behavior. Shirking is purportedly attenuated in this way. Williamson argues that nonseparabilities may explain monitoring in some circumstances, but nonseparabilities hardly qualify as an explanation for complex organizations. Also missing from Alchian and Demsetz's discussion is the impact of workers' attitudes to their work as the level of monitoring increases (e.g., as discussed by the human relations school of management in the 1940s).

Understanding the Employment Relation. Williamson (1975) observes that an important form of nonhomogeneity—job idiosyncrasy—is a key

concept for understanding the employment relation. Workers acquire, during the course of their employment, significant firm-specific capabilities, job-specific skills, and related task-specific knowledge.

Williamson (1975) notes that Alchian and Demsetz (1972) take the intellectual position that it is a delusion to characterize the relationships between employers and employees by reference to fiat, authority, or the like. Rather, it is Alchian and Demsetz's (1972) contention that the relationships between employers and employees are identical to that which exists between shoppers and their grocers in fiat and authority respects:

The single consumer can assign his grocer to the task of obtaining whatever the customer can induce the grocer to provide at a price acceptable to both parties. That is precisely all that an employer can do to an employee. To speak of managing, directing, or assigning workers to various tasks is a deceptive way of noting that the employer continually is involved in renegotiation of contracts on terms that must be acceptable to both parties. Long-term contracts between employer and employee are not the essence of the organization we call a firm. (p. 777)

Williamson (1975) points out that implicit in Alchian and Demsetz's (1972) argument is an assumption that the transition costs associated with employee turnover are negligible. Employers, therefore, adapt easily to changing market circumstances by filling jobs on a spot market basis. Although job incumbents may continue to hold jobs for a considerable period of time and may claim to be subject to an authority relationship, all that job incumbents are essentially doing is continuously meeting bids for their jobs in the spot market. Williamson argues, however, that where tasks are idiosyncratic, in nontrivial degree, the worker-employer relationship is no longer contractually equivalent to the grocer-customer relationship, and the feasibility of spot market contracting breaks down. Alchian and Demsetz's frictionless fiction blurs and neglects essential (incomplete contracting) elements of the organizational economics problem.

Vertical Integration. Williamson (1975) notes that earlier industrial organization research literature on vertical integration emphasized technological interdependency (or, in modern terms, economies of scope) as an important rationale for vertical financial ownership. The standard example is the integration of iron and steel making, where thermal economies are said to be available through vertical integration (Bain, 1968). Williamson (1971, 1975) insightfully (and originally) argues that were it possible to write and enforce a complex contingent claims contract between blast furnace and rolling mill stages, the vertical integration of these activities for thermal economy reasons would be

unnecessary. The prohibitive transaction costs of such contracting are what explain the decision to integrate vertically.

Limits of Vertical Integration and Firm Size. Williamson (1975) also analyzes sources of efficiency distortion in internal organization:

- *Internal procurement bias.* Subgoals of a group are easily given greater weight in relation to objective profitability considerations. A norm of reciprocity easily develops.
- *Internal expansion bias.* This bias is partly attributable to its dispute-settling characteristics. A common method of dealing with internal system conflict is to adopt a compromise solution where concessions are made to subsystems rather than require these subsystems to give up essential functions or resources. This size-preserving tendency is favored by the positive association of both pecuniary and nonpecuniary rewards with size, at least among the functional parts of the organization.
- *Persistence.* If the organizational system has committed itself in advance to the correctness and efficacy of its reform, the organizational system will not tolerate learning of failure.
- *Communication distortion.* Communication distortion can include serial reproduction loss (a bounded rationality problem) and deliberate distortion (an opportunistic behavior problem).

The Multidivisional Structure. Williamson (1975) reminds us that Berle and Means (1932, p. 121) noted that a separation of ownership and control existed and inquired “have we any justification for assuming that those in control of a modern corporation will also choose to operate it in the interests of the stockholders?” Williamson’s response is that not only have external controls been put in place (e.g., by the market for corporate control) but also internal controls have developed (and, as outlined by Chandler [1962], multidivisional organizations that emerged in the 1920s at General Motors, DuPont, Standard Oil, and Sears, among others, served as internal controls). In particular, the multidivisional (M-form) structure followed the corporate strategy of diversification (Chandler, 1962). Williamson emphasizes the following characteristics of the M-form organizational innovation:

- Responsibilities for operating decisions are assigned to (essentially self-contained) operating divisions, or quasi-firms.

- The staff attached to the general office performs both advisory and auditing functions. Both functions have the effect of securing greater coordination and control over operating division behavior.
- The general office is mainly concerned with strategic decisions, involving planning, appraisal, and control, including the allocation of resources among the (competing) operating divisions.
- The separation of the general office from operations provides general-office executives with the psychological commitment to be concerned with the overall performance of the organization rather than to become absorbed in the tactical decisions of the functional parts.
- The resulting M-form structure displays both rationality and synergy: The whole is greater (i.e., more effective and more efficient) than the sum of the parts.

Williamson (1975) argues that in relation to the functional form organization of the same activities, the M-form organization of the large, complex organization served both to economize on bounded rationality and to attenuate opportunism. Operating decisions were no longer forced to the top of the hierarchy but were resolved at the divisional level, which relieved the communication load.

Williamson (1975) notes that imitation of the M-form organizational innovation was at first rather slow. However obvious its superior properties may have been to the organizational innovators, others were naturally skeptical. Organizational innovation may more likely lead to sustainable, competitive advantage than to product innovation. Organizational innovation may be more complex, not easily articulable, and (for a time) tacit.

Williamson (1975) takes issue with the frictionless fiction of perfectly functioning capital markets. Traditional capital market controls are relatively crude because of asymmetric information with respect to internal conditions in the firms, and, because of sorting out difficulties, the risk of opportunism on the part of would-be takeover agents is great. Given asymmetric information, outsiders can usually make confident judgments that the firm is not adhering to profit-maximizing standards only at great expense. The large firm is a complex organization, and its economic performance is jointly a function of exogenous economic events, rival behavior, and internal decisions. Causal inferences are correspondingly difficult to make, and, hence, opportunism is costly to detect.

Williamson (1975) argues that in a general sense, the most severe limitation of the capital market is that the capital market is an external control mechanism. The capital market has limited constitutional powers to conduct audits, and the capital market has limited access to the firm's incentive and resource allocation fine-tuning mechanisms. The multidivisional firm can serve as a miniature capital market and can perform the following functions:

- *Manipulate incentives.* Salaries and bonuses can be adjusted to reflect differential operating performance.
- *Perform internal audits.* There are at least two advantages of the general office over the capital market in auditing respects. First, division managers are subordinates, and both their accounting records and backup files are appropriate subjects for review. Stockholders, by contrast, are more limited in what they can demand in the way of disclosure. Second, the general office can expect knowledgeable parties to be more cooperative than can an outsider, and such cooperation is often rewarded accordingly.
- *Provide cash flow allocation.* The multidivisional (see Figure 2.2) may be thought of as an internal competition among divisions where the general staff can perform a capital market function—assigning cash flows to high-yield uses. In many respects, this assignment of cash flows to high-yield uses is the most fundamental attribute of the M-form organization in the comparison of internal with external processes. What the multidivisional lacks in breadth of knowledge it compensates for in its in-depth knowledge of its resources and capabilities. Effective performance of multidivisional firms with heterogeneous resources is a result not necessarily of having better resources but of knowing more accurately the relative productive performances of those resources under different business contexts.

Williamson (1975) describes characteristics of effective divisionalization:

- Identifying separable economic activities within the firm
- According quasi-autonomous standing (usually of profit-center nature) to each division
- Monitoring the efficiency performance of each division
- Awarding incentives
- Allocating cash flows to high-yield uses
- Performing strategic planning (e.g., diversification, acquisition, and related activities)

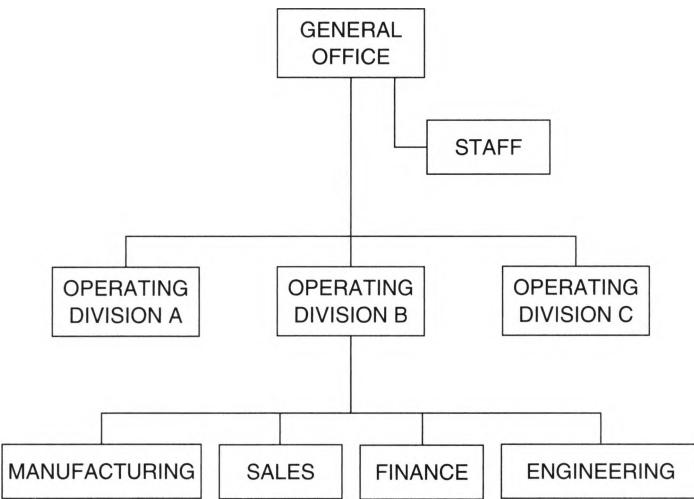


Figure 2.2 Multidivisional Form

SOURCE: Williamson (1975, p. 138)

Williamson (1975) emphasizes that care must be exercised lest the general management and its staff become overinvolved in operating decisions and fail to perform the high-level planning and control functions on which the M-form organization relies for its continuing success. Otherwise, Gresham's law of planning may occur where daily routine drives out strategic planning.

Williamson (1975) observes that although evolutionary change is to be expected, the hierarchical decomposition principles on which the M-form is based are robust. The concept of hierarchy is given in cybernetics as a necessary structural attribute of any viable organism. All viable systems do in fact exhibit hierarchical organizations. Moreover, not only does Simon's (1962) review of complex biological, physical, and social systems reaffirm this fact, but Simon emphasizes that hierarchies commonly factor problems in such a way that higher frequency dynamics are associated with the subsystems, the lower frequency dynamics are associated with the larger systems, and intracomponent linkages are generally stronger than intercomponent linkages. Hierarchical systems of this sort may be referred to as nearly decomposable. It is not merely fortuitous that the M-form structure factors problems very much in this way. In organization theory, these ideas are usually discussed under the heading of "loosely coupled systems." The

importance of subsystems can hardly be overemphasized in an analysis of organizational design.

Williamson (1975) provides the M-form hypothesis: The organization and operation of the large firm along the lines of the M-form favor goal pursuit and least-cost behavior more nearly associated with the neoclassical profit maximization hypothesis than does the functional form organizational alternative. The basic argument comes down to this: Just as the vertical integration of production is to be explained in large measure by reference to (comparative governance) failures in the market for intermediate goods, the affirmative aspects of multidivisional organization are to be understood in terms of (comparative governance) failures in the capital market. Williamson also points out that pressures due to the market for corporate control induced the more efficient multidivisional form. Once this multidivisional form had been extensively established in the 1960s, this organizational structure allowed for greater unrelated diversification as well as a more active takeover market.¹

Williamson (1975) concludes by highlighting the distinctive features of the approach in *Markets and Hierarchies*:

1. The approach makes evident that the characteristics of transactions rather than technology ultimately influence governance choice.
2. A comparative institutional analysis is maintained; markets and hierarchies are regarded as alternative governance modes.
3. The approach makes provisions for rudimentary attributes of human nature (bounded rationality and opportunism) and relates these attributes to a set of environmental factors (complexity, uncertainty, and small numbers) in the context of an organizational failures framework.
4. The approach is much more microanalytic than previous organizational treatments, focusing as it does on the transactional details of recurrent contracting under alternative modes of organization.
5. However useful the fiction of frictionless exchange is for some purposes, it is an impediment to understanding the efficiency properties of alternative modes of economic organization. The frictionless fiction is accordingly abandoned.

¹For empirical research in organization theory and strategic management on the multidivisional form see, for example, Armour and Teece (1978); Mahoney (1992a); Palmer, Friedland, Jennings, and Powers (1987); and Rumelt (1974). For example, Mahoney (1992a) finds strong empirical evidence that diversification (as measured by the Berry index) and geographic dispersion (as measured by the log of the number of states/cities in which the enterprise had plants) each induce the adoption of the multidivisional organizational form.

6. Organizational form, which is concerned with the decomposition principles of hierarchy, is introduced as an internal organizational counterpart to the familiar market structure measures of industrial organization for explaining economic performance.
7. New questions, or different perspectives on old questions, are afforded across a wide range of strategic management issues—including the employment relationship, vertical integration, and multidivisional organization.
8. The approach is comparatively value-free, in terms of the governance choice of markets and hierarchies—the approach is biased neither for nor against unfettered market modes of governance.
9. Supplying a satisfying exchange relation is made part of the economic problem by introducing the concept of atmosphere. Attitudinal interaction effects are brought to the foreground. Transactions that affect self-esteem or perceptions of collective well-being, or both, are those for which attitudinal considerations are especially important.
10. The employment relationship is the leading instance where the influence of metering intensity on work attitudes needs to be assessed with care. More specifically, efforts to divide the employment relation into parts and assess each separately in strictly calculative, instrumental terms can have, for some individuals at least, counterproductive consequences.

The Economic Institutions of Capitalism (*Williamson, 1985*)

Now that we have completed our studies of *Markets and Hierarchies* (Williamson, 1975), we next consider *The Economic Institutions of Capitalism* (Williamson, 1985), where the concept of asset specificity receives greater attention both theoretically and empirically. Indeed, transaction costs theory emphasizes the condition of asset specificity. The existence of durable, firm-specific assets is held to be widespread, and the choice of governance structure is held to be economically important.

Williamson (1985) dedicates his research book to four teachers: Kenneth Arrow, Alfred Chandler, Ronald Coase, and Herbert Simon. From Arrow (1974), Williamson learned the importance of information and not to shoehorn difficult problems into orthodox boxes. Chandler (1962, 1977) taught Williamson that organizational innovation was an important and often neglected phenomenon that had pervasive ramifications for understanding American industry. Coase (1937, 1960) taught Williamson that transaction costs were central to the study of economic organization and that such studies should be performed in a comparative institutional manner. Simon (1947, 1957) taught

Williamson that behavioral assumptions are important and not to be intimidated by disciplinary boundaries.

Williamson (1985) notes that, like transaction costs research literature, the recent principal-agent research literature is similarly oriented to the study of contract. But there are real differences as well. The principal-agent literature focuses on the *ex ante* (or economic incentive alignment) side of contract, assumes that contractual disputes are routinely referred to courts, and assumes that courts dispense justice effectively (indeed, costlessly). In contrast, transaction costs theory maintains that the governance of contractual relations is primarily achieved through private ordering rather than through legal centralism. Although the importance of *ex ante* incentive alignment is acknowledged, primary attention is focused on the *ex post* institutions of contract. A compact statement of the economic problem of organization is to devise contract and governance structures that have the purpose and effect of economizing on bounded rationality while simultaneously safeguarding transactions against the economic hazards of opportunism.

Williamson (1985) argues that understanding the economic institutions of capitalism poses deep and enduring challenges to law, economics, and organization. Transaction costs theory is, by design, an interdisciplinary undertaking. According to Williamson (1985), the transaction costs approach maintains that the institutions of capitalism have the main purpose and effect of economizing on transaction costs.

A transaction occurs when a good or service is transferred across a technologically separable interface. Transaction costs are analogous to frictions in mechanical systems. Transaction costs analysis requires a comparative institutional analysis among imperfect organizational alternatives. In particular, transaction costs analysis involves the comparative costs of planning, adapting, and monitoring task completion under alternative governance structures. Furthermore, an important task of effective governance is to reduce opportunistic behavior. The motivation and meaning for those studying deeply the economics of organization is that business practice should not be organized to the advantage of the opportunistic against those who are more inclined to keep their promises.

Williamson (1985) emphasizes that the concept of the fundamental transformation—a large-numbers condition at the outset (*ex ante* competition) transforming into a small-numbers condition during contract execution and at contract renewal (*ex post* competition)—has a pervasive importance for the study of the economics of organization.

Rather than characterize the firm as a production function, transaction costs theory maintains that the firm is more usefully regarded as a governance structure. Nonstandard contracting—customer and territorial restrictions, tie-ins, block booking, and related vertical coordination mechanisms—have often been presumed to have an anticompetitive purpose and effect. The transaction costs approach discloses that this formulation is simplistic: Many nonstandard or unfamiliar contracting practices serve legitimate transaction cost economizing purposes. Often the parties are engaged in an effort to devise contractual safeguards that promote more efficient exchange. Commercial equivalents of hostages (i.e., mutual sunk cost commitments) arise in this way.

Transaction Cost Economics. Williamson (1985) argues that firms, markets, and relational contracting (e.g., franchising²; Macneil, 1980) are important economic institutions. These economic institutions are also the evolutionary product of a fascinating series of organizational innovations. Williamson advances the proposition that the economic institutions of capitalism have the main purpose and effect of economizing on transaction costs. Compared to other economic approaches to the study of economic organization, transaction costs theory has the following characteristics:

- Is more microanalytic
- Is more self-conscious about behavioral assumptions
- Introduces and develops the economic importance of asset specificity
- Relies more on comparative institutional analysis
- Regards the business firm as a governance structure rather than a production function
- Places greater weight on the ex post institutions of contract, with special emphasis on private ordering (as compared with court ordering)

Williamson (1985) notes that ex ante and ex post transaction costs are usefully distinguished: Ex ante transaction costs include drafting, negotiating, and safeguarding an agreement. Safeguards can take many forms, the most obvious of which is common ownership. Ex post transaction costs include maladaptation costs, haggling costs, running

²Key theoretical and empirical papers in the economics and strategic management research literature on franchising include Carney and Gedajlovic (1991), Caves and Murphy (1976), Lafontaine and Shaw (1999), Michael (2000), Norton (1988), Rubin (1978), Shane (1996), and Sorenson and Sorenson (2001).

of governance structures where disputes are referred, and the economic bonding costs of effecting secure commitments. Note that ex post transaction costs are related to the agency costs of Jensen and Meckling (1976): monitoring expenditures of the principal; bonding expenditures of the agent; and the residual loss. The ex ante transaction costs and the ex post transaction costs of contracting are interdependent. These transaction costs must be addressed simultaneously rather than sequentially.

In some sense, there are three economic problems posed in Williamson (1985):

1. Hold technology (and production costs) constant and minimize transaction costs
2. Minimize the sum of production costs and transaction costs
3. Consider the problem where asset specificity also enters the revenue function, while taking account of production and transaction costs

For the most part, we focus here on the second problem of minimizing the sum of production and transaction costs.

Williamson (1985, p. 24) provides a tree structure for a cognitive map of contract: For the efficiency branch of organizational economics, there is the economic incentive branch and the transaction costs branch. For the economic incentive branch, there is property rights theory (see Chapter 3) and agency theory (see Chapter 4). The economic incentive research literature looks mostly at the ex ante side of contracts:

- a. Property rights literature emphasizes that ownership matters, where the rights of ownership of an asset take three parts: the right to use an asset; the right to appropriate returns from an asset; and the right to change the form or substance, or both, of an asset. The economic problem for resource allocation becomes one of getting the property rights correct.
- b. Agency literature emphasizes that principals contract in full awareness of the economic hazards that contract execution by agents poses. For example, although the separation of ownership and control attenuates profit incentives, it is anticipated at the time separation occurs and is fully reflected in the price of new shares (Jensen & Meckling, 1976). The structure of the economic problem holds no surprises; all of the relevant contracting action is packed into ex ante incentive alignments.

Positive agency theory emphasizes the minimization of monitoring and bonding costs and asserts that natural selection processes are reliably efficacious.

Principal-agent theory is a relatively mathematical literature that features *ex ante* incentive alignments in superlative degree (Holmstrom, 1979). This line of research is akin to the earlier contingent claims contracting research literature but moves beyond it by admitting contracting complications in the form of asymmetric (private) information. Complex economic problems of incentive alignment are posed (which the contingent claims contracting literature had ignored) if full and candid disclosure of asymmetric information cannot be assumed. In other respects, however, the principal-agent research literature and contingent claims contracting are similar: Both resolve all the relevant contracting issues in a comprehensive *ex ante* bargain, and both assume that court ordering is efficacious. Economic efficiency drives the argument.

The transaction costs branch also maintains the refutable presumption that nonstandard forms of contracting have economic efficiency purposes. Greater attention is shifted, however, to implementation at the contract execution stage.

Governance. The governance branch is the major emphasis of Williamson's (1985) work. In common with the property rights research literature, transaction costs theory agrees that ownership does matter. In common with agency literature, transaction costs theory economizing acknowledges that *ex ante* incentive alignment matters. But whereas the property rights and principal-agent approaches work within the tradition of legal centralism, transaction costs theory disputes that court ordering is efficacious. Attention is shifted instead to private ordering. Governance is especially important when there is bounded rationality coupled with uncertainty and opportunism coupled with asset specificity.

Measurement. First, there can be difficulty of measuring input. Low task programmability reduces the effectiveness of monitoring workers. Second, there can be difficulty of measuring output. For example, Alchian and Demsetz (1972) highlight the high nonseparabilities problem of team production.

Williamson (1985) notes that asset specificity is a key contractual dimension. An awareness of asset specificity can be traced at least to Marshall's (1920) discussion of quasi-rents. Similarly, Marschak (1968) suggested that there exist almost unique, irreplaceable research workers, teachers, and (firm-specific) personnel, just as there exist unique choice locations for plants and harbors.

Asset specificity refers to durable investments that are undertaken in support of particular transactions. The redeployability of such

investments is much lower in best alternative uses or by alternative users should the original transaction be prematurely terminated, and the specific identity of the parties to a transaction matters in these circumstances, which is to say that continuity of the relationship is valued. Thus, contractual and organizational safeguards arise in support of such transactions. Four types of asset specificity are identified in Williamson (1985):

- Site specificity (e.g., an electric plant, a coal mine)
- Physical asset specificity (e.g., specialized tools)
- Human asset specificity (e.g., firm-specific knowledge)
- Dedicated assets (e.g., body dies produced by Fisher Body for GM cars)

Williamson (1985) also classifies uncertainty into three categories:

- Primary (uncertainty about the state of nature)
- Secondary (lack of effective communication)
- Behavioral (opportunism)

Williamson (1985) argues that, with frequent transactions, the interaction effects between uncertainty and asset specificity are important in an understanding of economic organization, and empirical analysis of transaction cost features is complicated as a result (see Figure 2.3).

*Vertical Integration.*³ Williamson (1985) argues that the main economic differences between market and internal organization are the following:

³For mathematical models of vertical integration see Balakrishnan and Wernerfelt (1986), Blair and Kaserman (1983), Riordan and Williamson (1985), and Tirole (1988). For conceptual papers on vertical integration, see Harrigan (1984); Klein, Crawford, and Alchian (1978); Mahoney (1992b); Phillips and Mahoney (1985); Teece (1980); and Williamson (1971, 1979). For example, Mahoney (1992b) shows that in the absence of transaction costs, both vertical contacting (e.g., tying contacts, resale price maintenance, exclusive territories, exclusive dealing contacts) and vertical financial ownership would be equivalent for achieving economic value creation (via lower costs, higher revenues, or reduction of risks in ways that cannot be duplicated by shareholders). However, in a business world of positive transaction costs, a combination of high task programmability (Eisenhardt, 1985), high nonseparability (Alchian & Demsetz, 1972) and high asset specificity (Williamson, 1985) is expected to lead to vertical financial ownership because of its comparative governance efficiency advantages. For doctoral students studying the economics of organization, empirical studies that measure all three constructs for explaining and predicting governance structures appear warranted.

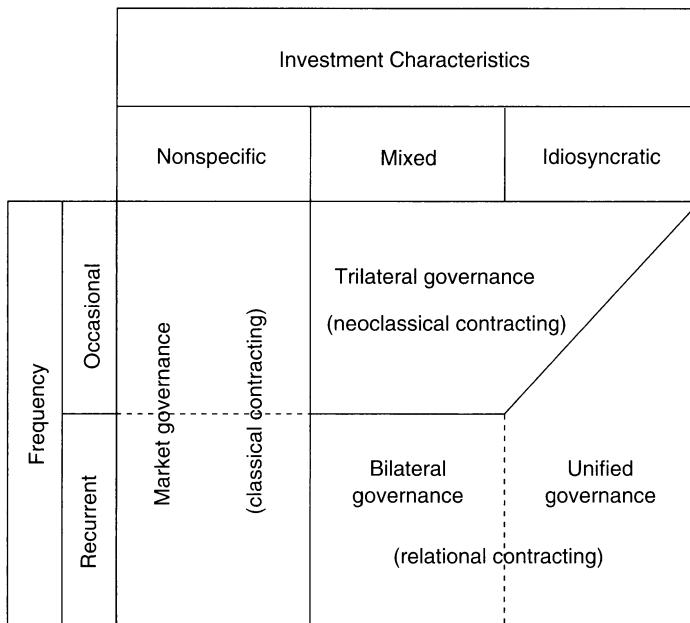


Figure 2.3 Efficient Governance

SOURCE: Williamson (1985, p. 79)

- Markets promote high-powered economic incentives and restrain bureaucratic distortion more effectively than internal organization.
- Markets can sometimes aggregate demands to realize economies of scale and scope. Therefore, market procurement has advantages in both scale and governance respects where optimal asset specificity is slight.
- Internal organization has access to distinctive governance mechanisms and enjoys the advantage where asset specificity is substantial.

Williamson (1985) and Williamson and Masten (1999) note that much of the empirical evidence in the research literature that tests transaction costs theory is corroborative:

- Statistical models using field data include, for example, Anderson, 1985; Anderson and Schmittlein, 1984; Bensaou and Anderson, 1999; Coles and Hesterly, 1998; Heide and Miner, 1992; John and Weitz, 1988; Lyons, 1994; Monteverde and Teece, 1982; Poppo and Zenger, 1998, 2002; Subramani and Venkatraman, 2003; Walker and Poppo, 1991; and Walker and Weber, 1984, 1987.

- Statistical models using secondary data include, for example, Balakrishnan and Wernerfelt, 1986; Caves and Bradbury, 1988; Leiblein and Miller, 2003; Leiblein, Reuer, and Dalsace, 2002; and Levy, 1985.
- Focused case studies and studies of contracts include, for example, Argyres, 1996; Crocker and Masten, 1988; Crocker and Reynolds, 1993; Goldberg and Erickson, 1987; Hennart, 1988b; Joskow, 1987; Mahoney and Crank, 1995; Masten and Crocker, 1985; Palay, 1984, 1985; Pirrong, 1993; and Williamson, 1985.
- Focused industry studies include, for example, Afuah, 2001; Dyer, 1996, 1997; Lieberman, 1991; Masten, 1984; Masten, Meehan, and Snyder, 1989, 1991; Monteverde, 1995; Nickerson, Hamilton, and Wada, 2001; Novak and Eppinger, 2001; Pisano, 1989; Richardson, 1993; and Stuckey, 1983.
- Business history includes, for example, Argyres and Liebeskind, 1999, and Chandler, 1962, 1977.

Business History. Consumer durables requiring information aids, credit, and follow-on service, and producer durables requiring the same were likely candidates for forward integration (e.g., Eastman Kodak, McCormick Harvester, National Cash Register, and Singer Sewing Machine; Chandler, 1977). While the governance branch emphasizes the role of asset specificity for explaining this pattern, the measurement branch emphasizes (positive and negative) externalities. Externality concerns arise in conjunction with a branded good or service that is subject to quality debasement. Whereas a manufacturer can inspect, and thereby better control, the quality of components and materials it purchases from earlier stage and lateral suppliers, it is less easy to exercise continuing quality controls over items sold to distributors. If quality enhancement (debasement) efforts of distributors give rise to positive (negative) interaction effects, in which the benefits (costs) are incompletely appropriated by (assigned to) the originators, then failure to extend quality controls over distribution results in suboptimization (e.g., free-rider problems in a franchise system).

The Limits of Firms: Incentive and Bureaucratic Features. Williamson (1985) asks the following questions: Why can't a large firm do everything that a collection of small firms can do and more? What is responsible for limitations in firm size? Why not organize everything in one large firm? Williamson provides two main reasons:

- Economies of scale and scope may be sacrificed if the firm attempts to make for itself what it can procure in the market.
- Governance costs are increased.

Williamson (1985) points out that the first reason is not a thoroughly comparative explanation. If economies of scale are reached by the outside supplier, then the same economies of scale can be preserved upon merger by instructing the supplier to service the market in the future just as in the past. Thus, the fundamental limitation to firm size must turn on the governance costs disabilities of internal organization where asset specificity is insubstantial. These governance costs disabilities include the following:

- Diminishing returns to management (attention as a scarce resource)
- Control loss (serial reproduction loss; opportunism)
- Susceptibility to logrolling and subject to politicization
- Internal procurement biases
- Bureaucratic delays in large firms
- Impairment of economic incentives

Thus, Williamson (1985) holds that selective intervention, whereby organizational integration realizes adaptive gains but experiences no losses, is not feasible.

Credible Commitments: Unilateral and Bilateral Applications. Williamson (1985) notes that credible (sunk-cost) commitments and credible (sunk-cost) threats share this common attribute: Both appear mainly in conjunction with irreversible, specialized investments (i.e., sunk costs). But whereas credible commitments are undertaken in support of joint ventures (Hennart, 1982, 1988a) and strategic alliances (Hennart, 1993; Koja & Prescott, 2002; Robertson & Gatignon, 1998) to promote exchange, credible threats (e.g., excess capacity as a barrier to entry) appear in the context of conflict and rivalry (Sutton, 1992).

Williamson (1985) reminds us that Schelling (1960) also noted that the exchange of human hostages served incentive purposes in an earlier age. Economic hostages, in our business context, involve asset specificity (i.e., sunk costs). *Economic hostages are an important component of self-enforcing agreements, having both ex ante (screening) and ex post (bonding) effects, and Williamson (1985) here focuses on the ex post bonding effects.*

For example, franchisers can better ensure quality by requiring franchisee investments in specific assets that at termination imply a capital loss economic penalty larger than can be obtained by the franchisee by cheating. For example, the franchiser may require franchisees to rent (rather than own) the land on which their outlet is located. This lease

arrangement creates a situation where termination can require the franchisees to move and thereby impose a capital loss on the franchisees up to the amount of their initial unsalvageable investment. Hence, a form of economic collateral to deter franchisee cheating is created. The arrangement is tantamount to the creation of economic hostages to restore integrity to an exchange. The use of economic hostages deters franchisees from exploiting externalities and is often regarded as an imposed (top down) solution. Franchisees are powerless; they accept economic hostage terms because no other contracting terms are available. Williamson (1985) counsels that such power arguments are often based on *ex post* reasoning. The use of economic hostages can be, and often is, an efficient systems solution and hence is independent of who originates the proposal. It is the franchise system that benefits from the control of externalities. Indeed, if franchisees recognize that the demand externality exists from the outset, if the franchiser refuses to make provision for the externality in the original contract, and if it is very costly to reform the franchise system once initial contracts are set, franchisees will bid less for the right to a territory than they otherwise would. It should not therefore be concluded that perceptive franchisers, who recognize the demand externality in advance and make a provision for it, are imposing objectionable *ex ante* economic terms on unwilling franchisees. Franchisers are taking steps to realize the full economic value of the franchise system. Here, as elsewhere, contracts must be examined in their entirety.

Critical Commentary on Agency Theory. Williamson (1985) points out that enthusiasts of laissez-faire capitalism are loathe to confront, and are sometimes schizophrenic on the subject of, managerial discretion. Focusing on any given time, these enthusiasts commonly deny the existence of managerial discretion. Comparing current practices with the past, however, those same enthusiasts point with pride to the development of new governance mechanisms that brought managerial discretion under more effective control. Williamson (1985) submits that the problem of managerial discretion, due to the separation of ownership and control, is alive and well. Or put differently, agency problems are pervasive. The continuing tension between management and stockholder interests is evident in the numerous efforts that incumbent managements have taken to protect target firms against takeover.

Williamson (1985) concludes that transaction costs theory assumes risk neutrality and investigates organizational form as a means for minimizing transaction costs. Williamson (1985) notes that organizational forms (e.g., franchising) may have risk-sharing purposes as well.

Williamson (1985) maintains that formal mathematical models of verbal arguments that lose in the translation are scarcely to be counted as gains. Formalization is not wanted at any cost. Sometimes, however, effort at formalization discloses gaps or ambiguities that the verbal arguments did not. Anomalies and contradictions can, and should, push those who employ transaction costs analysis to develop better models.

Williamson (1985) argues that, even holding technology constant, at least three things happen when a transaction is transferred out of the market and is placed under unified ownership:

- Ownership changes (property rights change)
- Incentives change
- Governance structures (ability to monitor and reward) change

Williamson (1985) maintains, “‘Flawed’ modes of economic organization for which no superior feasible mode can be described are, until something better comes along, winners nonetheless” (p. 408).

I conclude this chapter with an application of transaction costs theory to vertical integration in the aerospace industry (Masten, 1984), followed by *The Mechanisms of Governance* (Williamson, 1996). Here, Williamson further contributes to the evolving science of organization. Indeed, we learn that one answer to why we observe so many kinds of organization is that contractual hazards come in many forms, for which nuanced governance structures are devised and chosen or selected.

Application: Vertical Integration in the Aerospace Industry

Masten (1984) studied the make-versus-buy decisions for nearly 2,000 components of a major aerospace contractor. The firm made many products for the United States government. The company had to choose between making each component or subcontracting the component for production by another firm. Transaction costs theory suggests that internal production is more likely when the assets are specific and the uncertainties in contracting are large.

Masten (1984) used two measures of asset specificity for each component. The first measured design (physical asset) specificity, that is,

(Continued)

(Continued)

the extent to which the component was used exclusively by the company or could be easily adapted for use by other aerospace firms or by firms in other industries. A transistor would be an example of a nonspecific item; a circuit board designed to individual specifications would be an example of a component with high asset specificity. The second measure of asset specificity was site specificity.

Masten (1984) also measured the complexity of the product design, which was intended to proxy for uncertainties in contracting. Consistent with the theory, Masten found that products that were highly design-specific and highly complex were more likely to be produced internally. When the product was both design specific and complex, there was a 92% probability of internal production. If the product was design specific but not complex, the probability of internal production was 31%. The probability of internal production was only 2% when the product was neither design specific nor complex. For this particular company, site specificity was unimportant for explaining the organization of production.

SOURCE: Adapted from Masten (1984)

The Mechanisms of Governance (*Williamson, 1996*)

Williamson (1996) begins with the following: "Institutions. What are they? How do institutions differ? To what purpose and effect? Where does the action reside? What are the mechanisms? What are the refutable implications? What are the public-policy ramifications? What do the data support?" (p. 3). Williamson argues that organizational economics has the following characteristics:

- Holds that institutions matter and are susceptible to analysis
- Is different from but not hostile to orthodox microeconomic theory
- Is an interdisciplinary combination of law, economics, and organization

Williamson (1996) develops the argument that many puzzles of economic organization turn on an examination of the mechanisms of ex post governance. Williamson appeals to law (especially contract law) and organization (which is broadly construed to include organization theory, sociology, and political science) as well as economics and contends that

the main purpose and effect of nonstandardized contractual forms are to economize on transaction costs. The identification, explication, and mitigation of contractual hazards are central to transaction costs analysis. The analytical action resides in the details of transactions and governance. Williamson proposes a logical structure of organization in which the alignment of transactions with governance structures is the source of refutable implications. Farsighted contracting, credible commitments, and contractual hazard mitigation figure prominently in the analysis (Williamson & Winter, 1991). Williamson (1996) submits that the institutional environment (laws, polity, and so forth) and the institutions of governance (markets, hierarchies, and so forth) matter in ways that are pertinent to industrial organization and much else, such as economic history, comparative economic systems, labor economics, business strategy, multinational business, corporate finance, and organization theory (Scott, 1995). According to North (1990), institutions are the humanly devised constraints that structure political, economic, and social interactions. Institutions consist of both informal constraints (e.g., sanctions, taboos, customs, traditions, and codes of conduct) and formal rules (e.g., constitutions, laws, property rights). North (1990) focuses on the level of the institutional environment, the so-called rules of the game.

Williamson (1996) points out that the second, more microanalytic level at which institutional economics works is at the level of the institutions of governance (see Figure 2.4). Williamson is primarily concerned with the institutions of governance (markets, hybrids, hierarchies, bureaus, and so forth). The institutions of governance framework mainly takes the institutional environment as given. The institutions of governance operate at the level of individual transactions, whereas the institutional environment is more concerned with composite levels of activity. Another difference is that the two frameworks operate differently with respect to intentionality. Although both the institutional environment and the institutions of governance have evolutionary origins, the ramifications of each are different. The immense difficulties of changing the institutional environment to promote economizing outcomes in the aggregate helps explain North's (1990) conclusion that economic history is overwhelmingly a story of economies that failed. By contrast, transaction costs theory contemplates success: Taking the institutional environment as given, economic agents purportedly align transactions with governance structures to effect economizing outcomes. Not only is such a prediction a source of numerous refutable implications, but also the data are largely corroborative.

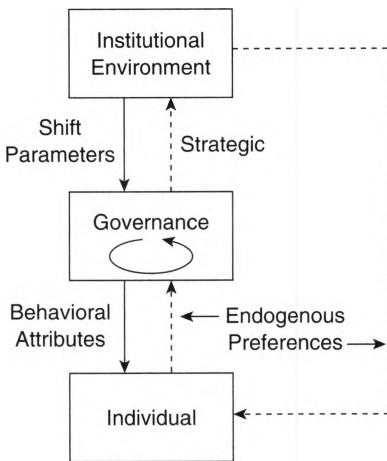


Figure 2.4 A Layer Schema

SOURCE: Williamson (1996, p. 223)

Law, Economics, and Organization. Williamson (1996) submits that parsimony, after all, is what science is after. As with most things, there are trade-offs. Simple theories that finesse or obfuscate core issues are unhelpful. If the action is in the details of transactions and governance, we need to meet the economic problems on terms that are responsive to the needs.

Williamson (1996) notes that orthodox microeconomic theory frequently assumes (often implicitly) that property rights are easy to define and that the courts knowledgeably enforce property rights and contracts at a negligible cost. In contrast, transaction costs theory treats property rights and contracting as problematic.

Remediableness. Williamson (1996) maintains that transaction costs theory avoids hypothetical ideals and insists that the relevant comparisons are with feasible alternatives, all of which are flawed. Within the feasible options, the relevant economic test is whether an alternative can be described that can be implemented with expected net gains. This test is the remediableness criterion. Claims of (path dependency arguments of) inefficiency (Arthur, 1994), which can be recognized only after the fact, or cannot be implemented with net gains, or both, have no operational importance.

Williamson (1996) argues that transaction costs theory holds that the main purpose and effect of economic organization is economizing on

transaction costs. Transaction costs theory concedes that comprehensive contracting is not a feasible option (by reason of bounded rationality), yet transaction costs theory maintains that many economic agents have the capacities both to learn and to look ahead, perceive contractual hazards, and factor these perceived hazards back into the contractual relation, thereafter to devise responsive institutions. Private ordering through ex post governance is where the main action resides.

Williamson (1996) notes various contractual hazards:

- Hazards of bilateral dependency (Williamson, 1971)
- Hazards that accrue to weak property rights (Libecap, 1989)
- Measurement hazards (Barzel, 1982; Ouchi, 1979)
- Weaknesses in the institutional environment (North, 1990)

Williamson (1996) submits that superior economic performance is realized by working out of a farsighted but incomplete contracting approach in which the object is to use institutions as (cost-effective) instruments for contractual hazard mitigation.

Williamson (1996) notes that Machiavelli advised his prince to breach contracts with impunity, when following the contract would be against his interest and when the reasons that made him bind himself no longer exist. This myopic approach to contract should be contrasted with a more farsighted (but nonetheless incomplete) approach to contract, according to which the prince is advised to mitigate ex post opportunism by crafting ex ante (mutual sunk cost) safeguards. Rather than reply to opportunism in kind, the wise prince is one who seeks both to give and to receive credible commitments. Partly, that entails economic incentive realignment, but mainly the need is to craft governance structures with superior adaptive properties. (See the section titled “Cooperative Game and Mutual (Sunk-Cost) Commitment” in this chapter for a game-theoretic illustration of the importance of crafting ex ante mutual sunk costs safeguards.)

Adaptation. Williamson (1996) maintains that adaptation is the central problem of economic organization, of which two types are distinguished: autonomous or Hayekian (Hayek, 1945, 1978) adaptation (in which markets enjoy the advantage) and cooperative or Barnardian (Barnard, 1938, 1948) adaptation (in which the advantage accrues to hierarchy). The study of incomplete contracting in its entirety implicates both ex ante incentive alignment and ex post governance.

Williamson (1996) notes that the invisible hand of Adam Smith (1776/1937) and the marvel of the market to which Hayek (1945) referred have spontaneous origins: The price system is one of those institutions that humans have learned to use after they stumbled on it without understanding it. What interested Barnard (1938), however, was not spontaneous cooperation but induced cooperation—the kind of cooperation that is conscious, deliberate, and purposeful. Barnard argues that authority is a solution to a complex problem of coordination/adaptation and that such authority arises out of mutual consent. Barnard maintains that both the decision of an individual to join an organization and the decision to continue reflect a comparative net benefit assessment. Presented with different employment scenarios, persons consciously choose whether or not they will enter into a specific cooperative system. Expanding a zone (of indifference or acceptance) to include greater (potential) burdens or sacrifices must be attended by greater inducements. Barnard argues that formal and informal organization always and everywhere coexist and that informal organization contributes to the viability of formal organization in three significant respects: communication, cohesiveness, and the maintenance of personal integrity. Williamson (1996) notes that left unmentioned by Barnard was the concept of influence activities—that is, subgoal pursuits (see Simon, 1947). Williamson concludes that to achieve the science of organization described by Barnard, economics and organization theory need to come together. Prospects for a science of organization are improved as a consequence.

The Analysis of Discrete Structural Alternatives. Williamson (1996) maintains that (1) firms employ different means than markets employ; (2) discrete contract law differences provide crucial support for and serve to define each generic form of governance; and (3) marginal analysis is typically concerned with second-order refinements to the neglect of first-order economizing.

Williamson (1996) argues that the implicit contract law of internal organization is forbearance. Thus, whereas courts routinely grant standing to firms should there be disputes over prices, the damages to be ascribed to delays and failures of quality, courts will refuse to hear disputes between one internal division and another over identical technical issues. If access to the courts is denied, the parties must resolve their differences internally. Accordingly, hierarchy is its own court of ultimate appeal. To review alleged mistakes of judgment or to adjudicate internal disputes would sorely test the competence of courts and would undermine the efficacy of hierarchy. Accordingly, the reason why

the market is unable to replicate the firm with respect to fiat is that market transactions are defined by legal rules of an altogether different kind. There is a logical structure to classical market contracting, and there is also a logical structure for forbearance law, and the choice of one regime precludes the other regime. The underlying rationale for forbearance law is twofold: (1) Parties to an internal dispute have deep knowledge—both about the circumstances surrounding a dispute as well as the efficiency properties of alternative solutions—that can be communicated to the court only at a great cost, and (2) permitting the internal disputes to be appealed to the court would undermine the efficacy of hierarchy. Williamson (1996) notes that this argument contradicts Alchian and Demsetz's (1972) claims that the firm has “no power of fiat, no authority, no disciplinary action any different in the slightest degree from ordinary market contracting” (p. 777). That claim is exactly wrong: Firms can, and do, exercise fiat that markets cannot. Prior neglect of contract law differences and their ramifications explain the disparity (see also, Masten 1988).

Williamson (1996) concludes that neoclassical economics was too preoccupied with issues of allocative efficiency, in which marginal analysis was featured to the neglect of organizational efficiency (first order economizing), in which discrete structural alternatives were brought under scrutiny. Market-favoring predispositions need to be disputed, lest the study of economic organization in all its forms be needlessly and harmfully truncated.

Corporate Finance and Corporate Governance. Williamson (1996) notes that the classical agency theory problem is the separation of ownership and control, and the classical transaction costs problem is vertical integration. Both agency theory and transaction costs theory adopt an efficient-contracting orientation to economic organization. However, there are important differences between agency theory and transaction costs theory:

- In agency theory the individual is the unit of analysis, and in transaction costs theory the transaction is the unit of analysis.
- Agency costs focus on ex ante costs; transaction costs emphasize ex post costs.
- There is a legal centralism assumption of agency theory and a private ordering assumption in transaction costs theory.

Calculativeness, Trust, and Economic Organization. Williamson (1996) submits that to craft credible commitments (through the use of

economic bonds, economic hostages, information disclosure rules, specialized dispute settlement mechanisms, and the like) is to create functional substitutes for trust. Albeit vitally important to economic organization, such substitutes should not be confused with (real) trust. Transaction costs theory refers to contractual safeguards, or their absence, rather than trust, or its absence. Williamson (1996) argues that it is redundant at best, and can be misleading, to use the term *trust* to describe commercial exchange for which cost-effective economic safeguards have been devised in support of more efficient exchange.

Williamson (1996) suggests reserving the term *trust* for the personal type. A deep and abiding trust relation cannot be created in the face of calculativeness. Calculative trust is a contradiction in terms. Personal trust is characterized by the absence of monitoring, favorable or forgiving predilections, and discreetness. Such relations are clearly very special. Trust, if it obtains at all, is reserved for very special relations between family, friends, and lovers. Such trust is also the stuff of which tragedy is made. It goes to the essence of the human condition. Not only is *calculated trust* a contradiction in terms, but also user-friendly terms, of which *trust* is one, have an additional cost. The business world is reorganized in favor of the cynics, and against the innocents, when social scientists employ user-friendly language that is not descriptively accurate—since only the innocents are taken in.

Strategizing, Economizing, and Economic Organization. Williamson (1996) notes that the beguiling language of strategizing—warfare, credible threats, and the like—notwithstanding, students of economic organization are better advised to focus on issues of economizing (e.g., Barney & Ouchi, 1986; Mahoney, 1992b; Seth & Thomas, 1994; Williamson, 1991; Yao, 1988)—of which credible commitments and adaptation are examples. Here, as elsewhere, the need is to get the priorities straight. Williamson (1996) notes that the leading efficiency approaches to business strategy are the resource-based and the dynamic capabilities approach, which he cautiously deems as promising. Williamson (1996) concludes by stating that he anticipates that the science of organization to which Barnard (1938) referred will be realized by this generation of organizational economics scholars.

Now that we have completed our studies on transaction costs theory, where the primary area of concern is governance of the transaction at the organizational level, we move to property rights theory, where we emphasize study of more macrophenomena at the institutional level, beginning with Libecap's (1989) lucidly written research book.

Cooperative Game and Mutual (Sunk-Cost) Commitment

Game theory is the analysis of rational behaviors in situations involving interdependent outcomes when the firm's payoff depends on what it does and what the other firm does (see Dixit & Nalebuff, 1991; McMillan, 1992; and Rasmusen, 1989). Because many strategic decisions involve interdependent outcomes, game-theoretic analysis can be applied, for example, to the study of vertical supplier-buyer relationships. In game-theoretic models each firm's optimal action depends on what the firm believes its counterpart will do. In other words, the game-theoretic analysis requires assumptions about the counterpart's rationality and the counterpart's belief about the firm's own rationality.

The prisoners' dilemma game provides a powerful metaphor for a fundamental conflict that arises in business situations involving vertical interdependence (Saloner, 1991; Scherer & Ross, 1990). In the prisoners' dilemma, what is best for the individual firm is to maximize its own economic profit, and this individual rationality is detrimental to group performance. The collective rationality is for both firms to cooperate and obtain a higher group payoff, but the individual rationality is for each firm to play their dominant strategy, given the current payoff matrix. No matter how much they preach the importance of the group (common) good, there is always the possibility that the poor group outcome will be the dominant strategy equilibrium as predicted in the prisoners' dilemma game.

One way out of the prisoners' dilemma occurs when the players take steps that change the payoff matrix. Paradoxically, worsening some of one's own payoff possibilities may improve the likely outcome of the game (Schelling, 1960). Consider the case of an information technology (IT) investment between Wal-Mart and P&G. The best strategies for the collective good are that both firms cooperate. Although this mutual cooperation is collectively rational, it is unfortunately not individually rational in terms of individual firm profitability. Thus, in Payoff Matrix 1(Box 2.1), we have an example of the Prisoners' dilemma situation where the Nash equilibrium point (91, 91) is predicted when each firm plays their dominant strategy of behaving opportunistically.

Let's suppose P&G posts an economic bond (e.g., investing in a relation-specific IT system) that P&G would lose if it defects from the joint profit-maximizing collaboration. In effect, this action unilaterally lowers the payoff associated with an opportunistic behavior by P&G

Box 2.1

Payoff Matrix 1		Wal-Mart	
		Cooperation	Opportunism
P&G	Cooperation	(112, 112)	(58, 123)
	Opportunism	(123, 58)	(91, 91)

Box 2.2

Payoff Matrix 2		Wal-Mart	
		Cooperation	Opportunism
P&G	Cooperation	(112, 112)	(58, 123)
	Opportunism	(-28, 58)	(-51, 91)

(i.e., from 123 to -28 and from 91 to -51, respectively; see Payoff Matrix 2, Box 2.2). Thus, P&G eliminates the attractiveness of defecting from the cooperative solution. Such a voluntary agreement is considered self-enforcing because third-party enforcement is not relied on (Telser, 1980). As a result, cooperation is now P&G's dominant strategy.

If Wal-Mart, in a myopic manner, decides not to make some form of commitment to P&G, then Payoff Matrix 2 will not become the long-run equilibrium. P&G would eventually abandon their commitment to Wal-Mart, and the situation would return to Payoff Matrix 1.

However, if Wal-Mart wisely emulates P&G's action (i.e., mutually commits to a relation-specific IT with P&G), then this action further transforms the situation to Payoff Matrix 3 (Box 2.3) to encourage mutual commitments to cooperation that increases both firms' payoffs to (112, 112). This example illustrates that firms involved in interdependent outcomes should seek both to give and receive mutual (sunk-cost) commitments that facilitate ongoing relationships and adaptation (Williamson, 1985). Thus, (sunk-cost) commitments are the dynamics of good strategy (Ghemawat, 1991).

Box 2.3

Payoff Matrix 3		Wal-Mart	
		Cooperation	Opportunism
P&G	Cooperation	(112, 112)	(58, -28)
	Opportunism	(-28, 58)	(-51, -51)

It should be noted here that this game-theoretic example provides a reconstructed logic (Kaplan, 1964) of the IT alliance between P&G and Wal-Mart. Whether this reasoning was the logic-in-use (Kaplan, 1964) of the managers of either or both of these firms is questionable. It is worth emphasizing that this game-theoretic reconstruction enables us to understand more fully the theoretical soundness of mutual commitments to support economic exchange.





3

Property Rights Theory

Coase (1960) initiated a flurry of property rights research that perhaps reached its peak with Alchian and Demsetz (1973). Barzel (1989) and Eggertsson (1990) provide useful discussions of the early property rights research literature. Much of this early property rights literature (with Demsetz [1967], serving as an exemplar of the neoclassical economics tradition) was quite optimistic about the evolution of property rights toward economic efficiency. Three important criteria for efficiency of property rights are (1) universality—all scarce resources are owned by someone; (2) exclusivity—property rights are exclusive rights; and (3) transferability—to ensure that resources can be allocated from low to high yield uses. In Demsetz's (1967) neoclassical economics framework, all three criteria are in place (in the long run).

In some sense, Libecap (1989), and especially North (1990), can be understood as providing historical accounts that challenge this earlier optimistic view of an inevitable evolution of property rights toward economic efficiency. The awarding of a Nobel Prize in economics to Douglass North suggests that, at the least, part of the economics profession has (implicitly) accepted that the evolution of institutional environment change toward economic efficiency often fails.

Students studying the economics of organization should take note that changes in theoretical views do take place. However, to make headway, you need to come prepared with the facts along with an analytical approach (and often a tough skin) to handle the almost inevitable initial resistance by others to new ideas that aim to overturn the conventional wisdom.

We begin this chapter on property rights with Libecap's (1989) *Contracting for Property Rights*, in which Libecap provides substantive research concerning the way property rights are formed. Libecap's research book is a synthesis of theory and history, which emphasizes the complexities of property rights formation. *Contracting for Property Rights*, in my judgment, is one of the best books in the property rights research literature,

a major contribution both to the theory of property rights and to our understanding of economic history. In particular, we learn from Libecap that distributional conflicts present political risks to politicians, giving these politicians incentives to propose regulations that do not seriously upset status quo rankings and that offer only limited relief from property rights economics inefficiencies due to common pool resource losses. Similar incentives and vested interests exist for regulatory agencies.

North (1990) applies his theories of the interplay between institutional evolution and political and economic organization to a range of historical examples, including the development of management structures, insurance, and financial markets. North offers a broad perspective on how institutions persist and change. In particular, North is concerned as much with explaining the evolution of institutional frameworks that induce economic stagnation and decline as with accounting for the successes.

Eggertsson (1990) emphasizes the variety of organizational forms and institutional arrangements that we observe in practice. Eggertsson's approach to explain such variety is to seek a new synthesis of neoclassical economic theory and institutional theory. As Eggertsson views the research literature, three important levels are identified. At the first level, the structure of property rights and forms of organization are explicitly modeled but are treated as exogenous. At the second level, organization form is endogenous, but the fundamental structure of property rights remains exogenous. At the third level, attempts are made to consider both social and political rules, and the structure of political institutions as endogenous in a positive transaction costs world. Eggertsson organizes his book on the basis of these three levels of analysis.

Barzel (1989), in the tradition of Coase (1960), provides a unified structure to analyze exchange, the formation of property rights, and organization. Barzel emphasizes that because of the costliness of measuring accurately all of an asset's attributes, rights are never fully delineated, and property is consequently in danger of appropriation by others due to adverse selection, free-riding behavior, and shirking, among other reasons.

Hart (1995) argues that contractual incompleteness and control are two concepts that can be brought together to understand a number of economic institutions and arrangements. Hart focuses on understanding firms and financial structures. For the purposes of the current book, I focus on the first half of Hart's work—understanding firms. Hart focuses on some fundamental questions: What does *ownership*

mean? What determines the boundaries of the firm? What are the economic implications of contractual incompleteness? What are the roles of nonhuman assets and the nature of authority?

In the 1990s, modern property rights theory (which provides more formalized mathematical models) has gained momentum in organizational economics, and Hart's (1995) work is an exemplar of this modern property rights framework. With the increasing importance of intellectual property rights in our current information age (both early and modern), property rights theory predictably will receive greater attention in strategic management and may prove to spur a growth area for research in the years ahead.

Contracting for Property Rights (*Libecap, 1989*)

How do institutions evolve in response to individual economic incentives, strategies, and choices? Libecap (1989) emphasizes that property rights matter.¹ Property rights provide the basic economic incentive system that shapes resource allocation. What has been largely missing is why property rights take the form that they do. Libecap argues that property rights are formed and enforced by political entities and that property rights reflect the conflicting economic interests and bargaining strength of those affected. Moreover, because today's choices are constrained by yesterday's decisions, history matters.

Property rights are the social institutions that define or delimit the range of privileges granted to individuals of specific resources, such as parcels of land or water. Private ownership of these resources may involve a variety of property rights, including the right to exclude nonowners from access, the right to appropriate the stream of economic rents from use of and investments in the resource, and the rights to sell or otherwise transfer the resource to others. Property rights institutions range from formal arrangements, including constitutional provisions, statutes, and judicial rulings, to informal conventions and customs regarding the allocations and uses of property. Such institutions critically affect decision making regarding resource use and, hence, affect economic behavior and economic performance.

¹Seminal works in classical property rights theory include Alchian (1965), Alchian and Demsetz (1973), Barzel (1989), Cheung (1969), Coase (1960), Demsetz (1967), and Furubotn and Pejovich (1972).

Because of the huge advantages of secure property rights, economic decision makers often are hypothesized to adopt, or to modify, property rights to mitigate the economic losses of the common pool, as soon as the private benefits of so doing outweigh the private costs. Forces that drive the adjustments in property rights include new market prices and production possibilities to which old arrangements are poorly attuned (Demsetz, 1988, 1995). Davis and North (1971) are explicit in the argument: "It is the possibility of profits that cannot be captured within the existing arrangement structure that leads to the formation of new (or the mutation of old) institutional arrangements" (p. 39).

Despite these optimistic assertions in the (neoclassical) property rights literature, the actual process by which property institutions change, and whether the changes represent an efficient economic solution to a particular social problem, have received much less attention. North (1981) notes, "But the fact that growth has been more exceptional than stagnation or decline suggests that efficient property rights are unusual in history" (p. 6).

Libecap (1989) argues that because certain property rights arrangements can reduce transactions costs in exchange and production, and encourage (sunk cost) investments to promote overall economic growth, such property rights have public goods aspects. As with all public goods, though, there are economic hazards in attempting to change property rights. For example, there may be shirking and uncooperative behavior among the bargaining parties that will affect the institutions that can be established. In bargaining over creating or modifying property rights, the positions taken by the various bargaining parties, including private claimants, bureaucrats, and politicians, will be molded by their private expected gains, as well as by the actions of the other parties.

Libecap (1989) emphasizes that property rights institutions are determined through the political process, involving either negotiations among immediate group members or the lobbying activities that take place at higher levels of government. The political process of defining and enforcing property rights can be divisive because of the distributional implications of different property rights allocations. If influential parties cannot be sufficiently compensated through share adjustments in the political process to win their support, beneficial institutional change (even as modified through contracting concessions) may not occur, and the potential economic gains fostered by the proposed arrangement will be forgone.

Even though society would be better off with the public goods provided by the new property rights, the distributional implications lead influential parties to oppose institutional change. In principle, it is possible to construct a side payment scheme that would compensate those who otherwise would oppose a desirable change in property rights. But in practice, devising perfectly compensating side payments to bring agreement encounters formidable obstacles, including questions of who would receive side payments, who should pay, what size the compensation should be, and what form the compensation should take. Libecap (1989) argues that distributional conflicts, and efforts to address such conflicts, can block institutional change or so influence the property rights arrangement that what ultimately emerges as institutional change bears little resemblance to that which was initially proposed.

The roles of time and precedent suggest that there may be historical path dependences for institutional change. Past property rights decisions serve to limit the menu of possible institutional solutions to varying economic problems. Libecap (1989) states that recent historical investigation suggests a less optimistic view of property rights change is in order. This conclusion is based on examination of the role of interest groups and conflicts among these groups over the distributional effects of property law and government regulation.

Analytical Framework. Libecap (1989) notes that the nature in which property rights are defined and enforced fundamentally impacts the performance of an economy for at least two reasons. First, by assigning ownership to valuable resources and by designating who bears the economic rewards and costs of resource-use decisions, property rights institutions structure incentives for economic behavior within the society. Second, by allocating decision-making authority, the prevailing property rights arrangement determines who the key actors are in the economic system.

In contracting over proposed property rights, the bargaining positions taken by the various parties depend on how these parties view their economic welfare under the new arrangement relative to the status quo. Estimates of the likely net economic gains or losses from institutional change faced by each party require an evaluation of the overall productive possibilities with the new property rights arrangement and the distribution of economic rents it authorizes. The bargaining parties must see their economic welfare improved, or at least made no worse off, for them to support institutional change, and each party has an incentive to seek as large a share of economic rents under the new arrangement as possible. This competition for the range of economic opportunities made

possible by changes in property rights is costly to society. Competition among the contracting parties uses resources, and such competition leads to changes in the definition and assignment of property rights that affect the nature and size of aggregate economic benefits that are possible. The side-payment schemes reached through the political process may be too incomplete to resolve the distributional conflicts needed for more than minimal institutional change to occur at any time.

Primary motivations for contracting for property rights are the aggregate (common pool) losses that arise under conditions of poorly defined property rights (e.g., open fisheries, oil field dissipation). In these circumstances, resource values fall for several reasons. First, because property rights to the resource are not assigned, individuals in their production decisions do not have to consider the full social costs of their activities. Individuals use the resource too rapidly at any time, relative to interest rate and price projections. Further, competitive pressures under conditions of poorly defined property rights encourage short-time horizons in production. The economic incentive to invest (e.g., in new technology) is reduced because investors cannot anticipate that they will capture any of the resulting economic returns due to insecure property rights.

Second, resource values fall because exchange and reallocation of the resource to higher valued uses become more costly and less effective if property rights are absent. Demsetz (1967) argues that an assignment of property rights is a prerequisite before decentralized price-making markets can form to define asset prices. Well-defined asset prices are needed to reflect underlying demand and supply conditions and to facilitate socially valuable exchange among economic agents. Without the more complete market signals possible when property rights are well defined, resources may not flow smoothly to higher valued uses as economic conditions change. Whether or not the more complete defining of property rights is socially beneficial depends on the magnitude of common pool losses, the nature of contracting costs to resolve such losses, and the economic costs of defining and enforcing property rights (Coase, 1960).

In Libecap's (1989) analytical framework, pressures to change existing property rights can emerge from the following factors:

- Shifts in relative prices
- Changes in production and enforcement technology
- Shifts in preferences and other political parameters

A number of implications can be drawn from Libecap's (1989) analytical framework:

1. All things being equal, the greater the size of the anticipated aggregate economic benefits of institutional change (the greater the economic losses of the common pool), the more likely new property rights will be sought and adopted because it is more likely that a politically acceptable share arrangement can be devised by politicians to make enough influential parties better off so that institutional change can proceed.
2. The larger the number of competing interest groups, the more likely distributional conflicts will block or delay institutional change because the greater the number of competing interest groups with a stake in the new definition of property rights, the more claims that must be addressed by politicians in building a consensus for institutional change.
3. The greater the heterogeneity of competing interest groups, the more likely distributional conflicts will block or delay institutional change. Important differences across the parties in information regarding the resource, as well as in production costs, size, wealth, and political experience, will make the formation of winning political coalitions, and a consensus on the proposed assignment or adjustment of property rights, more difficult.
4. Distributional conflicts will be intensified if there are known serious information asymmetries among the competing parties regarding the evaluation of individual claims. These distributional conflicts will occur quite aside from any strategic bargaining efforts if private estimates of the economic value of current property rights and of potential economic losses from the new system cannot be conveyed easily or credibly to politicians and the other bargaining parties.
5. The greater the concentration of wealth under the proposed property rights allocation, the greater the likelihood of political opposition and the less likely institutional change will be adopted without modification by politicians. In these circumstances, enough influential parties may see their economic welfare made worse, or at least not improved, by the change that political support for such change does not materialize.

Contracting for the Unitization of Oil Fields. Libecap (1989) observes that since the first discovery of petroleum in the United States in 1859, oil production has been plagued by serious common pool losses. These common pool losses arise as numerous firms compete for migratory oil lodged in subsurface reservoirs. Under the common rule law of capture, private property rights to oil are assigned only upon extraction. For each of the firms on a reservoir, a plan of dense-well drilling and rapid production allows the firm to drain oil from its neighbors and to take advantage of the low extraction costs that exist early in oil field development. In new, flush oil fields, subsurface pressures are sufficient to expel oil without costly pumping or injection of water or natural gas into the reservoir to drive oil to the surface.

Libecap (1989) notes that under these conditions, when there are multiple firms on a reservoir, each firm has an economic incentive to

drill competitively and to drain to increase its share of oil field economic rents, even though these individual actions lead to aggregate common pool losses. Economic rents are dissipated as capital costs are driven up with the drilling of excessive numbers of wells (more than geological conditions require or price and interest rate projections warrant) and with the construction of surface storage, where the oil can be held safe from drainage by other firms. Unfortunately, once in surface storage, oil is vulnerable to fire, evaporation, and spoiling. Rapid extraction also increases production costs as subsurface pressures are vented prematurely, forcing the early adoption of pumps and injection wells. Total oil recovery falls as pressures decline because oil becomes trapped in surrounding formations, retrievable only at high extraction costs. Finally, economic rents are dissipated as production patterns diverge from those that would maximize the economic value of output over time. Some estimates indicate that oil recovery rates of only 20% to 25% occur with competitive extraction, whereas recovery rates of 85% to 90% were thought possible with controlled withdrawal.

A complete solution to the common pool problem is oil field-wide unitization. Under unitization, production rights are delegated through negotiations to a single firm, the unit operator, with net revenues apportioned among all parties on the field (including those that would otherwise be producing). As the only producer on the field and a residual profit claimant, the unit operator has an economic incentive to maximize field rents. Accordingly, unitization results in important economic gains: a time stream of output that more closely approximates the rent-maximizing pattern, increased oil recovery (2 to 5 times greater than unconstrained production), and reduced wells and other capital costs. Despite these reasons for mitigating the substantial losses involved in common pool crude oil production, complete fieldwide unitization had not been widespread. As late as 1975, only 38% of Oklahoma production and 20% of Texas production came from fieldwide units.

Libecap (1989) argues that the key issue in blocking agreement on the voluntary unitization of oil fields is the distributional conflict over the share formula to divide the net proceeds of unit production among the various contractual parties. Uncertainties and information asymmetries regarding the economic valuation of individual firms oil leases, which are the basis for unit shares, are important contributors to the disagreements that block unitization, even in the presence of large and uncontroversial aggregate economic gains from unit formation. In share negotiations, two serious problems arise. First, unitization contracts must assign, once and for all, shares at the time the contract

is completed. This assignment is needed because, in reservoir dynamics after unitization, it is impossible to link unit production to particular leases, which would be necessary for adjusting shares. A second problem in unitization contracting is general uncertainty and asymmetrical information regarding relative preunitization lease values, which determine unit shares. These serious contractual problems block agreement on lease value estimates and proposed shares in unit economic rents.

Besides the information issues, small lease owners were given preferential drilling permits by regulatory authorities under prorationing controls adopted by states in the absence of widespread unitization. Differences in lease value estimates can block consensus on any side payments to draw potential holdouts into agreement. Under unanimity voting rules, small firms could delay or block the formation of fieldwide units. The empirical evidence that Libecap (1989) presents supports the notion that as fieldwide primary production nears an end, unitization agreements become more likely. By that time, information asymmetries among the firms become less important as all leases near primary depletion.

The failure of unitization to be widespread, despite significant aggregate economic gains from unitizing oil production, is another example of how distributional conflicts over rental shares can limit the adoption of property rights to increase economic efficiency. *The analysis presented by Libecap (1989) suggests that swift institutional responses to common pool losses to promote more rational resource use and greater economic growth cannot be taken for granted. Distributional conflicts inherent in any new property rights arrangement can block, or critically constrain, the institutions that can be adopted.* More attention, accordingly, should be directed to the distributional implications of property rights arrangements, to the identity and preferences of the various bargaining parties, and to the nature of the side-payment schemes adopted. And, perhaps even more important, attention should be directed to the history of past political agreements if the observed variations in property rights and associated economic and strategic behaviors are to be more fully understood.²

²Libecap and Wiggins (1985) provide empirical evidence of the influence of private contractual failure on regulation for the case of oil field unitization. Kim and Mahoney (2002) provide a fairly comprehensive collection of references on the property rights approach and provide resource-based and property rights perspectives concerning oil field unitization. Finally, elements of the oil field unitization case discussed in this chapter provide insight on the conflicts between Kuwait and Iraq that led to the Persian Gulf War of 1990–1991 (for some details, see Milgrom & Roberts, 1992, p. 296).

Libecap (1989) provides an exemplar for students studying the economics of organization on the use of case studies to build up and support a theoretical argument. Libecap, in my judgment, convincingly shows that the assertion that property rights will naturally move toward economic efficiency is frequently glib and inaccurate.

Institutions, Institutional Change, and Economic Performance (*North, 1990*)

Now that we have studied Libecap (1989), we next examine the work of North (1990). Early in his career (e.g., Davis & North, 1971), North held an (overly) optimistic view about the evolution of property rights toward economic value creation. In contrast, North (1990) later emphasized the persistence of inefficient property rights regimes throughout economic history to provide a main case explanation for why the whole world is not economically developed. The objective of North's research book is to provide an analytical framework to integrate institutional analysis into economics and economic history. North also provides us with a new understanding of historical change.

North (1990) examines the nature of institutions and the consequences of institutions for economic and societal performance and then outlines a theory of institutional change, not only to provide a framework for economic history but also to explain how the past influences the present and future, the way incremental institutional change affects the choice set of decision makers at a moment in time, and the nature of path dependencies. The primary objective of this research book is to achieve an understanding of the differential performance of economies through time.

North (1990) ties together the threads and illustrates the relationships between institutions, transaction costs, and transformation (production) costs. North then explores organizations and the way that they interact with institutions and argues that the nature of incremental institutional change, together with the imperfect way by which decision makers interpret their environment and make choices, accounts for path dependencies and makes history relevant.

North (1990) asks the following question: What combination of institutions best permits capturing the economic gains from trade? Institutions are defined as any constraint humans devise to shape their interactions and organizations, created to take advantage of the opportunities presented by institutions in shaping the development of

economies. The importance of institutions arises from the costliness of measuring what is valuable, from protecting rights, and from policing and enforcing agreements.

North (1990) emphasizes that history matters. History matters not just because we can learn from the past but also because the present and the future are connected to the past by the continuity of a society's institutions. Today's decisions and tomorrow's choices are shaped by the past. And the past can only be made intelligible as a story of institutional evolution.

For North (1990), the central focus is on the problem of human cooperation—specifically, the cooperation that permits economies to capture the economic gains from specialization and trade. The evolution of institutions that create a hospitable environment for cooperative solutions to complex exchange provides for economic growth.

North (1990) argues that institutions reduce uncertainty by providing a structure to everyday life. Institutions are a guide to human interactions, and these institutions define and limit the set of choices of individuals. Institutions include any form of constraint that humans devise to shape human interaction. Are institutions formal or informal? Institutions can be either, and North considers both formal constraints—such as rules that humans devise—and informal constraints—such as conventions and codes of behavior. Institutions may be created, as was the United States Constitution, or institutions may evolve over time, as does the common law. An essential part of the functioning of institutions is the costliness of ascertaining violations and the severity of punishment.

North (1990) makes an important distinction between institutions and organizations. Organizations include political bodies (e.g., political parties, trade unions, family farms, cooperatives), social bodies (e.g., churches, clubs, athletic associations), and educational bodies (e.g., schools, universities, vocal training centers). Organizations are groups of individuals bound by some common purpose to achieve objectives. Modeling organizations requires analyzing governance structures and organizational capabilities and understanding how learning by doing determines the organization's success over time (Oliver, 1997). The institutional framework fundamentally influences both what organizations come into existence and how organizations evolve. In turn, organizations influence how the institutional framework evolves.

North (1990) emphasizes that institutions are a creation of humans and suggests that integrating individual choices with the constraints that institutions impose on choice sets is a major step toward unifying

social science research. The major role of institutions in society is to reduce uncertainty by establishing a stable (but not necessarily efficient) structure to human interactions. Although formal rules may change overnight as the result of political or judicial decisions, informal constraints embodied in customs, traditions, and codes of conduct are much more impervious to deliberate policies. These cultural constraints not only connect the past with the present and future but also provide us with a key to explaining the path of historical change.

North (1990) maintains that the central puzzle of human history is to account for the widely divergent paths of historical change. North notes that although we do observe some convergence among leading industrial nations that trade with each other, an overwhelming feature of the last 10 millennia is that we have evolved into radically different religious, ethnic, cultural, political, and economic societies. Furthermore, the economic gap between rich and poor nations, between developed and underdeveloped nations, is as wide today as it ever was, and perhaps a great deal wider than ever before.

North (1990) then asks the following: What accounts for societies experiencing long-run stagnation or an absolute decline in economic well-being? North and Thomas (1973) make institutions the determinant of economic performance and relative price changes the source of institutional change. North and Thomas provide an essentially efficiency-based explanation: Changes in relative prices create economic incentives to construct more efficient institutions. North (1981), however, abandons the efficiency view of institutions. *Rulers devised property rights in their own vested interests, and transaction costs resulted in typically inefficient property rights prevailing. As a result, it was possible to account for the widespread existence of property rights throughout history (and in the present) that did not produce economic growth.*

North (1990) argues that institutions highly influence the opportunities in a society. Organizations are created to take advantage of those opportunities, and, as the organizations evolve, they alter the institutions. The resultant path of institutional change is shaped by the lock-in that comes from the tightly coupled relationship between institutions and organizations that have evolved as a consequence of the economic incentive structure provided by those institutions and the feedback process by which humans perceive, and react to, changes in the (subjective) opportunity set.

Actors frequently must act on incomplete information and process the information that they do receive through mental constructs, which can result in persistently inefficient paths. Transaction costs in political

and economic markets make for inefficient property rights, but the imperfect subjective models of the actors as they attempt to understand the complexities of the problems they confront can lead to the persistence of inefficient property rights.

North (1990) states that there is a persistent tension in the social sciences between the theories we construct and the evidence we compile about human interaction in the world around us. This tension is most striking in economics, where the contrast between the logical implications of neoclassical microeconomic theory and the performance of economies (however defined and measured) is startling. *North argues that the coercive power of the state has been employed throughout most of history in ways that have stymied economic growth.*

North (1990) maintains that the traditional behavioral assumptions of orthodox microeconomic theory have prevented economists from coming to grips with some fundamental issues that have impeded progress in the social sciences. In particular, North argues that the motivation of actors is more complicated (and their preferences less stable) than assumed in the received wisdom. Further, microeconomic theory implicitly assumes that actors possess cognitive systems that provide true models of the worlds about which they make choices. North insists that this implicit assumption is patently wrong for most of the important problems with which institutional economics and organizational economics are concerned. Individuals make choices based on subjectively derived models that diverge among individuals and the information of actors is so incomplete that in most cases these divergent subjective models show no tendency to converge. Only when we understand these modifications in the behaviors of the actors can we hope to make sense out of the existence and structure of institutions and to explain the direction of institutional change.

North (1990) argues that institutional analysis requires that we delve into two particular aspects of human behavior: motivation and deciphering the environment. Many cases are not simply of wealth-maximizing behavior but of altruism and of self-imposed constraints, which radically change the outcomes with respect to the choices that people actually make. Similarly, we find that people decipher the environment by processing information through preexisting mental constructs through which they understand the environment and solve the problems they confront.

North (1990) notes that the work of Simon (1982) captures the essence of why the subjective and incomplete processing of information plays a crucial role in decision making. Simon's work is useful for

accounting for ideology, based on subjective perceptions of reality, playing a major role in humans' choices. Simon's work brings into play the complexity and incompleteness of our information and the fumbling efforts we make to decipher information. North concludes that the regularized interactions we call institutions may be inadequate to deal with the economic problems at hand.

Culture can be defined as the transmission from one generation to the next, the teaching and replication of knowledge, values, and other factors that influence behaviors. North (1990) argues that culture provides a language-based conceptual framework for encoding and interpreting the information that the senses are presenting to our brain. Importantly, the cultural filter provides continuity and stability. Order is the result of a dense social network where people have an intimate understanding of each other. In the short term, culture defines the way individuals process and use information and hence may affect the way informal constraints are specified. Conventions are culture specific, as indeed are informal rules.

Formal Constraints. North (1990) observes that formal rules can complement and increase the effectiveness of informal rules. Formal rules also may be enacted to modify, revise, or replace informal constraints. Formal rules include political (and judicial) rules, economic rules, and contracts. Economic rules define property rights and, as a crude approximation, economic rules are derived from economic self-interest. Property rights are specified and enforced by political decision making, but the structure of economic interests will also influence the political structure. Indeed, there is a substantial amount of property rights literature that looks on the development of property rights as a simple function of changes in economic costs and economic benefits. North argues that this simplified approach needs modification to account for the obvious persistence of inefficient property rights.

Enforcement. North (1990) argues that the inability of societies to develop effective, low-cost enforcement of contracts is the most important source of both historical stagnation and contemporary underdevelopment in the third world. In developed countries, effective judicial systems include well-specified bodies of law and agents, such as lawyers, arbitrators, and mediators, and one has some confidence that the merits of a case rather than private payoffs will decisively influence outcomes. In contrast, enforcement in the third world economies is uncertain not only because of ambiguity of legal doctrine (a measurement cost) but also because of uncertainty with respect to behavior of the judicial system.

Institutions. North (1990) observes that it takes resources to define and protect property rights and to enforce agreements. Institutions together with the technology employed determine those transaction costs. It takes resources to transform inputs of land, labor, and capital into the output of goods and services, and that transformation is a function not only of the technology employed but of the institutions as well. Therefore, institutions play a key role in the costs of production. The interplay between techniques, institutions, transformation costs, and transaction costs makes clear that the relationships between them are complex.

North (1990) submits that contrasting the institutional framework in countries such as the United States, England, France, Germany, and Japan with Third World countries makes clear that the institutional framework is the critical success factor of economies, both cross-sectionally as well as through time. North further argues that the institutional framework shapes the direction of the acquisition of knowledge and capabilities, and that direction will be the decisive factor for the long-run development of that society. Path dependence is the key to an analytical understanding of long-run change in property rights. Property rights and economic incentives are the underlying determinants of economic performance. Bringing property rights and economic incentives to the foreground focuses attention where it belongs, on the key success factors for the economic performance of societies. One gets efficient institutions by a polity that has built-in economic incentives to create and enforce efficient property rights.

North (1990) concludes that we need to know much more about culturally derived norms of behavior and how such norms of behavior interact with formal rules to get better answers to such issues. We are just beginning the serious study of institutions in organizational economics and strategic management. The promise is there. We may never have definitive answers to all our questions. But students in the next generation of research can do better in both institutional economics and organizational economics research, which will contribute greatly to the evolving science of organization.

An Economic Analysis of Property Rights (Barzel, 1989)

The third property rights book for discussion in this chapter is by Barzel (1989), *An Economic Analysis of Property Rights*, and the fourth book is an overlooked classic by Eggertsson (1990), *Economic Behavior*

and Institutions. The Barzel book is complementary to Libecap (1989), and the Eggertsson book is especially complementary to North (1990). In fact, Eggertsson notes his intellectual debt to Douglass North: "North's vision that the economic approach, augmented by transaction costs and property rights, is a general tool for the study of society at all levels has inspired this book" (p. xiv).

Barzel (1989) notes that because transacting is costly, as an economic matter property rights are never fully delineated. Property rights of individuals over resources consist of the rights, or the powers, to consume, obtain income from, and alienate those resources. Obtaining income from and alienating resources require exchange, and exchange is the mutual ceding of rights. Legal rights, as a rule, enhance economic rights, but legal rights are neither necessary nor sufficient for the existence of the economic rights. The rights people have over resources (including themselves and other people) are not constant; they are a function of their own direct efforts at protection, of other people's capture attempts, and of government protection.

Barzel (1989) views the concept of property rights to be closely related to that of transaction costs. Transaction costs are defined as the economic costs associated with the transfer, capture, and protection of rights. When transaction costs are positive, rights to resources cannot be perfectly delineated. Exchange that otherwise would be attractive may be forsaken because of such exchange costs.

What underlies this costliness of transacting? What are the factors that prevent people from realizing the full economic value of their resources? Commodities have many attributes whose levels vary from one specimen of a commodity to another. The measurement of these levels is too costly to be comprehensive or entirely accurate. How difficult it is to obtain full information in the face of variability fundamentally determines how difficult it is to delineate rights. *Because it is costly to measure commodities fully, the potential of wealth capture is present in every exchange.* The opportunity for wealth capture is equivalent to finding property in the public domain; in every exchange, then, some wealth spills over in the public domain, and individuals spend resources to capture this economic wealth. Whereas people always expect to gain from exchange, they also always spend resources on the capture of economic wealth. Individuals maximize their (expected) net gains, the gains from exchange as conventionally perceived net of the economic costs of effecting exchange.

The sale of cherries illustrates the phenomenon of wealth capture. Obvious problems of information present themselves when cherries are

exchanged. Customers must spend resources to determine whether a store's cherries are worth buying and to determine which particular cherries to buy. Store owners who allow customers to pick and choose cannot easily prevent these customers from eating cherries after they have decided whether or not to buy the cherries, nor can store owners prevent customers' careless handling of cherries. Indeed, the process of picking and choosing itself allows wealth capture in the form of excess choosing. Some of the cherries' attributes, then, are placed in the public domain. The high cost of information results in transaction costs: economic costs that would not arise if the owner and the consumer of cherries were the same person. If information about the cherries was costless, their initial owner would not have to relinquish any rights, and pilfering, damage, and excess choosing would be avoided. In business reality, such public domain problems are unavoidable; people can take steps, however, to reduce the associated economic losses.

Contracts govern the exchange of property rights and are central to the study of such rights. The exchange value of a resource is a function of the gross income the resource can generate and of the transaction costs of measuring and policing its exchange. These economic costs also determine the pattern and the degree of ownership. The ownership of a resource's attributes is expected to gravitate into the hands of those people who are most inclined to affect the income flows that the attributes can generate.

Barzel (1989) maintains that the property rights transaction costs model can generate a better understanding of the allocation of resources and of the interaction of this allocation with economic organization. The research literature that assumes that the economic costs of transactions are zero and that all property rights are perfectly well delineated is incapable of dealing with a vast array of actual observed practices. Particularly glaring is the inability of such an approach to explain why exchange parties would ever impose restrictions on each other. The property rights approach is capable of addressing such issues, and we continue our property rights study with Eggertsson (1990).

Economic Behavior and Institutions (*Eggertsson, 1990*)

Eggertsson (1990) considers the costs of transacting and the allocation of resources; transaction costs and efficiency; the quality dimensions of goods and the costs of measurement, property rights, and their

dimensions; the partitioning of property rights; property rights and contract theory; the emergence of property rights; competition and the costs of alternative economic organizations; and economic outcomes. This research book provides a clear structure to and balanced overview of the property rights literature. Eggertsson (1990) provides a mature yet compact presentation of property rights research.

Eggertsson (1990) observes that organizations and institutions are not invariant; organizations and institutions vary with time and location, with political arrangements and structures of property rights, with technologies employed, and with physical qualities of resources and services that are exchanged. In fact, production involves not only the physical transformation of inputs into outputs but also the transfer of property rights between the owners of resources and labor services.

Eggertsson (1990) refers to the rights of individuals to use resources as property rights. A system of property rights is a method of assigning to particular individuals the authority to select, for specific goods, any use from an unprohibited class of uses. The rights of individuals to the use of resources (i.e., property rights) in any society are supported by the forces of etiquette, social custom, ostracism, and formal laws that are backed up by the states' power of coercion.

It is common to distinguish three categories of property rights: First, there are the rights to use a resource, including the right to transform physically a resource. Second, there is the right to earn income from a resource and contract over the terms with other individuals. Third, there is the right to transfer permanently to another party ownership over a resource—that is, to alienate or sell a resource.

The enforcement of property rights includes excluding others from the use of scarce resources. Exclusive ownership calls for costly measurement and delineation of resources and enforcement of ownership rights. The economic value of exclusive ownership rights depends, *ceteris paribus*, on the costs of enforcing those rights—that is, the costs of excluding others, which ultimately depends on coercion. The enforcement of exclusive rights is usually undertaken by both individual owners and by the state.

An economic problem arises when property rights over a valuable resource—for example, the rights to the air over the factory and the neighborhood—have not been fully delineated. In fact, the dispute between the factory and the neighborhood community involves a struggle over access to a common property resource. Once ownership over the atmosphere is established, the economic problem can be

resolved. In the real business world, we often find that rights to valuable resources are not fully delineated. Reasons for why these property rights are not fully delineated include a weak state, high measurement costs relative to the economic value of a resource, rapid economic change, and struggles over the distribution of wealth.

Property rights to a resource are often partitioned. For example, in the case of land, person A and person B may possess the right to grow wheat on the land. Person C may possess the right to dump ashes on the land. Person D may possess the right to fly an airplane over the land. And each of these rights may be transferable. In sum, private property rights to various partitioned uses of land are “owned” by different persons.

According to the so-called Coase (1960) theorem, the initial partitioning of property rights does not matter for the allocation of resources (ignoring wealth effects) when all rights are freely transferable and the costs of transacting are zero. But when transaction costs are introduced, the role of the state can have a crucial effect on resource allocation. Negotiation costs and other transaction costs may block the reassignment of rights, and the initial partitioning of property rights by the state may have important consequences for the output of an economy. Thus, the property rights approach is not complete without a theory of the state.

Eggertsson (1990) notes that the structure of a contract depends on the legal system, social customs, and the technical attributes of the resources involved in exchange. The more detailed the legal framework and the stronger the ties of custom and social control, the less specific the written contracts. The state, by using its police power and the courts, assists private individuals in enforcing legitimate contracts and thus lowers the costs of exchange, particularly when the state uses its power to enforce contracts in a systematic and predictable manner. In a business world of positive transaction costs, the distribution of political power within a country and the institutional structure of its rule-making institutions are critical success factors in economic development.

Demsetz (1967) offers an optimistic theory of property rights: “Property rights develop to internalize externalities when the gains of internalization become greater than the cost of internalization” (p. 350). Eggertsson (1990) notes that *characteristic of this optimistic view, the formulation of decision making with regard to property rights is solely in terms of private benefits and private costs. The theory does not deal with the free-riding problems that plague group decision, nor is there*

an attempt to model political processes. However, as Libecap (1989) demonstrated earlier in this chapter, the state does not always act to minimize costs and maximize economic value. In particular, the state governments of Texas and Oklahoma failed to design rules that encouraged the unitization of oil fields.

Eggertsson (1990) argues that a rudimentary knowledge of economic history or modern economic systems rules out Demsetz's (1967) optimistic model as a general theory. One of the first steps to modify the optimistic model of property rights involves linking this model to the interest-group theory of legislation and government. Eggertsson (1990) refers to this extension of the optimistic model as the *interest-group theory of property rights*.

The interest-group theory of property rights takes the fundamental social and political institutions of the community as given and seeks to explain the structure of property rights, in various industries, in terms of interactions between interest groups in the political market. Property rights, which serve the narrow self-interest of special interest groups but cause substantial output losses to the community as a whole, typically are explained in terms of transactions costs, free-riding, and asymmetrical information. Eggertsson (1990) concludes (along with North [1990]) that there is overwhelming historical evidence to support the proposition that states typically do not supply structures of property rights that are appropriate for placing the economy close to the technical production frontier.

Firms, Contracts, and Financial Structure (Hart, 1995)

The first four books in this chapter have been in the classical property rights literature. I conclude this chapter with the modern (more formalized) property rights theory (e.g., Grossman & Hart, 1986; Hart & Moore, 1990) and the exemplar work of Hart (1995), *Firms, Contracts, and Financial Structure*. Hart's works (1989, 1995) focus on the boundary and scope of the firm in the market economy and describe an incomplete contracting or property rights approach to both explain and predict firm-level vertical integration decisions. Hart (1995) emphasizes the meaning and importance of asset ownership.

Hart (1995) provides a framework for thinking about firms and other kinds of economic institutions. The basic idea is that firms arise in situations where people cannot write complete contracts and where the allocation of control is therefore important. Given that people write an

incomplete contract, it is clear that revisions and renegotiations will take place. In fact, the contract is seen as a suitable starting point for such renegotiations rather than specifying the final outcome. Hart (1995) submits that because contracts are incomplete, the ex post allocation of control matters. Indeed, these two ideas, contractual incompleteness and the ex post allocation of control, can be used to understand a number of economic institutions.

Property rights theory focuses on how control rights are allocated in a contractual relationship when contracts are incomplete. Hart (1995) notes that in principal-agent theory, it is supposed that it is costless to write a contract. An implication is that an optimal contract will be comprehensive in the sense that the optimal contract will stipulate each person's obligations in every conceivable eventuality and impose large economic penalties if anybody fails to live up to these obligations. Control issues are irrelevant in the principal-agent model since an optimal comprehensive contract will not be renegotiated.

Hart (1995) also observes that transaction costs theory comes closest to the framework of the modern property rights theory. However, although transaction costs theory puts a lot of emphasis on the economic costs of writing contracts and the consequent contractual incompleteness, less attention is paid to the idea that institutional arrangements are designed to allocate control rights among agents.

The Meaning of Ownership. Hart (1995) points out that scholars have written a great deal about why property rights are important and, in particular, why it matters whether a machine, say, is privately owned or is common property. However, there has been less success in explaining why it matters who owns a piece of private property. To understand the difficulty, consider a situation where I want to use a machine initially owned by you. One possibility is for me to buy the machine from you; another possibility is for me to rent the machine from you. If contacting costs are zero, we can sign a rental agreement that is as effective as a change in ownership. In particular, the rental contract can specify exactly what I can do with the machine, when I can have access to it, what happens if the machine breaks down, what rights you have to use the machine, and so on. Given this possibility, however, it is unclear why changes in asset ownership ever need to take place.

In a business context where there are positive transaction costs, however, renting and owning are no longer economically equivalent. If contracts are incomplete, not all the uses of the machine will be specified in all possible eventualities. The economic question then arises:

Who chooses the unspecified uses? A reasonable approach is that the owner of the machine has this property right; that is, the owner has the residual rights of control over the machine. For example, if the machine breaks down or requires modification and the contract is silent about this contingency, the owner can decide how and when the machine is to be repaired or modified. It is now possible to understand why it might make sense for me to buy the machine from you rather than to rent the machine from you. If I own the machine, I will have all the residual rights of control. To put it another way, if the machine breaks down or needs to be modified, I can ensure that the machine is repaired or modified quickly, so that I can continue to use the machine productively. Knowing this possibility, I will have a greater economic incentive to look after the machine, to learn to operate the machine properly, and to acquire other machines that create a synergy with this machine.

The Boundaries of the Firm. Once we recognize that contracts are incomplete and transaction costs are positive, then the boundaries of the firm matter for economic efficiency. Specifically, Hart (1995) argues that firm boundaries are chosen to allocate control rights optimally among the various parties to a transaction. A merger between firms with highly complementary assets enhances economic value. If two highly complementary firms have different owners, then neither owner has real control since neither can do anything without the other. It is better to give all the control rights to one of the owners through a merger.

Agency Theory. Hart (1995) observes that neoclassical microeconomic theory ignores all economic incentive problems within the firm. Over the last 20 years or so, a branch of the organizational economics research literature—principal-agent theory—has developed that tries to rectify this neglect of an essential organizational economic problem. I discuss in more detail principal-agent theory in the next chapter. Hart (1995) argues that principal-agent theory leads to a richer and more realistic portrayal of firms but that principal-agent theory leaves unresolved the basic issue of the determinants of firm boundaries.

Hart (1995) notes that there is now a vast research literature that analyzes the form of the optimal economic incentive scheme under specified circumstances. Moreover, the basic principal-agent problem described has been extended in a number of directions. Among other things, agency theorists have allowed for repeated relationships, several agents, several principals, several dimensions of actions for the agent, career concerns, and reputation effects.

As a result of all this research, a rich set of results about optimal economic incentive schemes has been obtained. However, although these results can throw important light on the determinants of managerial compensation packages and on certain aspects of the organization of production, the agency approach does not pin down the boundaries of the firm (or say much about the internal organization of firms).

Hart (1995) points out that agency theory does not distinguish an optimal contract written by independent firms and internal transfers between divisions of a firm, and yet economically they are quite different. The principal-agent theory is consistent with there being many small, independent firms linked by optimal arm's length contracts, but this theory is also consistent with there being one large firm, consisting of a large number of divisions linked by optimal economic incentive contracts. Clearly, there is something missing from the agency theory of the firm (just as there is something missing from the neoclassical theory of the firm).

The Distinction Between Comprehensive and Incomplete Contracts. Hart (1995) argues that one important factor missing from the principal-agent view is the recognition that writing a (good) contract is itself costly (Coase, 1988; Williamson, 1985). Hart (1995) maintains that although the optimal contract in a standard principal-agent model will not be first-best (since it cannot be conditioned directly on variables like effort that are observed by only one party), the optimal contract in a standard principal-agent model will be comprehensive in the sense that the principal-agent model will specify all parties' obligations in all future states of the world, to the fullest extent possible. As a result, there will never be a need for the contractual parties to revise or renegotiate the contract as the future unfolds. The reason is that, if the contractual parties ever changed or added a contract clause, this change or addition could have been anticipated and built into the original (comprehensive) contract. One would also not expect to see any legal disputes in a comprehensive contracting world. The reason is that, since a comprehensive contract precisely specifies everybody's obligations in every eventuality, the courts should simply enforce the contract as it stands in the event of a dispute.

The Sources of Transaction Costs. Hart (1995) notes that in business reality, contracts are not comprehensive and are revised and renegotiated all the time. According to the transaction costs research literature, renegotiation is a consequence of three factors missing from the standard principal-agent model:

- In a complex and highly unpredictable business world, it is difficult for people to think far ahead and to plan for all the various contingencies that may arise.
- Even if individual plans can be made, it is hard for the contracting parties to negotiate about these plans, not least because the contractual parties have to find a common language to describe states of the world and actions with respect to which prior experience may not provide much of a guide.
- Even if the contractual parties can plan and negotiate about the future, it may be difficult for them to write their plans down in such a way that, in the event of a dispute, an outside authority—a court, say—can figure out what these plans mean and enforce these plans.

Hart (1995) concludes that as a result of these three contracting costs, the parties will write a contract that is incomplete. That is, the contract will contain gaps and missing provisions.

The Economic Implications of Contractual Incompleteness. Hart (1995) notes that, as observed, an incomplete contract will be revised or renegotiated—or both—as the future unfolds. In fact, given that the contractual parties can fill in the gaps as they go along, one may ask why contractual incompleteness matters. The reason is that the renegotiation process imposes several transaction costs. Some of these costs are *ex post* costs incurred at the renegotiation itself, and others are *ex ante* costs incurred in anticipation of renegotiation.

First, the contractual parties may engage in a great deal of haggling over the terms of the revised contract. Argument about division of surplus serves no overall productive purpose, and, to the extent that haggling is time-consuming and wastes resources, such haggling is inefficient. Second, there may be costly legal disputes because an incomplete contract will be ambiguous, and the contractual parties will look to the courts to resolve the ambiguity. Third, not only may the process of *ex post* bargaining be costly, but also, to the extent that the contractual parties have asymmetric information, the contractual parties may fail to reach an efficient agreement.

Hart (1995) argues that if these three costs are high, it must be because there is something binding the partners together and making it difficult for them to switch at the recontracting stage. The leading candidate for that “something” is an *ex ante* relationship-specific investment, that is, a prior strategic commitment, which creates economic value if the contractual parties’ economic relationship extends over time.

Hart (1995) maintains that once the existence of relationship-specific investments is recognized, it becomes apparent that there can be a third cost of contractual incompleteness that may dwarf the haggling and

ex post inefficiency costs. Specifically, because contracts are incomplete, the contractual parties may be deterred from making the relationship-specific (sunk cost) investments that would be optimal in a first-best world. Given each contractual party's fear that the other party will hold it up at the renegotiation stage, the contractual parties are likely to make investments that are relatively nonspecific. Such decisions sacrifice some of the efficiency benefits of specialization, but, in a world of incomplete contracting, these efficiency losses are more than offset by the security that a nonspecific investment provides for each contractual party.

Hart (1995) asks the following: How would these costs change if the two independent (i.e., nonintegrated) firms merged and became a single firm? If there is less haggling and hold-up behavior in a merged firm (as transaction costs theory submits), it is important to provide reasons why. The modern property rights approach developed by Grossman and Hart (1986) and Hart and Moore (1990), the so-called Grossman-Hart-Moore (GHM) model, focuses on this efficient boundaries question.³

The Property Rights Approach. Hart (1995) maintains that (in contrast with the principal-agent approach) the property rights approach tries to address head on the question of why there are less haggling and hold-up problems in a merged firm than between two independent (i.e., nonintegrated) firms. Why does ownership of physical or non-human assets matter? The answer, Hart submits, is that ownership is a source of control rights when contracts are incomplete.

Given that a contract will not specify all aspects of resource usage in every contingency, who has the property rights to decide about missing usages? According to the property rights approach, it is the owner of the resource in question who has these property rights. That is, the owner of a resource has residual control rights over the resource: the property

³The modern property rights approach, discussed in Hart (1995), builds on Grossman and Hart (1986), Hart (1989), and Hart and Moore (1990). Byrnjolfsson (1994) provides an insightful extension and application of the GHM model to information assets and information technology. Both extensions and critiques of the GHM model are many (see, e.g., Holmstrom & Roberts, 1998; Williamson, 2000). For the relevance of the property rights approach to strategic management see, for example, Liebeskind (1996) and Mahoney (1992c). Mahoney (1992c) notes an isomorphism between the Coase (1960) theorem that in the absence of transaction costs liability rules do not matter for achieving economic efficiency and the idea expressed in the previous chapter that, in the absence of transaction costs, organizational form (governance structure) does not matter for achieving economic efficiency. Of course, the main point of Coase (1937) is that in a world of positive transaction costs, organizational form choice does impact economic efficiency, and the main point of Coase (1960) is that in a world of positive transaction costs, initial property rights assignments do impact economic efficiency (as well as income distribution).

rights to decide all usages of the resource in any way not inconsistent with a prior contract, custom, or law. In fact, possession of residual control rights is taken to be the definition of ownership in the modern property rights approach.

Hart (1995) concludes that the economic benefit of integration is that the acquiring firm's economic incentive to make relationship-specific investments increases since, given that the firm has more residual control rights, the firm will receive a greater fraction of the ex post surplus created by these relationship-specific investments. One implication of the property rights theory is that, *ceteris paribus*, a party is more likely to own a resource if he or she has an important (sunk cost) investment decision.

Another strategic implication of the property rights theory is that highly complementary assets should be under common ownership. For example, Joskow (1985) has investigated the ownership arrangements governing electricity-generating plants that site next to coal mines. Such relationship-specific assets are highly complementary, and Joskow finds a high incidence of common ownership. Stuckey (1983) has investigated the case of aluminum refineries that site next to bauxite mines. In this business situation, the degree of complementarity is arguably even greater because, in addition to the two entities being located next to each other, the refinery also installs equipment that is specific to the particular bauxite mine. Stuckey finds that vertical integration occurs in essentially every case. I submit that students studying the economics of organization who provide further case studies along the lines of Joskow and Stuckey that empirically test this modern property rights perspective would enrich the organizational economics research literature.

The Role of Nonhuman Assets and the Nature of Authority in Property Rights Theory. The crucial economic features of the property rights approach are that contracts are incomplete and that there are some significant nonhuman assets in the economic relationship. So far, I have focused on why contractual incompleteness is important to the modern property rights approach. I now discuss why (at least some) nonhuman assets are an essential economic feature of a property rights theory of the firm. These nonhuman assets might include tangible assets, such as machines, inventories, or buildings, or intangible assets, such as patents, brand names, or the firm's reputation.

To understand better the role of nonhuman assets, consider a situation where Firm 1 acquires Firm 2, which consists entirely of human capital. What is to stop Firm 2's workers from quitting? In the absence

of any physical assets (e.g., buildings) Firm 2's workers would not even have to relocate physically. For example, if the workers are linked by telephone or computer terminal (assets that the workers own themselves), workers could announce that they have become a new firm.

For Firm 1's acquisition of Firm 2 to make any economic sense, there must be some source of Firm 2's economic value over and above the workers' human capital. This source of economic value may consist of (a) a place to meet, (b) the firm's reputation, (c) a distribution network (assets that might be relevant to newspapers, journals, or publishing houses), (d) the firm's files containing important information about its operations or its customers (assets that might be relevant for insurance companies or law firms), or (e) a contract that prohibits Firm 2's workers from working for competitors or from taking existing clients with them when they quit (such a contract may be relevant for accounting firms, public relations firms, advertising agencies, or R&D labs, as well as law firms). Thus, a firm's nonhuman assets represent the glue that keeps the firm together.

Hart (1995) notes that it is important to emphasize that there is no inconsistency between defining a firm in terms of nonhuman assets and recognizing that a large part of a firm's economic value derives from human capital. Suppose Firm 2 consists of nonhuman asset a_2 and one worker W_2 . Assume that W_2 can make \$300,000 a year using a_2 and only \$200,000 in its absence, and suppose that W_2 is the only person who knows how to operate a_2 and that the scrap value of a_2 is zero. Then, under the assumption of Nash bargaining, asset a_2 is worth \$50,000 to an acquirer since the acquirer will be able to obtain 50% of W_2 's incremental \$100,000 by threatening to deny W_2 access to the asset. That is, the economic value of the firm to an acquirer is significant even though the value of a_2 in its next-best use (its scrap value) is zero.

Hart (1995) argues that the concept of nonhuman assets is also helpful for clarifying the concept of authority. Coase (1937), Simon (1947), and Williamson (1975) have argued that a distinguishing feature of the employer-employee relationship is that an employer can tell an employee what to do, whereas one independent contractor must explicitly compensate another independent contractor to do what he or she wants. However, as Alchian and Demsetz (1972) point out, the source of an employer's authority over an employee is unclear. It is the case that an employer can tell an employee what to do, but it is also the

case that one independent contractor can tell another independent contractor what to do. The pragmatically interesting question is why the employee acts accordingly, whereas the independent contractor (perhaps) does not pay attention.

When nonhuman assets are present, there is a pragmatic difference between the employer-employee situation and the independent contractor situation. In the employer-employee case, if the employment relationship breaks down, the employer walks away with economically relevant nonhuman assets, whereas in the independent contractor case, each independent contractor walks away with nonhuman assets. This pragmatic difference gives the employer leverage. Put compactly, *control over economically relevant nonhuman resources leads to control over human resources*. This argument connects the behavioral theory of the firm (March & Simon, 1958; Simon, 1947), transaction costs theory (Coase, 1937; Williamson, 1975), and modern property rights theory.

I next discuss an application: the vertical merger of Fisher Body and General Motors (Klein, Crawford, & Alchian, 1978) in light of modern property rights theory. The next chapter covers agency theory. I first discuss the classic work by Berle and Means (1932) concerning the potential agency problem due to the separation of ownership and control. I then develop the basic foundations for the mathematical principal-agent model.

Application: The Vertical Merger of Fisher Body and General Motors

Originally, automobiles were constructed of open wooden bodies. By about 1919, however, closed metal bodies were being manufactured using giant presses to stamp the body parts. Making closed bodies required stamping dies that were in large measure specific to the particular requirements of the model to be produced. In the early period of the automobile industry, the producers of the dies were independent of the automobile manufacturers themselves. Soon after the shift toward closed bodies, which entailed a large specific investment on the part of the die manufacturers, long-term contracts appeared.

Because Fisher Auto Body had to develop specialized production devices that could only be used for General Motors (GM) cars, Fisher Body was reluctant to sign a short-term contract because at renegotiation time, Fisher Body would be at the mercy of General Motors. On the other hand, GM was reluctant to depend so heavily on one supplier, fearing that, with a short-term contract, at renegotiation time, GM would be at the mercy of Fisher Body. Because each party feared a short-term contract would leave it at the mercy of the other firm, Fisher Body and General Motors signed a long-term contract for ten years, according to which GM agreed to buy virtually all of its closed bodies from Fisher Body. This clearly protected Fisher Body from being held up by GM. But now opportunities have been created for Fisher Body to take advantage of GM. At what price would GM buy? Suppose demand conditions change greatly and GM wants to renegotiate the contract? How would quality be assured? Contract negotiations became increasingly complex, until by 1926, the two firms merged as a final attempt to mitigate bargaining difficulties, thereby replacing the transaction costs in the marketplace with internal organization. Vertical financial ownership replaced long-term contracting, which allowed the parties to adjust in an adaptive, sequential manner.

An important aspect of this case, based on the Grossman and Hart (1986) property rights theory of ownership, is that much of the asset specificity came from investment in relationship-specific know-how by the Fisher Body workers, which would have made it difficult for General Motors to find another supplier if Fisher Body had tried to engage in holdup. Thus, vertical integration via financial ownership is persuasively explained in these property rights/transaction costs terms.

SOURCE: Adapted from Klein, Crawford, & Alchian (1978)





4

Agency Theory

Agency theory is an influential approach to the study of corporate governance in strategic management (Eisenhardt, 1989; Kosnik, 1987; Oviatt, 1988; Rediker & Seth, 1995). While Berle and Means (1932) tended to be pessimistic about the economic effects of the separation of ownership and control, modern agency theorists (e.g., Fama, 1980; Fama & Jensen, 1983a; Jensen & Meckling, 1976) shift the emphasis from capital market failures to capital market efficiencies (Jaffe & Mahoney, 1999). Thus, modern agency theory is quite distinct, in this regard, from transaction costs theory.

Modern agency theorists tend to be (overly) optimistic that various governance mechanisms (e.g., the market for corporate control, the market for managers) have solved agency problems. In fact, some agency theorists (e.g., Fama & Jensen, 1983b) make the strong suggestion that these governance mechanisms have reached a high degree of refinement—on which account there is not now, if indeed there ever has been, an organization control problem with which scholars and others are legitimately concerned.

I suggest that students who are studying the economics of organizations should have grave doubts about the overly optimistic (Chicago School) view that agency problems are solved. This view, in my judgment, leads down a blind alley. My suggestion to have doubts parallels my arguments in the previous chapter to be cautious in accepting the overly optimistic (Chicago School) view that property rights typically evolve toward economic efficiency. My own interpretation of the research literature is that the strong ideological fervor of Chicago School economists in espousing the virtues of the market has led these economists more to thinking wishfully than to coming to terms with agency (and property rights) problems that are well documented in the world of experience.

In addition to the descriptive agency theory of Jensen and Meckling (1976) and Fama and Jensen (1983a), there is also a more formal

branch of agency theory where mathematical principal-agent models (e.g., Holmstrom, 1979) are developed. For the purposes of this research book, we focus more on the principal-agent model than on the descriptive theory. Another section in this chapter, however, covers the classic paper by Jensen and Meckling (1976).

The mathematical principal-agent problem in its moral hazard form stems from a basic conflict between insurance and economic incentives. On the one hand, the theory of optimal insurance demonstrates that the optimal division of a pie of a random size (e.g., the profit) between a risk-neutral party (the shareholders) and a risk-averse one (e.g., the manager) has the risk-neutral party bearing all the risk, if economic incentives are left aside (see Tirole, 1988).

On the other hand, if the risk-averse party (the manager) is given an income that does not depend on effort, then the risk-averse party (the manager) has no economic incentive to exert effort. Thus, the goal of providing full insurance conflicts with the goal of providing economic incentives. Specifically, for a risk-averse agent the insurance objective conflicts with the residual claimant status. Indeed, the trade-off between providing insurance and providing economic incentives generally leaves the contractual parties with both suboptimal insurance and suboptimal profits (a second-best contract).

In summary, the basic principal-agent model discussed in this chapter shows that efforts, if they are not observed, must be induced through economic incentives. The manager's compensation must grow with the realized profit. Because such economic incentive structures destroy insurance, the expected compensation is higher under nonobservability. This fact, in turn, may make the principals (the shareholders) not wish to induce effort. Thus, the principals (the shareholders) may tolerate slacking.

Before focusing on this mathematical principal-agent model, I begin with a classic statement of the agency problem by Berle and Means (1932). I then consider the mathematical principal-agent model in Pratt and Zeckhauser (1985), Arrow (1985), and Levinthal (1988). Finally, I discuss positive (i.e., descriptive) agency theory. Because the Chicago School optimistic view is mentioned throughout this chapter, the classic paper by Jensen and Meckling (1976) is included in the section titled "Theory of the Firm: Managerial Behavior, Agency Costs, and Capital Structure."

Berle and Means (1932) argue that the separation of ownership and control exists in varying degrees and that this separation becomes almost

complete when not even a substantial minority shareholder exists. Berle and Means note, for example, that in the American Telephone and Telegraph Company, the largest shareholder owned less than 1% of the company's stock. Berle and Means raise concerns about the divergence of economic interests between shareholders and managers.

Pratt and Zeckhauser (1985) maintain that the issue of the separation of ownership and control is but a subset of a host of economic problems that can be classified as principal-agent problems. Indeed, the theme of Pratt and Zeckhauser's work is that businesses, workers, and consumers regularly struggle to deal with agency problems. Overall, Pratt and Zeckhauser take a more optimistic view than Berle and Means (1932) concerning the severity of the agency problem, although Pratt and Zeckhauser do acknowledge the possibility that with creativity in the governance area we could do better.

Similar to Pratt and Zeckhauser's (1985) arguments, Arrow (1985) notes that the agency relationship is a pervasive fact of economic life and that the principal-agent relationship has significant scope and economic magnitude. Arrow makes the useful distinction between hidden action models (e.g., moral hazard models) and hidden information models (e.g., adverse selection models). Arrow also considers the roles of multiple principals, multiple agents, monitoring, and repeated relations in attenuating the agency problem. Finally, Arrow considers the extent to which the principal-agent relationships in actuality differ from that developed in the models. Most important, the theory tends to predict very complex fee functions to align economic interests between principals and agents. In fact, we do not observe such complex relations in reality. Although not mentioned by Arrow, one obvious explanation of this divergence is that the mathematical-principal agent model assumes away the problem of bounded rationality.

Levinthal (1988) characterizes and critiques the research on agency models of organizations to broaden the set of readers of such models and to stimulate the production of new research. In many ways, this article highlights the ways in which the strategic management field can contribute to the evolving science of organization, using specific mathematical functional forms that highlight the implications of these models. Moreover, Levinthal skillfully uses ideas from Simon (1947), March and Simon (1958), and Cyert and March (1963) to critique agency models from the perspective of the behavioral theory of the firm. A question left open to current students engaged in the evolving science of organization is whether agency theory and the behavioral theory of

the firm will proceed in a dialectical manner, with each posing questions and problems for which the other perspective must respond, or whether a more synthetic approach will emerge.

Finally, I conclude with the Chicago School view espoused by Jensen and Meckling (1976). Jensen and Meckling note that if a manager has no debt or equity holdings in a firm, then the manager will bear none of the costs directly. In this situation, managers clearly face an incentive to make decisions that are contrary to the best, wealth-maximizing interests of outside investors. Of course, outside investors may engage in a wide variety of activities to monitor the actions and decisions of managers. To go further, managers can post economic bonds whereby they will be penalized economically should they make decisions that violate the interests of outside investors. These continuing conflicts between managers and investors, together with the monitoring and economic bonding mechanisms designed to reduce these conflicts, are costly.

Jensen and Meckling (1976) refer to the sum of these costs as agency costs. Total agency costs consist of monitoring expenditures made by the principal to regulate and monitor the behavior of the agent, bonding expenditures made by the agent to reassure principals, and residual agency costs, or costs due to unresolved conflicts of interest between agents and principals—this third category being a very expansive category. Jensen and Meckling argue that in a world of efficient capital markets, managers in an attempt to raise funds from outside sources bear all the economic burden of these agency costs. Jensen and Meckling use these conclusions to investigate the possibility of an optimal capital structure for a firm. Because managers bear all the agency costs in their attempt to raise capital, these managers have a strong economic incentive to choose the capital structure that minimizes total agency costs.

Thus, Jensen and Meckling (1976) argue that although the separation of ownership and control lessens profitability incentives, that incentive issue is fully anticipated at the time the separation of ownership from control occurs and is therefore fully reflected in the price of new shares. The future therefore holds no surprises; all of the relevant contracting action is packed into *ex ante* incentive alignments.

Students should note that Jensen and Meckling (1976) examine the consequences of diluting a 100% equity position in an entrepreneurial firm. However, their real theoretical interest is the diffusely owned modern corporation, but the basis for moving from one scenario to the other scenario is not described. Jensen and Meckling (1976) expressly acknowledge this condition:

One of the most serious limitations of this analysis is that as it stands we have not worked out in this paper its application to the very large modern corporation whose managers own little or no equity. We believe our approach can be applied to this case but . . . [these issues] remain to be worked out in detail. (p. 356)

As a general criticism, Williamson (1996), commenting on Jensen and Meckling (1976), notes that “the logic that connects tractable micro[economic] models and the composite uses to which they are put is often asserted but is rarely fully worked out” (p. 188). Clearly, challenges to contemporary students are posed.

The Modern Corporation (*Berle & Means, 1932*)

Berle and Means (1932) submit that corporations have ceased to be merely legal devices through which the private business transactions of individuals may be carried on and, in fact, have become both a method of property tenure and a means of organizing economic life.

For Berle and Means (1932), the direction of industry by persons other than those who have ventured their own wealth has raised the question of the motive force behind such direction and the effective distribution of the economic returns from business enterprise. The private corporation has given way to an essentially different form, the quasi-public corporation, a corporation in which a large measure of the separation of ownership and control has taken place through the multiplication of owners.

Berle and Means (1932) argue that the separation of ownership and control produces a condition where the interests of owner(s) and managers may, and often do, diverge and where many of the checks that formerly operated to limit the use of such discretionary managerial power disappear. Physical control over the instruments of production has been surrendered in ever growing degrees to centralized groups who manage the property in bulk, supposedly, but by no means necessarily, for the economic benefit of the security holders.

Berle and Means (1932) maintain that it has been assumed that, if individuals are protected in the rights both to use their property as they see fit and to receive the full fruits of its use, their desire for personal gains, for profits, can be relied on as an effective economic incentive to their efficient use of any industrial property they may possess. In the quasi-public corporation, such an assumption no longer holds.

Berle and Means (1932) take the position that where the separation of ownership and control is substantial, “control may be held by the

directors or titular managers who can employ the proxy machinery to become a self-perpetuating body, even though as a group they own but a small fraction of the stock outstanding" (p. 5). Berle and Means argue that ownership of wealth without appreciable control, and control of wealth without appreciable ownership, appear to be the logical outcome of corporate development. Berle and Means ask:

Have we any justification for assuming that those in control of the modern corporation will also choose to operate it in the interests of the owners? The answer to this question will depend on the degree to which the self-interest of those in control may run parallel to the interests of ownership and, insofar as they differ, on the checks on the use of power which may be established by political, economic or social conditions. (p. 121)

Berle and Means (1932) in their empirical study found that 88 of the 200 largest American nonfinancial corporations were management controlled in 1929 because no individual, family, corporation, or group of business associates owned more than a 20% share of all outstanding voting stock and because evidence of control by a smaller ownership group was lacking. Berle and Means judged only 22 of the corporations to be privately owned or controlled by a group of stockholders with a majority interest.

Berle and Means (1932) note that it is traditional for a corporation to run for the economic benefit of its owners, the stockholders, and that any profits should be distributed to stockholders. However, Berle and Means state that a controlling group may hold the power to divert profits into their own pockets, submitting that there is "no longer any certainty that a corporation will in fact be run primarily in the interests of the stockholders" (p. 333).

Berle and Means (1932) conclude that *the rise of the modern corporation has brought a concentration of economic power that can compete on equal terms with the modern state—economic power versus political power*. The state seeks, in some respects, to regulate the corporation, whereas the corporation, steadily becoming more powerful, makes every effort to avoid (or shape) such regulation. Where its interests are concerned, the modern corporation even attempts to dominate the state. The future may see the economic organization, typified by the corporation, not only on an equal plane with the state but possibly even superseding the state as the dominant form of social organization. The law of the corporation, accordingly, might well be considered a potential constitutional law for the new economic state, whereas

business practice is increasingly assuming the aspect of economic statesmanship.

“Principals and Agents: An Overview” (*Pratt & Zeckhauser, 1985*)

We next consider an overview of the principal-agent research literature by Pratt and Zeckhauser (1985). This article provides us with the reason why agency theory is relevant to contemporary management practice. Pratt and Zeckhauser emphasize that the real world has been inventive in developing subtle mechanisms, such as career expectations and product reputations, to overcome the difficulties posed by informational asymmetries. Thus, mathematical model builders need to learn from the world of business practice, and the insights provided by model builders can reciprocate providing insights for managerial practice.

The Agency Relationship. Pratt and Zeckhauser (1985) note that a predominant concern for an economy, discussed since the time of Adam Smith (1776/1937), is to ensure that production is conducted in the most efficient manner, taking advantage of the economic benefits of specialization while appropriately conserving scarce resources. But even if we could figure out, or were willing to let the market figure out, the most efficient way to produce goods and services, there would be the problem of ensuring that each individual performs his or her agreed-on task. Full information rarely is freely available to all contractual parties, so the economic problem becomes how to structure an agreement that will induce agents to serve the principal’s interest even when the principal does not observe the agents’ actions and information.

Whenever one individual depends on the action of another, an agency relationship arises. The individual taking the action is called the *agent*. The affected party is the *principal*. In common parlance, the doctor is the agent, and the patient is the principal. The corporate executive is the principal, and the subordinates are the agents. The corporate executive, in turn, is an agent for the shareholders.

Challenges in the agency relationship arise whenever the principal cannot perfectly and costlessly monitor the agent’s action and information. The problems of inducement and enforcement then come to the fore. Given information asymmetries—agents typically know more about their tasks than their principals do—we cannot expect any business enterprise or business institution to function as well as it would if all information were costlessly shared or if the economic incentives

of principals and agents could be costlessly aligned. This shortfall is sometimes called the *agency loss* or *agency costs*. The economic challenge in structuring an agency relationship is to minimize agency costs. In economic parlance, since the first-best outcome could be achieved only in the unrealistic world of costless information flow, our goal must be to do the best we can, to achieve what is sometimes called the *second-best solution*. Pratt and Zeckhauser (1995) note that businesses, workers, consumers, and indeed all participants in society at large regularly struggle to deal with the intractable problems that arise in agency relationships, that organizational forms evolve to deal with such agency costs, and, on average, these organizational forms perform reasonably well. The building blocks of agency theory are information and economic incentives.

Information. Pratt and Zeckhauser (1985) note that in most social and business relationships, the contractual parties have different information available to them. Relationships vary in the degree of informational asymmetry they involve. At one extreme is the perfect-market transaction, with standardized products and full disclosure of all information. At the other end of the continuum are situations in which the agent has full discretion and is not observed at all by the principal.

Monitoring. Pratt and Zeckhauser (1985) provide several agency theory predictions concerning monitoring:

- We tend to get less monitoring, or monitoring of poorer quality, when monitoring is expensive or when substitutes for monitoring are cheap, or both.
- The agency loss is more severe when the economic interests or economic values of the principal and agents diverge substantially and information monitoring is costly.
- In a range of real-world situations, more limited monitoring—say of an indicator of output—is relatively or fully successful.
- A large stock of economic value that could be lost through bad behavior, such as reputation or assets subject to legal suit, is a strong economic incentive for good behavior.
- Long-term relationships, among other benefits, develop the stocks of economic value needed for enforcement and make limited monitoring more effective.
- The economic benefits of any reductions in agency loss will be shared by principal and agent in most market situations.
- The principal and agent have a common economic interest in defining a monitoring-and-incentive structure that produces economic outcomes as close as possible to the economic outcome that would be produced if information monitoring were costless.

Incentives. Pratt and Zeckhauser (1985) note that each person performs his or her task in Adam Smith's (1776/1937) pin factory. The specialization of labor that has created the possibilities for productive modern industrial societies has also created the need for organizations larger than the crafts shop or family. The new challenge becomes how to motivate the participants within the organizations that make up society so that these participants will be as productive as they would be if they were the owners.

Pratt and Zeckhauser (1985) argue that the simple prescription of neoclassical microeconomic theory is difficult to apply in business practice. It is difficult to measure an individual's marginal product. If Company A's sales go up, is it because of the perseverance of the salesperson, the workers' greater attention to quality, or blind luck? The problem of measuring marginal product has become so intractable that the vast majority of Americans are paid based on salaries. If work input is difficult to measure, the employee may have an economic incentive to shirk. An indicator of inputs, such as hours on the job, is likely to be used in evaluating the employee instead of a measure of actual inputs, which would include the diligence and quality of those hours. Contemporary wage compensation systems, in which compensation is not closely tied to productivity, pose a sizeable agency problem for the motivation of workers.

Pratt and Zeckhauser (1985) maintain that the corporation and the modern economy have found ways to respond to this separation of rewards and productivity. Part of the solution is monitoring by supervisors. Promotions, for example, offer an individual higher future wages for the indefinite future. And dismissal typically imposes a major cost on the worker. If motivation is particularly important and monitoring is costly, then wages may be set above the amounts workers could earn elsewhere; the small chance of termination will then be an influential economic incentive (Klein & Leffler, 1981).

Pratt and Zeckhauser (1985) note that, unlike labor, capital need not be motivated. But informational asymmetries still can create agency problems for capital. The quality of a machine can be overstated—consider the used car or the new computer. Machines that can be abused or overused, but whose condition is difficult to monitor, are sold rather than leased. And a production process is likely to be vertically integrated if there are severe quality uncertainties at various supply-chain stages.

Modern accounting attempts to keep track of capital and the profits such capital produces. However, accounting techniques cannot

accurately assess many contributors to a firm's long-term profitability, such as reputation for quality, condition of equipment, or research accomplishments. An agency loss may be the consequence, with managements pursuing measured outputs, such as reported profits, at the expense of those accomplishments that are difficult to tally. Some critics of modern business enterprises have identified this economic distortion as a major factor contributing to lagging American productivity.

Pratt and Zeckhauser (1985) emphasize that the types of informational asymmetries that lead to agency loss make it difficult to reward labor or capital with the economic value of its product. Having the factors of production in a single business entity, such as a corporation, is advantageous. Long-term residence there allows for more creative reward structures, such as internal promotion opportunities. Long-term relationships also encourage individuals to work toward a common purpose.

The Design of Agency Structures. Pratt and Zeckhauser (1985) note that in considering the agency problem, it is perhaps natural to focus on the question of how the principal can reap the greatest advantage through economic incentives that influence agents' behavior yet reward the agents enough so that they will not quit. Two central assumptions are implicit in this formulation: first, that the principal is in a position to design the monitoring and incentive mechanism and, second, that all the economic benefits from improvements in performance go to the principal. Pratt and Zeckhauser argue that agency situations often satisfy neither of these assumptions and that the most important issues do not depend on them. In fact, the agent and principal are merely two (or more) individuals (or organizations) in some sort of explicit or implicit contractual relationship. Oversight and reward mechanisms may be designed by either the agent (e.g., most lawyers define the terms of their relationships with clients) or the principal (e.g., the store manager probably sets the employment conditions for retail clerks). There may be a joint negotiation, as I elsewhere show with labor agreements under collective bargaining. Or some external party may set the terms, as the government does in structuring many of the rules by which corporations are governed in the United States.

Pratt and Zeckhauser (1985) conclude that *while agency theory has provided modest aid in improving business practice, the real world of business has been inventive in developing subtle mechanisms such as career expectations and product reputations to overcome the difficulties posed by informational asymmetries.* Also, human environments can

change quickly, and there is no assurance that the institutions we currently observe are best. Therefore, sound conceptual thinking concerning institutions by students studying the economics of organization can be worthwhile for reducing agency loss in the future.¹

“The Economics of Agency” (Arrow, 1985)

We next consider Arrow’s (1985) review of the principal-agent model. Arrow also concludes with a thoughtful evaluation of the agency model.

Arrow (1985) observes that the agency relationship is a pervasive fact of economic life. The principal-agent relation is a phenomenon of significant scope and economic magnitude. A common element for early principal-agent models is the presence of two individuals. One (the agent) must choose an action from a number of alternatives. The action affects the economic welfare of both the agent and another person, the principal. The principal, at least in the simplest cases, has the additional function of prescribing payoff rules; that is, before the agent chooses the action, the principal determines a rule that specifies the fee to be paid to the agent as a function of the principal’s observations of the results of the action.

The economic literature has focused primarily (but not exclusively) on the case in which the agent’s action is not directly observable by the principal, and the outcome is affected but not completely determined by the agent’s action. Were it not for the second condition, the principal could infer the agent’s action by observing the outcome. In more technical language, the outcome is a random variable whose distribution depends on the actions taken.

Arrow (1985) notes that, more generally, a single principal may have many agents. Each agent takes an action, and the output of the system is a random function of all the actions. The principal cannot observe

¹For example, a number of mechanisms have evolved to attenuate the agency problem—the market for corporate control, the market for managers, use of outside board of directors, monitoring by institutional investors, concentrated ownership, and so on (see Mahoney, 1992c). For both economic and strategic management theoretical and empirical contributions to the agency theory literature, see Eisenhardt (1985, 1989); Fama (1980); Fama and Jensen (1983a, 1983b); Holmstrom (1979); Jensen and Meckling (1976); Kosnik (1987); Lajili, Barry, Sonka, and Mahoney (1997); Mahoney and Mahoney (1993); Mahoney, Sundaramurthy, and Mahoney (1996, 1997); Rediker and Seth (1995); Seth (1990); Seth and Thomas (1994); Sundaramurthy and Lewis (2003); Sundaramurthy, Mahoney, and Mahoney (1997); Walsh and Seward (1990); and Zenger (1994).

the actions themselves but may make some observations, for example, of the output. Again, the principal sets in advance a schedule stating the fees to be paid to the individual agents as a function of the observations made by the principal. The principal-agent theory is in the standard economic tradition. Both principal and agent are assumed to be making their decisions optimally in view of their constraints, and intended transactions are realized. As is usual in economic theory, the theory functions both normatively and descriptively. The theory offers insights used in the construction of contracts to guide and influence principal-agent relations in the real business world, and at the same time the theory represents an attempt to explain observed phenomena in the empirical economic world, particularly exchange relations that are not explained by more standard economic theory.

Arrow (1985) finds it useful to distinguish two types of agency problems, hidden action (e.g., moral hazard) and hidden information (e.g., adverse selection). A typical hidden action is the effort of the agent. Effort is a disutility to the agent, but the agent's effort has an economic value to the principal in the sense that such effort increases the likelihood of a favorable outcome (technically, the distribution of the outcome to a higher effort stochastically dominates that to a lower effort; that is, the probability of achieving an outcome that exceeds a given level is higher with higher effort). An example of hidden action is the relation between stockholders and management. The stockholders are principals who cannot observe in detail whether the management, their agent, is giving high effort and is making appropriate decisions.

Fire insurance dulls economic incentives for caution and even creates economic incentives for arson; this economic problem is the origin of the term *moral hazard*. Health insurance creates similar economic problems because health insurance creates an economic incentive to use excessive medical care. Employment relationships typically are relationships in which effort and the ability acquired through training and self-improvement are difficult to observe. In one view, firms exist as a means of measuring effort (Alchian & Demsetz, 1972). In hidden information problems, the agent has made some observation that the principal has not made. The agent uses (and should use) this observation in making decisions; however, the principal cannot check whether the agent has used his or her information in the way that best serves the principal's interest. A case much studied from various points of view in the economic research literature is that of a decentralized socialist economy. The productive units may well have economic incentives not to reveal their

full potentiality because it will be easier to operate with less stringent requirements. The economic problem for the central planning unit (the principal) is how to tap the agent's information. A similar problem occurs in decentralization within a firm. This topic in the literature has acquired the name of incentive compatibility (Miller, 1992).

The problem of adverse selection was originally noted in insurance of several kinds. The population being insured is heterogeneous with respect to risk; in the case of life insurance, for example, some have a higher probability than others of dying young. In at least some cases, those who are insured have better knowledge of this probability than the insurance company, which is unable to differentiate. If the same premium is charged to everyone, then the high-risk individuals will purchase more insurance, and the low-risk individuals will purchase less insurance. This outcome leads to an inefficient allocation of risk bearing. Arrow (1985) notes that the hidden information principal-agent model becomes more complicated when multiple principals compete for agents.

The Hidden Action Model. Arrow (1985) provides a simple formulation of the hidden action model. The agent (for the moment, assume there is only one) chooses an action a . The result of this choice is an outcome x , which is a random variable whose distribution depends on a . The principal has chosen beforehand a *fee function* $s(x)$, to be paid to the agent. For the simplest case, assume that the outcome x is income; that is, a transferable and measurable quantity. Then the net receipts of the principal will be $x - s(x)$. Because the principal and agent are both, in general, risk averse, each values whatever income he or she receives by a utility function with diminishing marginal utility. Let U be the utility function of the principal and V that of the agent. Further let $W(a)$ be the disutility the agent attaches to action a . It will be assumed separable from the utility of income; that is the marginal utility of income is independent of the action taken (the amount of effort). Note that the action is taken before the realization of the uncertainty and is therefore not uncertain to the agent, though the agent's action is unknown to the principal.

Because, even for a given action, the outcome x is uncertain, both principal and agent are motivated to maximize the expected value of their own utility. Given the principal's choice of fee function $s(x)$, the agent wishes to maximize the expected value of $V[s(x)] - W(a)$. In effect, therefore, the principal can predict the action taken for any given fee schedule. The choice of fee schedules is, however, restricted by

competition for agents, who have alternative uses for their time. Hence, the principal must choose a fee schedule that offers the agent a utility at least equal to what the agent could achieve in other activities (i.e., there is a participation constraint). *The principal-agent problem combines two inextricable elements: risk sharing and differential information.* Even if there was no economic problem of differential information, there would be some sharing of the outcome if both contractual parties are risk averse. Indeed, if the agent was risk neutral, the principal-agent problem would have a straightforward solution, the agent would bear all the risks, and then the differential information would not matter. That is, the principal would retain a fixed amount and would pay the remainder to the agent, who therefore would have no dilution of economic incentives. In the terminology used previously, the fee function would equal the outcome less a fixed amount, $s(x) = x - c$, where the constant c is determined by the participation constraint. Thus, a landlord renting land to a tenant farmer would simply charge a fixed rent independent of output, which in general depends on both the tenant's effort, which is unobservable to the landlord, and the vagaries of the weather.

However, this economic solution is not optimal if the agent is risk averse. Because all individuals are averse to sufficiently large risks, the simple solution of preserving economic incentives by assigning all risks to the agent fails as soon as the risks are large compared with the agent's wealth. For example, the president of a large company can hardly be held responsible for the company's income fluctuations.

In the general case of a risk-averse agent, the fee will be a function of the outcome—to provide economic incentives—but the risk will be shared. Generally, there is a trade-off between economic incentives and the efficiency of the system (considering both principal and agent). For a business application, consider the case of insurance with moral hazard. Some insurance will be written, but insurance will not be complete. In the terminology of the insurance industry, there will be coinsurance; that is, the insured will bear some of the economic losses against which the insurance is written. Coinsurance is customary in health insurance policies where the insured has considerable control over the amount of health expenditures.

Monitoring. Arrow (1985) notes that some principal-agent research literature has emphasized the possibility of monitoring. By this is meant the possibility that the principal has certain information in addition to the outcome. If this observation y is costless and conveys any information

about the unobserved action a beyond that revealed by the outcome, x (technically, if x is not a sufficient statistic for the pair x, y with respect to action a), then the principal can always improve the second-best contract by making the fee depend on y as well as x .

Multiple Agents and Repeated Relations. Arrow (1985) maintains that new possibilities for economic incentives arise when there are many agents for a single principal or, alternatively, repeated relations between agent and principal. One can use the ordinal ranking of agents' outcomes as a basis for fees.

Repeated relations between a principal and agent provide new opportunities for economic incentives. Experience rating in insurance illustrates the situation; the premium rate charged today depends on past outcomes. In effect, the information on which the fee function is based is greatly enriched.

Evaluation of Agency Theory. Arrow (1985) submits that principal-agent theory gives a good reason for the existence of sharecrop contracts, but principal-agent theory is a very poor guide to their actual terms. Indeed, as John Stuart Mills pointed out long ago, the terms tend to be regulated by custom. Sharecrop contracts are remarkably uniform from farm to farm and from region to region. Principal-agent theory, in contrast, suggests that the way the produce is divided between landlord and tenant would depend on the probability distribution of weather and other exogenous uncertainties and on the relation between effort and output, both of which certainly vary from one region to another. The relation between effort and output would vary over time as well. Similarly, the coinsurance provisions in health insurance policies are much simpler than could possibly be accounted for by principal-agent theory.

Arrow (1985) observes that in some cases where principal-agent theory seems clearly applicable, real-world business practice is very different from the model. In many respects, the physician-patient relation exemplifies the principal-agent relationship almost perfectly. The principal (the patient) is certainly unable to monitor the efforts of the agent (the physician). The relation between effort and outcome is random, but presumably there is some connection. Yet, in practice, the physician's fee schedule is in no way related to outcome. In general, compensation of professionals shows only a few traces of the complex fee schedules implied by agency theory.

Why is this divergence between theory and practice so stark? Arrow (1985) argues that one basic problem is the costs of specifying complex

relations. Second, superiors judge executives on criteria that could not have been stated in advance. Outcomes and even supplementary objective measures simply do not exhaust the information available on which to base rewards. A third limitation of the present model is the restricted reward or penalty system used. It is virtually always stated in terms of monetary payments. *Further extensions of the agency model are needed to capture some aspects of reality because there is a whole world of rewards and penalties that take social rather than monetary forms.* Arrow (1985) submits that professional responsibility is clearly enforced in good measure by systems of ethics, internalized during the education process, and enforced in some measure by formal punishments and more broadly by reputations. Ultimately, of course, these social systems have economic consequences, but they are not the immediate ones of current principal-agent models.

I conclude the agency chapter with an application by Wolfson (1985), followed by a section that discusses Levinthal's (1988) instructive article on principal-agent models and their limitations, and Jensen and Meckling's (1976) classic article on agency theory.

Application: Empirical Evidence of Incentive Problems and Their Mitigation

Wolfson (1985) examines the oil and gas industry, a business context with high agency problems due to asymmetric information. Moreover, the United States tax code has encouraged investment in oil and gas drilling by limited partnerships—a classic agency relationship. The limited partners (the principals) put up most of the capital, whereas the general partner (the agent) makes most of the decisions. The limited partners are typically far from the scene of the action, and most of them have no experience or expertise in the business.

There are severe conflicts of interests in this business context. The general partner frequently sells services to his or her limited partners. The general partner may also have related investments nearby and may benefit personally from the information gained from drilling. Moreover, the partnership typically imposes all of the immediately deductible costs on the limited partners (to capitalize on the tax advantages), whereas completion costs are shared

between the limited and general partner, or may even be the full responsibility of the general partner. Therefore, the general partner has too much economic incentive to drill and, for some revenue splits, too little economic incentive to complete.

The agency problems in drilling are well known to both operators and limited partners. Prospectuses detail these agency problems at length and point out that monitoring by the limited partners is not really feasible. To reassure prospectus investors, general partners explain that they are most concerned about maintaining their long-term business reputation.

Wolfson (1985) tests a number of empirical assertions about the economic outcomes that, based on agency theory, we might expect in this business context. For example, the problem of distorted economic incentives is less severe in exploratory than developmental drilling. Hence, many limited partnerships should operate in the business context of exploratory drilling, and this economic outcome does hold empirically.

Wolfson (1985) also investigates the relationship between reputation and reduced agency loss. Wolfson shows empirically that if an operator general partner can build a reputation for success, then the general partner can secure investment funds on more favorable economic terms. To build reputation, the general partner takes actions to benefit the limited partners at the expense of the general partner's own immediate profitability. Wolfson's data reveal that operator general partners with records of success for limited partners do receive more favorable economic terms from investors. Wolfson shows that even, or perhaps particularly, in the potentially opportunistic business environment of oil and gas drilling, business reputations can help overcome contracting difficulties and the resulting economic inefficiencies.

SOURCE: Adapted from Wolfson (1985)

"A Survey of Agency Models of Organizations" (Levinthal, 1988)

Levinthal (1988) provides the insightful perspective that agency theory can be viewed as the neoclassical economic response to the questions raised many years earlier by March and Simon (1958) and Cyert and

March (1963) regarding the behaviors of an organization of self-interested agents, with conflicting goals, in a world of incomplete information.

An agency relationship is said to exist between two (or more) parties when one contractual party, designated as the agent, acts on behalf of another contractual party, designated as the principal. For example, in an attempt to bring management's interests in line with those of stockholders, top management is given complex compensation packages, consisting of salary, performance bonuses, stock ownership, and pension plans.

Agency theory views the economic problem of contract design as maximizing the economic payoff to the principal, taking several factors into account:

- The relationship between output and the economic incentive scheme offered
- The allocation of risk associated with different compensation schemes
- The preferences of the principal and agents with respect to income and nonpecuniary outcomes

Agency Model

Consider an organization consisting of two utility maximizing individuals. One is the owner, the other a hired manager. The owner's economic problem is to design a compensation package that elicits an appropriate effort level from the manager. This economic problem is complicated by the fact that the owner only has limited information. The owner only observes the firm's revenue, which is a function of two factors that the owner cannot directly observe: the manager's effort and some exogenous variable such as consumer demand. Given the observed revenue level, the owner can only make probabilistic inferences about the manager's effort level. To make the example more concrete, suppose that Table 4.1 indicates the relationship between revenue, the manager's effort a , and the exogenous state variable S_i .

For instance, if the manager exerts high effort (high) and state S_4 occurs, then the cash flow is \$25,000. The expected cash flow from the agent's effort level is derived by assuming that the four states are equally likely to occur; that is, the expected cash flow corresponding to a high level of effort is computed as $0.75 (\$50,000) + 0.25 (\$25,000) = \$43,750$. The agent's effort is assumed to be productive, in the sense that the expected cash flow increases as the agent's effort level rises from low to medium to high.

More generally, we can think of the observed outcome (revenue in the previous example) as a random variable whose distribution depends on

Table 4.1 Calculations of Expected Cash Flow

	State				Expected Cash Flow
	S_1	S_2	S_3	S_4	
Effort Level (a)					
High	50,000	50,000	50,000	25,000	43,750
Medium	50,000	25,000	50,000	25,000	37,500
Low	50,000	25,000	25,000	25,000	31,250

the action (effort) chosen by the agent. Let $f(x/a)$ represent a probability distribution of x , conditional on the agent's action. Table 4.1 presents an example of a particular distribution of $f(x/a)$. The usefulness of the agent's efforts is modeled by the assumption that increased effort shifts the distribution of outcomes to the right in the sense of first-order stochastic dominance. This characteristic of the model is illustrated in Table 4.1, where for any state S_i , increased effort weakly increases revenue (i.e., the revenue either increases or stays the same).

Because the owner cannot directly enforce a particular effort level, the owner must influence the agent's self-interest by the choice of compensation scheme. This choice of economic incentive structure is also constrained by the fact that the owner cannot force the manager to work for the firm. As a result, the manager's alternative employment opportunities must be considered—a participation constraint, to use Arrow's (1985) term.

To characterize the optimal compensation scheme for this example, we must specify the manager's preferences. It is assumed that the manager has preferences over income and effort represented by a utility function $W(z, a) = U(z) - V(a)$, where z is the level of income and a is the level of effort. It is assumed that the utility function is separable into the utility for money $U(z)$ and the disutility of effort $V(a)$.

In particular, $U(z) = \sqrt{z}$ and

$$V(a) = V(\text{high}) = 40$$

$$V(\text{medium}) = 20$$

$$V(\text{low}) = 5$$

The function $U(z)$ is one in which the agent's utility increases with z but at a diminishing rate (i.e., this is a concave function) and implies that the agent is risk averse. Second, for a given income level, the agent's utility is decreasing in the agent's effort level. More generally in agency analyses, it is assumed that $U(z)$ is concave and $V(a)$ is convex with respect to the continuous effort variable. The agent is risk averse, and the agent's disutility of effort increases at an increasing rate. In addition, the agent's opportunity cost of working on behalf of the principal must be stated. In our example, it is assumed that the agent's best alternative yields a utility measure of 120 units.

Finally, to complete the model, the owner's preferences must be specified. It is assumed that the owner derives utility only from the net earnings (i.e., the cash flow specified in Table 4.1 less the agent's compensation). The owner's utility function is represented by $G(\cdot)$; in the example, it is assumed that the principal is risk neutral and $G(x) = x$.

Suppose that the owner wants to elicit an effort of *high* from the manager. The owner faces two constraints. First, the contract must yield the manager an expected utility of at least 120 units if the manager is to work in the owner's behalf. Second, it must be in the manager's own interest to choose an effort level of high over medium over low. The only device under the owner's control is the compensation scheme that the owner offers the manager. The agent's share of compensation is not a fixed payment but a schedule of payments expressing the agent's reward as a function of the revenue of the firm. The agent's compensation is represented by the sharing rule $z(x)$. Let Z_{50} be the payment to the manager if the outcome $x = \$50,000$ is observed, and Z_{25} the payment if the outcome $x = \$25,000$ is observed. The expected utility of a manager who sets $a = \text{high}$ is the following:

$$E [U(\text{high})] = 0.75(\sqrt{z_{50}}) + .25(\sqrt{z_{25}}) - 40$$

$$E [U(\text{medium})] = 0.50(\sqrt{z_{50}}) + .50(\sqrt{z_{25}}) - 20$$

$$E [U(\text{low})] = 0.25(\sqrt{z_{50}}) + .75(\sqrt{z_{25}}) - 5$$

If the manager is to find it to his or her interest to choose $a = \text{high}$ over the other alternative effort levels, then the following relationships must be satisfied:

$$E [U(\text{high})] > E [U(\text{medium})] \text{ and}$$

$$E [U(\text{high})] > E [U(\text{low})]$$

In addition, if the manager is to accept employment with the owner then

$$E[U(\text{high})] > 120$$

We can write the owner's decision problem as the following constrained optimization problem:

$$\text{Min } 0.75(Z_{50}) + 0.25(Z_{25})$$

$$\text{st. } 0.75(\sqrt{z_{50}}) + .25(\sqrt{z_{25}}) - 40 > 120$$

$$0.75(\sqrt{z_{50}}) + .25(\sqrt{z_{25}}) - 40 > .50(\sqrt{z_{50}}) + .50(\sqrt{z_{25}}) - 20$$

$$0.75(\sqrt{z_{50}}) + .25(\sqrt{z_{25}}) - 40 > .25(\sqrt{z_{50}}) + .75(\sqrt{z_{25}}) - 5$$

The solution to this optimization problem is $Z_{50} = \$32,400$ and $Z_{25} = \$10,000$.

However, the owner's calculations are not over. This economic incentive scheme is the solution to the economic problem of how an owner can elicit an effort level of high with least cost, but high effort may not be the optimal choice from the owner's perspective. While it is true that high effort generates the highest expected revenue, the cost of eliciting this effort level may be such that the additional revenue is not worth the additional cost. In fact, high effort is not optimal in this mathematical principal-agent model example.

The economic incentive scheme that at minimum expected cost motivates the agent to choose medium effort is to set $Z_{50} = \$28,900$ and $Z_{25} = \$12,100$. This economic incentive scheme yields the owner expected net earnings of $0.5(\$50,000 - \$28,900) + 0.5(\$25,000 - \$12,100) = \$17,000$.

In contrast, the expected net earnings from the cost minimizing scheme that elicits an effort of high is $0.75(\$50,000 - 32,400) + 0.25(\$25,000 - 10,000) = \$16,950$. The increase in the expected compensation required to elicit this higher effort level dominates the increase in revenue.

Finally, if the principal wishes to elicit a low effort the optimal payment scheme is $Z_{50} = \$15,625$ and $Z_{25} = \$15,625$. Because the agent's utility is decreasing in the agent's effort level, paying the agent a constant wage elicits the minimum effort level. In addition, the constant wage is the most efficient risk-sharing arrangement because the manager is risk averse and the owner is risk neutral.

To elicit a higher effort level, the manager must bear some of the uncertainty regarding the state of nature. This trade-off between providing motivation and risk sharing is a general feature of agency problems. To

Table 4.2 Calculations of Net Earnings

Effort	Expected Revenue	Required Wage	Net Earnings
High	\$43,750	\$25,600	\$18,150
Medium	\$37,500	\$19,600	\$17,900
Low	\$31,250	\$15,625	\$15,625

provide an economic incentive for the manager to choose anything but the minimum effort level, the compensation schedule must deviate from the efficient risk-sharing contract (Holmstrom, 1979).

Levinthal (1988) notes that it is instructive to contrast the optimal contract under incomplete information with that under complete information. In this example, that the owner is risk neutral and the manager risk averse suggests that in the first-best contract all the risk regarding the state of nature should be borne by the owner. An alternative arrangement is inefficient because the manager is willing to take a salary reduction to ensure against uncertainty in the manager's income, whereas the risk-neutral owner is indifferent to any uncertainty in income.

As a result, the optimal compensation scheme is calculated under the assumption of complete information. For instance, if the manager is to exert a high effort level, the manager must receive sufficient compensation to yield a net utility level of 120 units. Therefore, if z is the wage, then it must satisfy the following relationship: $\sqrt{z} - V(\text{high}) = 120$. Since $V(\text{high})$ is equal to 40 utility units, then $Z = \$25,600$. Table 4.2 indicates the required payment level for all three of the effort levels and the resulting net earnings of the owner.

Examining the owner's net earning reveals that high effort is the preferred effort level, yielding the owner an expected net income of \$18,150 (with complete information and where the agent's effort is directly observable). Contrast this first-best contract under complete information with the optimal contract under incomplete information. There are two losses compared to the first-best solution: First, *the manager's effort level is reduced from the first-best level; second, there is inefficient risk sharing because the manager bears some of the risk associated with the exogenous state of nature*. The second-best compensation scheme provides an economic incentive for the agent to work, but this compensation scheme comes at the expense of the benefits of risk

sharing. Agents are, to some extent, held accountable for events over which they have no control.

Levinthal (1988) notes that the fundamental insight of the characterization of the second-best contract relates to the imposition of risk on the agents. The first-best contract is analogous to a wage contract. In jobs where supervisors can directly observe labor input, a first-best wage contract is possible, whereas in jobs in which the worker's input is not readily observable the optimal contract imposes risk on the worker. The latter case is clearly applicable to managerial and sales personnel. For these workers, compensation is frequently based on uncertain outcomes. In the case of managers, compensation is often based on divisional profits; similarly, for sales personnel, compensation may be based on the number of products sold. These outcomes are determined in part by the worker's actions but are also significantly influenced by factors beyond the worker's control. However, given the nonobservability of the worker's actions, compensation must be based on these partially random outcome measures if the worker is to have an economic incentive to choose any but the lowest level of effort.

Levinthal (1988) points out that the first-best contract can be achieved in a trivial manner if the agent is risk neutral. Risk neutrality of the agent implies that there is no economic welfare loss in having the agent absorb all the risk associated with the uncertain outcome. In this case, the optimal contract takes the form of the principal receiving a fixed payment and the agent receiving the residual outcome. Essentially, this amounts to the agent purchasing the firm from the principal.

Monitoring. Levinthal (1988) notes that there is a trade-off between imposing risk on the agent and carrying out costly monitoring activities. Thus, one can view inefficient risk sharing as a substitute for monitoring. For example, a sales commission is a substitute for measuring how hard and cleverly a salesperson works. Methodologically, this substitution effect has the strategic implication of joint consideration in choice of monitoring mechanism and compensation scheme.

Consider an information system that provides imperfect information as to the state of nature. The reduction in risk imposed on the agent allows the principal to reduce the expected compensation to the agent and still satisfy the requirement that the agent's expected wage is at least as large as the agent's opportunity cost. Monitoring helps the principal

distinguish whether the outcome results from the agent's action or the state of nature.

Levinthal (1988) points out that an agency theory setting is a noncooperative game in which the actions are chosen by players in the game, rather than by nature, as in a statistical decision problem. That is, agency theory is game theoretic rather than decision theoretic.

Role of Time. Levinthal (1988) notes that the repetition of an agency relationship over time tends to improve its efficiency. Holmstrom (1979) suggests that "when the [agency relationship] repeats itself over time, the effects of uncertainty tend to be reduced and dysfunctional behavior is more accurately revealed, thus alleviating the problem of moral hazard" (p. 90).

Multi-Agent Models and Tournament Contracts. Levinthal (1988) maintains that the risk imposed on an agent can be reduced by basing individual compensation on individual performance relative to that of other agents, who face similar states of nature. For example, in tournaments, the reward is a function of the rank order of performance relative to other agents.

Conclusions. Levinthal (1988) concludes that the economic incentive conflict between stockholders and management does not appear to stem primarily from effort aversion on the part of management. *It is not the lack of industriousness of top management that is typically at issue but the qualitative nature of the decisions that managers make.* For instance, it is frequently claimed that management's risk aversion leads to more conservative behavior than the maximization of shareholder wealth would imply. Models based on effort aversion may be appropriate in some contexts, but the emphasis placed on them in the principal-agent research literature is excessive. The extent of the relevance of this research literature to the economics of organization depends on the connections that the agency theory literature makes to empirical work (see, e.g., Eisenhardt, 1988).

"Theory of the Firm: Managerial Behavior, Agency Costs, and Capital Structure" (Jensen & Meckling, 1976)

I conclude this section with a discussion of Jensen and Meckling (1976). This article provides a positive (i.e., descriptive) theory of agency issues. Jensen (1983) refers to the positive theory of agency, arguing the following:

Capital intensity, degree of specialization of assets, information costs, capital markets, and internal and external labor markets are examples of factors in the contracting environment that interact with the costs of various monitoring and bonding practices to determine contractual forms. (pp. 334–335)

This positive agency branch repeatedly asserts that natural selection processes are reliably efficacious (Fama, 1980; Fama & Jensen, 1983a, 1983b; Jensen, 1983).

My own reading is that the positive agency theory (e.g., Fama, 1980; Jensen, 1983; Jensen & Meckling, 1976) provides a far too optimistic view concerning the attenuation of conflicts of interests between agents and principals. Be that as it may, the Chicago School view is given its voice in the final section of this agency theory chapter.

Jensen and Meckling (1976) integrate elements from agency, the theory of property rights, and the theory of finance to develop a theory of the ownership structure of the firm. Jensen and Meckling define an agency relationship as a contract under which one or more persons (the principals) engage another person (the agent) to perform some service on their behalf, which involves delegating some decision-making authority to the agent. If the contractual parties to the relationship are utility maximizers, there is good reason to believe that the agent will not always act in the best interests of the principals. The principals can limit divergences from their interests by establishing appropriate incentives for the agent and by incurring monitoring costs designed to limit the aberrant activities of the agent. In addition, in some situations, it will pay the agent to expend resources (e.g., economic bonding costs) to guarantee that this agent will not take certain actions that would harm the principals or to ensure that the principals will be compensated if the agent does take such actions.

In most agency relationships the principals and agent incur positive monitoring and economic bonding costs (nonpecuniary as well as pecuniary), and in addition there is some divergence between the agent's decision and those decisions that maximize the economic welfare of the principal. The dollar equivalent of the reduction in economic welfare experienced by the principals due to this divergence is also a cost of the agency relationship and is referred to as the residual loss. Jensen and Meckling (1976) define agency costs as the monitoring costs by the principal, the economic bonding expenditures by the agent, and the residual loss.

Jensen and Meckling (1976) note that because the relationship between stockholders and the manager of a corporation fits the

definition of a pure agency relationship, the issues associated with the separation of ownership and control in the modern diffuse ownership corporation are intimately associated with the general problem of agency. Jensen and Meckling also point out that the problem of inducing an agent to behave as if this agent were maximizing the principals' welfare is quite general. Agency exists in all organizations and in all cooperative efforts—at every level of management in firms, universities, mutual companies, cooperatives, governmental authorities and bureaus, unions, and so forth. Jensen and Meckling focus on the analysis of agency costs generated by the contractual arrangement between the owners and top management of the corporation.

Jensen and Meckling (1976) argue that contractual relations are the essence of the firm, not only with employees (Alchian & Demsetz, 1972) but also with suppliers, customers, creditors, and so on. The problem of agency costs and monitoring costs exists for all these contracts. Jensen and Meckling submit that most organizations serve as a nexus for a set of contracting relationships among individuals.

Jensen and Meckling (1976) argue that agency costs (i.e., monitoring costs, economic bonding costs, residual loss) are an unavoidable result of the agency relationship. *Although Jensen and Meckling argue that agency costs are nonzero, these costs are not regarded as nonoptimal in their framework.* In fact, Jensen and Meckling posit that because agency costs are borne entirely by the decision maker, the decision maker has the incentive to see that agency costs are minimized (because the decision maker captures the benefits from the reduction in agency costs).

Jensen and Meckling (1976) conclude that agency costs are as real as any other costs. The level of agency cost depends among other things on statutory and common law and human creativity in devising better contracts. Both the law and the sophistication of contracts relevant to the modern corporation are the products of an historical process in which there were strong economic incentives for individuals to minimize agency costs.

Having completed our survey of the agency literature, I now turn to the fifth chapter on resource-based theory, dynamic capabilities, and real options. Resource-based theory is the dominant economic perspective in the research area of strategic management. I submit that students studying the economics of organization who want to be first-rate scholars in this research area benefit greatly by appreciating the contributions of organization theory (e.g., the behavioral theory of the firm) and economics (e.g., transaction costs, property rights, and

agency theory). During the next decade, professionals should see the development of resource-based theory by a new generation of enthusiastic and gifted scholars.

I begin Chapter 5 with the seminal book by Edith Penrose (1959), *The Theory of the Growth of the Firm*. This book has had moderate success in the industrial organization research literature over the years (Scherer & Ross, 1990) and has greatly affected the strategic management field from the mid-1980s to the present.





5

Resource-Based Theory, Dynamic Capabilities, and Real Options

Although early contributions to resource-based theory and dynamic capabilities came from the discipline of economics (e.g., Demsetz, 1973; Gort, 1962; Marris, 1964; Penrose, 1959; Richardson, 1960, 1972; Rubin, 1973; Slater, 1980), during the last 20 years the business field of strategic management has made significant contributions to resource-based theory and dynamic capabilities (e.g., Foss, 1997; Heene & Sanchez, 1997; Volberda & Elfring, 2001). Logic dictates that (organizational) economic theory will continue to play an important role in the study of economic value creation and sustainable competitive advantage. After all, sustainable competitive advantage requires an understanding of market frictions, and there is a large and well-developed economics research literature on market failures that students studying the economics of organization can draw on.

Although the market-failures literature is well developed, the organizational-failures literature is comparatively less developed, thereby providing research opportunities for students studying the economics of organization. Furthermore, resource-based theory and dynamic capabilities and real options research may develop into a paradigmatic approach to strategic management, an important contribution to the evolving science of organization. Clearly, there is a need for rigorous empirical research to establish both the nature and the impact of dynamic capabilities on sustainable competitive advantage. Capabilities that can prove especially useful in dynamic business environments are operational and strategic flexibility.

I begin this chapter with the seminal work of Penrose (1959), who provides (1) a general theory of the growth of the firm, (2) a theory of entrepreneurship based on the subjective opportunity set of the firm, (3) expansion based on indivisibility and the balance of processes, (4) a

resource-based theory of diversification, and (5) a theory of expansion through acquisition and merger. In addition, Penrose provides a theory of the limits to the rate of the growth of the firm, in particular, arguing that the binding constraint on the firm's rate of the growth is provided by the capacities of its existing management—the so-called Penrose effect.

Chandler's (1990) *Scale and Scope* represents a culmination of Chandler's long quest to chart the evolution of modern industrial enterprise. The book provides the reader with an extraordinary breadth and depth of knowledge concerning the development of managerial capitalism. The essence of successful firm strategy, Chandler argues, is the making of three interrelated investments: (1) investment in production to achieve the cost advantages of scale and scope; (2) investment in product-specific marketing, distribution, and purchasing networks; and (3) investment in managerial talent and management structure to plan, coordinate, and monitor the firm's often dispersed operations. Chandler argues that such three-pronged investment enables firms to develop organizational capabilities, which, in turn, provides the dynamic for the continuing growth of the enterprise.

Itami and Roehl (1987) emphasize the dynamic fit between resources and the environment. Itami and Roehl build on the work of Penrose (1959) concerning corporate growth and move the arguments forward by emphasizing the role of invisible assets of a firm, which are based on information. Invisible assets include intellectual property rights of patents and trademarks, trade secrets, proprietary data files, personal and organizational networks, reputation, and culture. Itami and Roehl argue that these invisible assets are often a firm's only real source of sustainable competitive advantage.

Nelson and Winter (1982) consider the promise and the problems of evolutionary modeling of economic change. Among the many benefits that may be derived from Nelson and Winter's theoretical approach that reconciles economic analysis with real-world business firm decision making, the most important relate to improved understanding of technological change and the dynamics of the competitive process. Nelson and Winter's evolutionary theory is intrinsically dynamic, in which the heterogeneity of firms is a key feature.

This chapter on dynamic resource-based theory concludes with a research book that some in the strategic management field may find to be a curious choice. Over time, however, I anticipate that it will become abundantly clear that a key category in developing dynamic capabilities involves strategies that enhance adaptability and strategic flexibility.

Viewed in this light, Trigeorgis's (1996) research book is an important contribution to the dynamic capabilities research literature. Real options research has the potential to make a significant difference to our understanding of resource accumulation and capability-building processes and investment decision making under uncertainty. Finally, supplementing the real options analysis with game-theoretic tools that capture competitive dynamics is promising for future research by students pursuing the evolving science of organization.

The Theory of the Growth of the Firm (*Penrose, 1959*)

Penrose (1959) is concerned with the growth of firms and only incidentally with the size of the firm. Penrose argues that firm size is only a by-product of the process of growth and that there is no optimum, or even most profitable, size of the firm. Penrose is primarily concerned with a theoretical analysis of the growth process of the firm.

Penrose (1959) emphasizes the internal resources of a firm—on the productive services available to a firm from its own resources, particularly the productive services available from management with experience within the firm. The (firm-specific) experience of management affects the productive services that all its other resources are capable of rendering. As management tries to make the best use of the available resources, a dynamic interacting process occurs that encourages a continuous, but limited, rate of growth of the firm. To focus attention on the crucial role of the firm's inherited resources, the environment is treated, in the first instance, as an image in the entrepreneur's mind of the possibilities and restrictions with which it is confronted. For it is, after all, such an image that in fact determines a person's behavior. Whether experience confirms expectations is another story.

The Firm in Theory. Penrose (1959) notes that in a private enterprise industrial economy the business firm is the basic unit for the organization of production. Because of its complexity and diversity, a firm can be approached with many different types of analysis—sociological, organizational, engineering, or economic—and from whatever point of view that seems appropriate to the business problem at hand.

The theory of the firm—as it is called in the neoclassical economics literature—was constructed for the purpose of assisting in the theoretical investigation of one of the central problems of economic analysis—the way in which prices and the allocation of resources among different

uses are determined. It is but part of the wider theory of economic value. The equilibrium of the firm is, in essence, the equilibrium output. As Boulding (1950) notes, "The firm is a strange bloodless creature without a balance sheet, without any visible capital structure, without debts, and engaged apparently in the simultaneous purchase of inputs and the sale of outputs at constant rates" (p. 24).

Penrose (1959) points out that if we become interested in other aspects of the firm, then we ask questions that the theory of the firm is not designed to answer. Penrose wants to deal with the firm as a growing organization, not as a price-and-output decision maker for given products. Penrose argues that the essential difference between economic activity inside the firm and economic activity in the market is that economic activity in the firm is carried out within an administrative organization (see Simon, 1947), whereas economic activity in the market is not. Penrose refers to this court of last resort in the firm as central management. It is the area of coordination—the area of authoritative communication (Barnard, 1938; Simon 1947)—which defines the boundaries of the firm, and, consequently, it is a firm's ability to maintain sufficient administrative coordination to satisfy the definition of an industrial firm that sets the limit to its size as an industrial enterprise.

The Firm as a Collection of Productive Resources. According to Penrose (1959), a firm is more than an administrative unit; a firm is also a collection of productive resources, where the choice of different uses of these resources over time is determined by administrative decision. The physical resources of a firm consist of tangible things—plant, equipment, land, and natural resources; raw materials; semifinished goods; waste products and by-products; and even unsold stocks of finished goods. There are also human resources available in a firm—unskilled and skilled labor and clerical, administrative, financial, legal, technical, and managerial staff.

Penrose (1959) argues that, strictly speaking, it is never resources themselves, but only the services that the resource can render, that are the inputs in the production process. Resources consist of a bundle of potential services and can, for the most part, be defined independently of their use, whereas services cannot be so defined, the very word *service* implying a function, an activity. *It is largely in this distinction that we find the source of the uniqueness of each individual firm.*

The business firm, as Penrose (1959) defines it, is both an administrative organization and a collection of productive resources. The general

purpose of the business firm is to organize the use of its own resources together with other resources acquired from outside the firm for the production and sale of goods and services at a profit.

The term *entrepreneur* refers to individuals or groups within the firm providing the entrepreneurial services, whatever their position or occupational classification may be. Entrepreneurial services are those contributions to the operations of a firm that relate to the introduction and acceptance on behalf of the firm of new ideas, particularly with respect to products, location, and significant changes in technology; to the acquisition of new managerial personnel; to fundamental changes in the organization of the firm; to the raising of capital; and to the making of plans for expansion, including the strategic choice of expansion method.

Penrose (1959) submits that a versatile type of executive service is needed if expansion requires major efforts on the part of the firm to develop new markets or entails branching out into new lines of production. Here, the imaginative effort, the sense of timing, the instinctive recognition of what will catch on or how to make it catch on, become of overwhelming importance. These services are not likely to be equally available to all firms. Firms not only alter the environmental conditions necessary for the success of their actions, but, even more important, they also know that they can alter these conditions and that the environment is not independent of their own activities.

Expansion Without Merger: The Receding Managerial Limit. Penrose (1959) notes three classes of explanation for why there may be a limit to the growth of firms—managerial ability, product or factor markets, and uncertainty and risk. The first explanation refers to conditions within the firm, the second explanation refers to conditions outside the firm, and the third explanation is a combination of internal attitudes and external conditions. The capacities of the existing managerial personnel of the firm necessarily set a limit to the expansion of that firm in any given period of time, and such management possessing firm-specific abilities cannot be hired in the marketplace.

Penrose (1959) argues that an administrative group is something more than a collection of individuals; an administrative group is a collection of individuals who have had experience in working together—only in working together can teamwork be developed. Experiences these individuals gain from working within the firm, and with each other, enable them to provide services that are uniquely valuable for the operations of the particular group with which they are associated. Existing management limits the amount of new management that can be hired

at any point in time (after all, the services of existing management are required to instruct the new personnel).

Penrose (1959) submits that if a firm expands its organization more rapidly than the individuals in the expanding organization can obtain the experience with each other they need for effective operation of the group, the efficiency of the firm will suffer. Because the services from current managerial resources control the amount of new managerial resources that can be absorbed, they create a fundamental limit to the amount of expansion a firm can undertake at any point in time. The amount of activity that can be planned at a given time period limits the amount of new personnel that can be profitably absorbed in the next period. This idea over subsequent years is known as the *Penrose effect*. Moreover, as plans are completed and put into operation, managerial services absorbed in the planning processes will be gradually released and become available for further planning.

Penrose (1959) argues that knowledge comes to people in two different ways: It can be formally taught, and it can be achieved via learning-by-doing in the form of personal experience. Experience produces increased knowledge and contributes to objective knowledge insofar as its results can be transmitted to others. But experience itself can never be transmitted; experience produces a change—frequently a subtle change—in individuals and cannot be separated from them.

Increasing experience shows itself in two ways—changes in knowledge acquired and changes in the ability to use knowledge. There is no sharp distinction between these two forms because to a considerable extent the ability to use old knowledge is dependent on the acquisition of new knowledge. But it is not exclusively so dependent; with experience a person may gain wisdom, sureness of movement, and confidence—all of these become part of his or her very nature, and they are all qualities that are relevant to the kind and amount of services a person can give to the firm. Much of the experiences of business personnel are frequently so closely associated with a particular set of external circumstances that a large part of a personnel's most valuable services may be available only under these circumstances.

A person whose past productive activity has been spent within a particular firm, for example, can, because of his or her intimate knowledge of the resources, structure, history, operations, and personnel of the firm, render services to that firm that person could give to no other firm without acquiring additional experience. Penrose (1959) submits that, after it is recognized that the very processes of operating and of

expanding are intimately associated with a process by which knowledge is increased, it becomes clear that the productive opportunity of a firm will change, even in the absence of any change in external circumstances or in fundamental technological knowledge. New opportunities open up that did not exist at the time expansion plans were made. That is, the subjective opportunity set of the firm will change.

Penrose (1959) hastens to add that management not only is the source for expansion but also is a brake on expansion. A firm has a given amount of experienced managerial services available at any one time. Parts of these managerial services are needed for ordinary operation, and the rest of these managerial services are available for planning and executing expansion programs. The effect of uncertainty is to require that some of these available services be used to gather information, process the information, and reach conclusions about the possibilities of action in which the management team has confidence.

Inherited Resources and the Direction of Expansion. Penrose (1959) maintains that the external inducements to expansion include growing demand for particular products, changes in technology that call for production on a larger scale than before, discoveries and inventions with particularly promising uses, and opportunities to obtain a better market position. Inducements to expansion also include backward integration to control sources of supply, diversification of final products to spread risk, or expansion of existing or allied products to preclude the entry of new competitors. External obstacles to expansion include keen competition in markets for particular products that makes profitable entry or expansion in those markets difficult.

Penrose (1959) argues that whereas external inducements and obstacles have been widely discussed, little attention has been paid, in a systematic way at least, to the equally important internal influences on the direction of expansion. Internal obstacles arise when some of the important types of specialized services required for expansion in particular directions are not available in sufficient amounts within the firm. In particular, internal obstacles arise when not enough of the managerial capacity and the technical skills required for the planning, execution, and efficient operation of a new program can be obtained from among existing experienced personnel. Internal inducements to expansion arise largely from the existence of a pool of unused productive services, resources, and specialized knowledge, all of which will always be found within any firm. To Penrose, a resource can be viewed as a bundle of possible services. As long as resources are not used fully

in current operations, there is an economic incentive for a firm to find a way of using them more fully.

Penrose (1959) maintains that three significant obstacles preclude the attainment of a state of rest:

1. Those arising from the familiar difficulties posed by the indivisibility of resources, which Penrose (1959) calls "the jig-saw puzzle" (p. 69)
2. Those arising from the fact that the same resources can be used differently under different circumstances
3. Those arising because, in the ordinary processes of operation and expansion, new productive services are continually being created

Penrose (1959) then discusses how the division of labor (specialization) can lead to the growth of the firm and diversification. This process has been called the *virtuous circle*, in which specialization leads to higher common multiples and higher common multiples lead to greater specialization. Penrose also argues that diversification strategy can be driven by the desire to achieve multiproduct economies of scale (which in modern strategic management language is called *economies of scope*; Teece, 1980).

Penrose (1959) observes that for many purposes it is possible to deal with rather broad categories of resources, overlooking the lack of homogeneity in the members of the category. Economists usually recognize this simplification, stating that for convenience alone resources are grouped under a few headings—for example, land, labor, and capital—but Penrose (1959) points out that the subdivision of resources may proceed as far as is useful, according to whatever principles are most applicable for the business problem at hand. The heterogeneity of the productive services available or potentially available from its resources gives each firm its unique character.

Furthermore, the possibilities of using services change with changes in knowledge. Consequently, there is a close connection between the type of knowledge possessed by the personnel of the firm and the services obtainable from its material resources. The firm, then, is viewed as a collection of resources. Unused productive services shape the scope and direction of the search for knowledge. Knowledge and an economic incentive to search for new knowledge are built into the very nature of firms possessing entrepreneurial resources of even average initiative. Physically describable resources are purchased in the strategic factor markets for their known services, but, as soon as these resources

become part of a firm, the range of services they are capable of yielding starts to change. The services that resources yield depends on the capacities of the people using them, but the development of the capacities of people is partly shaped by the resources they deal with. The two together create the distinctive, subjective, productive opportunity set of a particular firm.

If resources were completely nonspecific, a firm could in principle produce anything. The selection of the relevant product markets is necessarily determined by the inherited resources of the firm—the productive services it already has. To be sure, the anticipation of consumer acceptance is a necessary condition of entrepreneurial interest in any product, but the original economic incentive to a great deal of innovation can be found in a firm's desire to use its existing resources more efficiently. There is a close relation between the various kinds of resources with which a firm works and the development of ideas, experience, and knowledge of its managers and entrepreneurs. Changing experience and knowledge affect not only the productive services available from resources but also the demand as seen by the firm.

Penrose (1959) further elaborates, noting that unused productive services are, for the enterprising firm, at the same time a challenge to innovate, an economic incentive to expand, and a source of sustainable competitive advantage. Unused productive services facilitate the introduction of new combinations of resources—innovations—within the firm. Unused productive services are a selective force in determining the direction of expansion. Therefore, analysis is required of internal and external inducements and internal and external obstacles for expansion.

The Economies of Diversification. Penrose (1959) argues that of all the outstanding characteristics of business firms, perhaps the most inadequately treated in economic analysis is the diversification of their activities. *Anticipating Teece (1982), Penrose argues that market imperfections are an important explanation of diversification strategy.* Diversification that involves a departure from the firm's existing areas may be one of three kinds:

- The entry into new markets with new products using the same production base
- Expansion in the same market with new products based on a different area of technology
- Entry into new markets based on a different area of technology

Penrose (1959) observes that a firm's opportunities are necessarily widened when the firm develops a specialized knowledge of a technology that is not very specific to any particular kind of product, for example, knowledge of different types of engineering or industrial chemistry. Diversification and expansion based primarily on a high degree of competence and technical knowledge in specialized areas of manufacture are characteristic of many of the largest firms in the economy. Penrose submits that this type of competence, together with the market position such competence and technical knowledge ensures, is the strongest and most enduring position a firm can develop.

Diversification through both internal and external expansion is likely to be extensive because of the wide variety of productive services generated within such firms and because the competitive advantages these firms possess are conducive to expansion. Opportunities for expansion both within existing resource bases and through the establishment of new resource bases are likely to be so prevalent that the firm has to choose carefully among many different courses of action.

The Firm as a Pool of Resources. Penrose's (1959) thesis is that a firm is essentially a pool of resources, the use of which is organized in an administrative framework. In a sense, the final products being produced by a firm at any given time represent one of several ways in which the firm could be using its resources, an incident in the development of its basic potentialities. The continual change in the productive services and knowledge within a firm and the continual change in external circumstances present the firm with a continually changing productive opportunity set.

The Rate of Growth of a Firm Through Time. Penrose (1959) notes that markets and firms are interacting institutions, each being necessary to the existence of the other. Penrose emphasizes that one of the more significant characteristics of entrepreneurial and managerial services is their heterogeneity, their uniqueness for every individual firm. The factors determining the availability of managerial services and the need for these services in expansion determine the maximum rate of growth of the firm, where rate of growth is defined as the percentage rate at which the size of the firm increases per unit of time. The possibility of acquiring other firms raises enormously the maximum rate of expansion, primarily because acquisition substantially reduces the managerial services required per unit of expansion.

Concluding Comments. Kor and Mahoney (2000) suggest the following list of key ideas derived from Penrose (1959):

- Firm growth can be usefully studied as a dynamic process of management interacting with resources.
- Firms are institutions created by people to serve the purposes of people.
- Services of resources are drivers of firm heterogeneity.
- Services that material resources will yield depend on the knowledge possessed by human resources. The two together create a subjective opportunity set that is unique for each firm.
- Firm growth is a function of firm-specific experiences in teams.
- Managerial capability is the binding constraint that limits the growth rate of the firm—the so-called Penrose effect.
- Excess capacity of productive services of resources is a driver of firm growth.
- Unused productive services of resources can be a source of innovation.
- Firm diversification is often based on a firm's competencies that can lead to a sustainable competitive advantage.
- An important component of the competitive process is experimentation.

Finally, it is noted that some criticize Penrose's resources approach for ignoring the business environment. Penrose (1959), in fact, addresses this issue, arguing that whether or not we treat the resources of the firm or its environment as the more important factor explaining growth depends on the question we ask: If we want to explain why different firms see the same environment differently, why some firms grow and some do not, or, to put it differently, why the environment is different for every firm, we must take the resources approach; if we want to explain why a particular firm or group of firms with specified resources grows the way it does, we must examine the opportunities for the use of those resources. Penrose calls these opportunities for small firms the *interstices* in the economy. The productive opportunities of small firms are thus composed of those interstices left open by the large firms that the small firms see and believe they can take advantage of. Penrose concludes that management's experiments with different types of corporate structures are in themselves an important aspect of competition.

In my judgment, Penrose (1959) is the seminal work in resource-based theory that anticipates the works of Chandler (1962, 1990), which document organizational innovations and organizational capabilities

that, in turn, provide an internal dynamic for the continuing growth of the modern industrial enterprise.¹

Scale and Scope: The Dynamics of Capitalism (*Chandler, 1990*)

I turn now to Chandler (1990), who provides a detailed but highly generalized description and analysis of the creation and dynamic evolution of the central institution of managerial capitalism—the modern industrial enterprise. These concepts and generalizations are then used to develop an explanatory theory of the evolution of the modern industrial enterprise. The richness of information provided in this research book can be helpful for students in the evolving science of organization in answering questions that have long concerned economists and business historians—questions about changes in internal organization and management, competition and cooperation among firms, growth through horizontal acquisitions and mergers, vertical integration, expansion into international markets, diversification into new product lines, and the effect of legal requirements, government rulings, and cultural values on firm growth and economic performance.

Chandler (1990) observes that in the last half of the 19th century a new form of capitalism appeared in the United States and Europe. The building and operating of rail and telegraph systems called for the creation of a new type of business enterprise. The massive investment required in constructing those systems, and the complexities of their operations, brought the separation of ownership and management. The

¹For further readings see Penrose (1955, 1960). Penrose (1960) provides a case study of the Hercules Powder Company to illustrate that growth is governed by a creative and dynamic interaction between a firm's productive resources and its market opportunities. Richardson (1972), Rubin (1973), Slater (1980), and Teece (1982) are influential journal articles in the economics research literature that build on Penrose (1959). For a recent assessment of Penrose (1959), see Kor and Mahoney (2000), which focuses on (1) the research process that led to Penrose's (1959) classic, (2) Penrose's (1959) contributions to the discipline of strategic management, (3) the generative nature of Penrose's (1959) research for current resource-based theory, and (4) suggested future research building on Penrose's (1959) resource approach. In addition, Kor and Mahoney (2004) and Rugman and Verbeke (2002) consider Penrose's (1959) direct and indirect contributions to the modern resource-based view within strategic management. Pettus (2001) studies the Penrose effect in the deregulated trucking industry, whereas Tan (2003) and Tan and Mahoney (2003) provide empirical tests of Penrose (1959) in the context of multinational firms. Finally, Pitelis (2002) provides an excellent edited collection of recent writings that document the legacy of Penrose (1959) on contemporary research on the growth of the firm.

enlarged enterprises came to be operated by teams of salaried managers who had little or no equity in the firm.

Chandler (1990) notes that the new forms of transportation and communication, in turn, permitted the rise of modern mass marketing and modern mass production. The unprecedented increase in the volume of production, and in the number of transactions, led the entrepreneurs who established the new mass-producing and mass-distributing enterprises—like the railroad personnel before them—to recruit teams of salaried managers.

Chandler (1990) examines the beginning and growth of global managerial capitalism, focusing on the history of its basic institution—the modern industrialized enterprise—in the world's three leading industrial nations. They each had been rural, agrarian, and commercial and each became industrial and urban, a transformation that, in turn, brought the most rapid economic growth in business history. At the center of the transformation were the United States, Great Britain, and Germany, which together accounted for just over two thirds of the world's industrial output in 1870. The industrial sector grew significantly in the United States and Germany; in Great Britain the development was slower but sustained. Further, whereas Great Britain experienced only a moderate change of employment structure after the 1880s, the United States—and Germany to a lesser degree—showed a dramatic transformation from an agrarian to a modern economy, in which almost half of the country's employment centered in industry.

Chandler (1990) maintains that as a result of the regularity, increased volume, and greater speed of the flows of goods and materials made possible by the new transportation and communication systems, new and improved processes of production developed that for the first time in business history achieved substantial economies of scale and scope. Large manufacturing works applying the new technologies could produce at lower costs than could the smaller manufacturing works.

Chandler (1990) observes that for entrepreneurs to benefit from the cost advantages of these new, high-volume technologies of production, the entrepreneurs had to make three sets of interrelated investments:

1. An investment in production facilities large enough to use a technology's potential economies of scale and scope
2. An investment in a national and international marketing and distribution network so that the volume of sales might keep pace with the new volume of production

3. Investment in management, which required entrepreneurs to recruit and train managers not only to administer the enlarged facilities and increased personnel in both production and distribution but also to monitor and coordinate those two basic functional activities and to plan and allocate resources for future production and distribution

Chandler (1990) submits that *this three-pronged investment in production, distribution, and management brought the modern industrial enterprise into being*. The first entrepreneurs to create such enterprises acquired substantive competitive advantages. Their industries quickly became oligopolistic—that is, dominated by a small number of first movers. These first-mover firms, along with a few challengers that subsequently entered the industry, no longer competed primarily on the basis of price. Instead, these firms competed through functional and strategic effectiveness. These firms did so functionally by improving their product, marketing, purchasing, and labor relations, and these firms did so strategically by moving into growing markets more rapidly and by divesting out of declining markets more quickly and effectively than did their competitors.

Such rivalry for market share and profitability honed the enterprise's functional and strategic capabilities. These organizational capabilities, in turn, provided an internal dynamic for the continuing growth of the enterprise. In particular, these organizational capabilities stimulated its owners and managers to expand into more distant markets in their own country and then to become multinational by moving abroad. These organizational capabilities also encouraged the firm to diversify by developing products competitive in markets other than the firm's original market, becoming a multiproduct enterprise.

Scale, Scope, and Organizational Capabilities. Chandler (1990) argues that the modern industrial enterprise can be defined as a collection of operating units, each with its own specific facilities and personnel, whose combined resources and activities are coordinated, monitored, and allocated by a hierarchy of middle and top managers. This hierarchy makes the activities and operations of the whole enterprise more than the sum of its operating units. The manufacturing enterprises became multifunctional, multiregional, and multiproduct because the addition of new units permitted these enterprises to maintain a long-term rate of return on investment by reducing overall costs of production and distribution, by providing products that satisfied existing demands, and by transferring facilities and capabilities to more profitable markets when economic returns were reduced by competition, changing technology, or altered market demand.

Chandler (1990) submits that whatever the initial motivation for its investment in new operating units, the modern industrial enterprise has rarely continued to grow or maintain its competitive position over an extended period of time unless the addition of new units (and to a lesser extent the elimination of old units) has actually permitted the visible hand of its managerial hierarchy to reduce costs, to improve functional efficiency in marketing and purchasing as well as production, to improve existing products and processes and to develop new ones, and to allocate resources to meet the challenges and opportunities of ever-changing technologies and markets. Such a process of growth has provided organizations with the internal dynamic that has enabled them to maintain their position of dominance as markets and technologies have changed. Chandler further argues that it was the development of new technologies and the opening of new markets that resulted in economies of scale and scope and reduced transaction costs, that made the large, multiunit enterprise come when it did, where it did, and in the way it did.

Chandler (1990) maintains that coordination demanded the constant attention of a managerial team or hierarchy. The potential economies of scale and scope are a function of the physical characteristics of the production facilities. However, the actual economies of scale and scope, as measured by throughput, are a function of organizational capabilities. The full fruition of economies of scale and scope depend on knowledge, skill, experience, and teamwork—on the organizational capabilities essential to use the full potential of technological processes. Further, in many instances, Chandler finds that the first company to build a plant of minimum efficient scale, and to recruit the essential management team to enable the enterprise to reach its full potential, often remained the leader in its industry for decades.

Chandler (1990) indicates that organizational capabilities included those of lower management and the workforce, in addition to the skills of middle and top management. Organizational capabilities also included the facilities for production and distribution acquired to use fully the economies of scale and scope. Such organization capabilities provided the economic profits that, in large part, financed the continuing growth of the enterprise. Highly product specific and process specific, these organizational capabilities affected—indeed, often determined—the direction and pace of first movers and challengers and of the industries and even the national economies in which they operated (Collis, 1994).

Chandler (1990) emphasizes that only if these facilities and organizational capabilities were carefully coordinated and integrated could the enterprise achieve the economies of scale and scope that were needed to

compete in national and international markets and to continue to grow. Middle managers not only had to develop and apply functional-specific and product-specific managerial skills, but they also had to train and motivate lower level managers and to coordinate, integrate, and evaluate their work. Such organizational capabilities, of course, had to be created, and, once established, these capabilities had to be maintained. Their maintenance was as great a challenge as their creation because facilities depreciate, individual skills atrophy, and organizational capabilities can diminish. Moreover, changing technologies and markets constantly make existing facilities, individual skills, and organizational capabilities obsolete. One of the more critical tasks of the top management team has always been to maintain these organizational capabilities and to integrate these facilities and skills into a coherent, unified organization—so that the whole becomes more than the sum of its parts. Such organizational capabilities, in turn, have provided the source—the dynamic—for the continuing growth of the enterprise. Organizational capabilities have made possible the earnings that supplied much of the funding for such growth.

As Chandler (1990, 1992) repeatedly emphasizes, in the collective individual industries that are so aptly documented, the first movers' initial, interrelated, three-pronged investments in manufacturing, marketing, and management created powerful barriers to entry (see also Porter, 1980). Challengers had to make comparable (sunk-cost) investments at a greater risk, precisely because the first movers had already learned the ways of the new processes of production, were already dominating the markets for the new or greatly improved products, and were already reaping substantial economic returns from their initial investments. As the first movers' functional and organizational capabilities were honed, the difficulties of entry by newcomers became even more formidable. In the sale of consumer products, particularly branded, packaged goods, barriers to entry were reinforced by advertising, vertical tying contracts, and exclusive franchising. In the more technologically advanced producer-goods industries, patents reinforced these entry barriers. In Europe, first movers strengthened their strategic positioning still further by arranging interfirm agreements as to price, output, and market territories.

A New Era of Managerial Capitalism? Chandler (1990) notes that the historian who has studied the past experience of the business enterprise is in a better position than most analysts to identify current business practices that are truly new. Chandler observes that of the many recent

changes in the growth, management, and financing of the modern industrial enterprise, the following six have no historical precedents:

- The adaptation of a new corporate strategy of growth—that of moving into new markets where the organizational capabilities of the enterprise do not provide competitive advantages
- The separation of top management in the corporate office from middle management in the operating divisions
- The extensive and continuing divestiture of operating units
- The buying and selling of corporations as a distinct business in its own right
- The role played by portfolio managers in the capital markets
- The evolution of those capital markets to facilitate the coming of what has been termed a market for corporate control

Chandler (1990) concludes that his research book has only begun to map the evolution of the industrial enterprise in the United States, Great Britain, and Germany from the 1880s to the 1940s. Valid description and analysis on which generalizations can be made must await an in-depth, industry-by-industry, country-by-country historical study.² Much more work needs to be done that certainly may modify the patterns of institutional change that Chandler has outlined. Clearly, there are research opportunities for those students studying the economics of organization who combine the craft of the business historian and the analytical skills derived from the resource-based/dynamic capabilities perspective. Indeed, Chandler provides insights that connect (company-specific) organizational capabilities and the economics of organization.

Mobilizing Invisible Assets (*Itami & Roehl, 1987*)

To develop further the dynamic capabilities perspective, I turn now to Itami and Roehl's (1987) contribution to dynamic capabilities theory. Itami and Roehl emphasize the role of environmental, corporate, and internal information flow. Environmental information flow includes discovering customer preferences and maintaining competitor intelligence.

²Resource-based/dynamic capabilities theory has recently been empirically corroborated in the context of international business studies (e.g., Anand & Delios, 2002; Anand & Singh, 1997). Peng (2001) documents the extent to which resource-based theory has diffused in international business research.

Corporate information flow includes reputation, brand image, and marketing know-how. Internal information flow includes corporate culture and managerial capabilities (e.g., routines).

Itami and Roehl (1987) provide a strategic logic that is heavily influenced by Penrose (1959) and emphasizes the vital contribution of accumulated experience and information to a corporation's strategic resources. Itami and Roehl emphasize that the intangible assets, such as a particular technology, accumulated consumer information, brand name, reputation, and corporate culture, are invaluable to the firm's competitive advantage. In fact, these invisible resources are often a firm's only real source of competitive edge that can be sustained over time.

Itami and Roehl (1987) emphasize that current strategy, because it can change the level of invisible assets, is more than the basis for short-term competitive advantage; current strategy provides the foundation for future strategy and adds to or erodes the invisible resource base. The competitive success of a strategy is dependent on the firm's invisible assets, but the dynamics of invisible assets (their accumulation and depreciation over time) is also largely determined by the content of that strategy. Itami and Roehl explore how invisible assets affect, and are affected by, the firm's strategy. Decisions made today can affect a firm's long-term capabilities and adaptability because such decisions often determine the accumulation of invisible assets.

Itami and Roehl (1987) maintain that many invisible resources are quite fixed. There is no easy way to obtain a well-known brand name or advanced technical production skills in the market, nor can money buy an instantaneous change in corporate culture and employee morale. Accumulation of these invisible resources requires ongoing, conscious, and time-consuming efforts; you cannot just go out and buy these resources off the shelf. For this reason, a firm can differentiate itself from competitors through its invisible resources. If a resource can be bought, competitors with sufficient financial resources can gain access to it. And if a resource can be created quickly, competitors will have ready access to such a resource through imitation. But competitors cannot do this easily with invisible resources.

The important features of invisible resources are as follows:

- Unattainable with money alone
- Time-consuming to develop
- Capable of multiple simultaneous uses
- Able to yield multiple, simultaneous benefits

These features of invisible resources make it crucial to consider carefully the strategies for accumulating invisible resources.

Information is at the heart of invisible resources. Information-based invisible resources include not only the stock of accumulated information in the firm but also the channels that handle the flow of information of importance to the firm. Information can be classified as environmental, corporate, or internal.

Environmental information flows from the environment to the firm, creating invisible assets related to the environment. This type of information flow includes production skills, customer information, and channels for bringing in information.

Corporate information flows from the firm to the environment, creating invisible assets stored in the environment. This category of information flow includes such invisible assets as corporate reputation, brand image, corporate image, and influence over the distribution and its parts suppliers, as well as marketing know-how.

Internal information originates and terminates within the firm, again affecting the invisible asset stock. This category of information flow includes corporate culture, morale of workers, and management capabilities, as well as the firm's ability to manage information, the employees' ability to transmit and use the information in decision making, and the employees' habits and norms of effort expended. *Successful accumulation of invisible resources comes down to control of the information flow.*

In my judgment, Itami and Roehl (1987) is a seminal contribution to resource-based theory and the dynamic capability approach. Invisible assets serve as a focal point of strategy development and growth. Students studying the economics of organization are served well in examining closely this often overlooked classic.

I turn next to a classic that almost everyone recognizes as the seminal and path-breaking book on evolutionary economics and dynamic capabilities.

An Evolutionary Theory of Economic Change (Nelson & Winter, 1982)

Nelson and Winter (1982) provide a wealth of strategic issues for consideration by current students who want to contribute to the evolving science of organization. Nelson and Winter provide the organization-theoretic foundations of economic evolutionary theory, the building blocks of which include individual skills and organizational capabilities. Nelson

and Winter develop an evolutionary model of economic growth and a perspective that emphasizes the role of Schumpeterian competition.

Nelson and Winter (1982) argue that much of firm behavior can be more readily understood as a reflection of general routines and strategic orientations coming from the firm's past than as the result of a detailed survey of the remote twigs of a decision tree extending into the future. Nelson and Winter acknowledge their intellectual debts to Joseph Schumpeter and Herbert Simon. Schumpeter (1934, 1950) points out the right problem—how to understand economic change—and many of the important elements of the answer. Simon (1982) provides a number of specific insights into human and organizational behavior that are reflected in Nelson and Winter's theoretical models. But, most important, Simon's (1947, 1982) works encourage Nelson and Winter in maintaining the view that there is much more to be said on the problem of rational behavior in the world of experience than can be adequately stated in the language of orthodox economic theory.

Nelson and Winter (1982) develop an evolutionary theory of the organizational capabilities and behaviors of business firms operating in a market environment. The firms in their evolutionary theory are treated as motivated by profitability and engaged in the search for ways to improve their profitability, but the firm's actions are not assumed to be profit maximizing over well-defined and exogenously given choice sets. Evolutionary theory emphasizes the tendency for the more profitable firms to drive the less profitable firms out of business. However, Nelson and Winter do not focus their analysis on hypothetical states of industry equilibrium, in which all unprofitable firms are no longer in the industry and profitable firms are at their desired size.

Relatedly, the modeling approach employed in Nelson and Winter (1982) does not use the familiar maximization calculus to derive equations characterizing the behavior of firms. Rather, firms are modeled as having, at any given time, certain organizational capabilities and decision rules. Over time, these organizational capabilities and decision rules are modified as a result of both deliberate problem-solving efforts and random events. And over time, the economic analogue of natural selection operates as the market determines which firms are and are not profitable and winnows out the unprofitable firms. Supporting Nelson and Winter's analytical emphasis on this sort of evolution by natural selection is a view of organizational genetics—the processes by which traits of organizations, including those traits underlying the capability to produce output and to make profits, are transmitted through time.

Nelson and Winter (1982) give attention to uncertainty, bounded rationality, the presence of large corporations, institutional complexity, and the dynamics of the actual adjustment process. Considerable attention is also given to imperfect information and imperfect competition, transaction costs, indivisibilities, increasing returns, and historical change.

Although Nelson and Winter (1982) stress the importance of certain elements of continuity in the economic process, they do not deny (nor does contemporary biology deny) that change is sometimes rapid. Also, some people who are particularly alert to teleological fallacies in the interpretation of biological evolution seem to insist on a sharp distinction between explanations that feature the processes of blind evolution and those that feature deliberate goal seeking. Whatever the merits of this distinction in the context biological evolution theory, such a distinction is unhelpful and distracting in the context of Nelson and Winter's theory of the business firm. It is neither difficult nor implausible to develop models of firm behavior that interweave blind and deliberate processes. Indeed, in human problem solving itself, both elements are involved and difficult to disentangle. Relatedly, Nelson and Winter describe their theory as unabashedly Lamarckian: The evolutionary economics theory of the firm contemplates both the inheritance of acquired characteristics and the timely appearance of variations under the stimulus of adversity.

Nelson and Winter's (1982) general term for all regular and predictable behavioral patterns of firms is *routine*. Nelson and Winter use this general term to include characteristics of firms that range from well-specified technical routines for producing things to procedures for hiring and firing, ordering new inventory, or stepping up production of items in high demand to policies regarding investment, research and development (R&D), or advertising to business strategies about product diversification and overseas investment. In Nelson and Winter's evolutionary theory, these routines play the role that genes play in biological evolutionary theory. They are a persistent feature of the organism and determine its possible behavior (though actual behavior is determined also by the environment).

Most of what is regular and predictable about business behavior is plausibly subsumed under the heading *routine*. That not all business behavior follows regular and predictable patterns is accommodated in evolutionary theory by recognizing that there are stochastic elements both in the determination of decisions and of decision outcomes. From

the perspective of a participant in business decision making, these stochastic elements may reflect the result of tumultuous meetings or of confrontations with complex problems under crisis conditions, but, from the viewpoint of an external observer seeking to understand the dynamics of the larger system, these phenomena are difficult to predict. Whereas in orthodox theory, decision rules are assumed to be the consequence of maximization, in evolutionary theory decision rules are treated as reflecting at any moment in time the historically given routines governing the actions of a business firm. Routine-changing processes are modeled as searches. Nelson and Winter's (1982) concept of search is the counterpart of that of mutation in biological evolutionary theory. Through the joint action of search and selection, the firms evolve over time, with the conditions of the industry in each period bearing the seeds of its condition in the following period.

Just as some orthodox microeconomic ideas seem to find their most natural mathematical expression in the calculus, the foregoing verbal account of economic evolution seems to translate naturally into a description of a Markov process—though one in a rather complicated state space. The process is not deterministic; search outcomes, in particular, are partly stochastic. Thus, what the industry condition of a particular period really determines is the probability distribution of its condition in the following period. Important antecedents of Nelson and Winter (1982) have been described in previous chapters:

- Behavioral theory of the firm (Cyert & March, 1963; Simon, 1947)
- Transaction costs theory (Williamson, 1975)
- Theory of the growth of the firm (Penrose, 1959)
- Business history (Chandler, 1962)

Chandler (1962) demonstrates that the organizational capabilities of a firm are embedded in its organizational structure, which is better adapted to certain strategies than to others. Thus, strategies at any point in time are constrained by the organization. Also, a significant change in a firm's strategy is likely to call for a significant change in its organizational structure.

Nelson and Winter (1982) build on the concept of Schumpeterian competition. Schumpeter's (1934) credentials as a theorist of bounded rationality could hardly be more incisively established than in the following passage:

The assumption that conduct is prompt and rational is in all cases a fiction. But it proves to be sufficiently near to reality, if things have time to hammer logic into men. Where this has happened, one may rest content with this fiction and build theories upon it. . . . Outside of these limits our fiction loses its closeness to reality. To cling to it there also, as the traditional theory does, is to hide an essential thing and to ignore a fact which, in contrast with other deviations of our assumptions from reality, is theoretically important and the source of the explanation of phenomena which would not exist without it. (p. 80)

Nelson and Winter (1982) observe that a consistent theme in retrospective studies is that failure occurs not because the intelligence system failed to acquire warning signals but because the intelligence system failed to process, relate, and interpret those signals into a message relevant to available choices. Intelligence analysts and decision makers have only a limited amount of time each day, limited communication channels to connect their systems, and limited assistance in the task of organizing, analyzing, and thinking about the available information. Sometimes, highly obvious and emphatic signals get lost in the noise as a result of these limitations. The events of September 11, 2001, are a compelling recent example. Nelson and Winter see no reason to think that economic decision making is any different in this regard.

There is similarly a fundamental difference between a situation in which a decision maker is uncertain about the state of X and a situation in which the decision maker has not given any thought to whether X matters or not. To treat them calls for a theory of attention, not a theory that assumes that everything always is attended to but that some things are given little weight (for objective reasons). In short, the most complex models of maximizing choice do not come to grips with the problem of bounded rationality.

Skills. Nelson and Winter (1982) develop the basic postulates about behavior in evolutionary theory. Although evolutionary economics theory is concerned with the behavior of business firms and other organizations, Nelson and Winter find it useful to begin the analysis with a discussion of some aspects of individual behavior. An obvious reason for doing so is that the behavior of an organization is, in a limited but important sense, reducible to the behavior of the individuals who are members of that organization. Regularities of individual behavior must therefore be expected to have consequences, if not counterparts, at the organizational level (see Dosi, Nelson, & Winter, 2000). Nelson and Winter (1982) propose that individual skills are the analogue of organizational routines and that an understanding of the role that routinization plays in organizational functioning is therefore obtainable by considering the role of skills in

individual functioning. By *skill* Nelson and Winter (1982) mean an ability to achieve a smooth sequence of coordinated behavior that is ordinarily effective relative to its objectives, given the context in which the skill normally occurs. Thus, the ability to serve a tennis ball is a skill, as is the ability to engage in competent carpentry, drive a car, operate a computer, set up and solve a linear programming model, or judge which job candidate to hire. Important characteristics of skills are as follows:

- Skills are programmatic (i.e., a sequence of closely followed steps).
- The knowledge that enables a skillful performance is, in large measure, tacit knowledge, in the sense that the performer is not fully aware of the details of the performance and finds it difficult or impossible to articulate a full account of those details.
- The exercise of a skill often involves the making of numerous choices, but to a considerable extent the options are selected automatically and without awareness that a choice is being made.

Nelson and Winter (1982) note that these three aspects of skilled behavior are closely interrelated. Skilled human performance is automatic in the sense that most of the details are executed without conscious volition. Indeed, a welcome precursor of success in an effort to acquire a new skill is the diminishing need to attend to details. Although impressiveness is obviously a matter of degree and relative to expectation, only the most unmoving can escape being impressed, at some point, by a skillful performance.

The late scientist and philosopher Michael Polanyi (1962) wrote extensively of the central place in the general scheme of human knowledge occupied by knowledge that cannot be articulated—tacit knowledge. On the simple observation that we know more than we can tell, Polanyi built an entire philosophical system. Polanyi notes that to be able to do something, and at the same time be unable to explain how it is done, is more than a logical possibility—it is a common situation. Polanyi offers a good example early in the discussion of skills:

I shall take as my clue for this investigation the well-known fact that the aim of a skillful performance is achieved by the observance of a set of rules, which are not known as such to the person following them. For example, the decisive factor by which the swimmer keeps himself afloat is the manner by which he regulates his respiration; he keeps his buoyancy at an increased level by refraining from emptying his lungs when breathing out and by inflating them more than usual when breathing in; yet this is not generally known to swimmers. (p. 49)

Nelson and Winter (1982) note that the tacitness of a skill, or rather of the knowledge enabling a skill, is a matter of degree. Words are probably a more effective vehicle for communicating the skills of elementary algebra than for those of carpentry and more effective for carpentry than for gymnastic stunts. Also, a trait that distinguishes a good instructor is the ability to discover introspectively, and then articulate for the student, much of the knowledge that ordinarily remains tacit. Skill involves the observance of a set of rules, which are not known as such to the person following them.

What are some determinants of the degree of tacitness? First, there is a limit imposed by the feasible time rate of information transfer through symbolic communication, which may be well below the rate necessary or appropriate in the actual performance.

A second consideration that limits the articulation of the knowledge underlying a skill is the limited causal depth of the knowledge. Polanyi's (1962) swimming example illustrates the point that the possession of a skill does not require theoretical understanding of the basis of the skill. Yet this does not imply that an attempt to articulate the basis of the skill would not benefit from the availability of this terminology. Perhaps some novice swimmers would be helped by Polanyi's brief explanation of the body's buoyancy.

The third aspect of the limitation of articulation is the coherence aspect—that of the whole versus the parts. Efforts to articulate complete knowledge of something by exhaustive attention to details and thorough discussion of preconditions succeed only in producing an incoherent message. This difficulty is probably rooted to a substantial extent in the related facts of the linear character of language-based communication, the serial character of the central processor of the human brain, and the relatively limited capacity of human short-term memory. Given these facts, the possibilities of articulating both the details and the coherent patterns they form—the relationships among the details—are necessarily limited. In short, much operational knowledge remains tacit because the knowledge cannot be articulated fast enough, because the knowledge is impossible to articulate all that is necessary to a successful performance, and because language cannot simultaneously serve to describe relationships and characterize the related things.

The knowledge contained in the how-to book and its various supplements and analogues tends to be more adequate when the pace of the required performance is slow and pace variations are tolerable; when a standardized, controlled context for the performance is somehow

assured; and when the performance as a whole is truly reducible to a set of simple parts that relate to one another in some very simple ways. To the extent that these conditions do not hold, the role of tacit knowledge in the performance may be expected to be large.

Finally, it should be emphasized that economic costs matter. Whether a particular bit of knowledge is in principle articulable or necessarily tacit is not the relevant question in most behavioral situations. Rather, the question is whether the economic costs associated with the obstacles to articulation are sufficiently high so that the knowledge, in fact, remains tacit. There is in a sense a trade-off between ability and deliberate choice, a trade-off imposed ultimately by the fact that rationality is bounded. The advantages of skills are attained by suppressing deliberate choice, confining behavior to well-defined channels, and reducing option selection to just another part of the program (March & Simon, 1958). Orthodox microeconomic theory treats the skillful behavior of the businessman as maximizing choice, and choice carries connotations of deliberation. Nelson and Winter (1982), on the other hand, emphasize the automaticity of skillful behavior and the suppression of choice that this skillful behavior involves.

Organizational Capabilities and Behavior. The organizations that Nelson and Winter (1982) envisage are those that face a substantial co-ordination problem, typically because these organizations have many members, performing many distinct roles, who make complementary contributions to the production of a relatively small range of goods and services. Nelson and Winter provide several salient functions of routines:

1. *Routine as Organizational Memory.* The routinization of activity in an organization constitutes the most important form of storage of the organization's specific operational knowledge. Basically, Nelson and Winter (1982) claim that organizations remember by doing. Exercise of a routine serves as parsimonious organizational memory. Recall that Arrow (1974) gave particular emphasis to the internal dialectic or code of an organization as a key resource of the economies that formal organization provides and as an important cause of persistent differences among organizations.
2. *Routine as Truce.* Routine operation involves a comprehensive truce in intraorganizational conflict (Cyert & March, 1963). Adaptations that appear obvious and easy to an external observer may be foreclosed because such adaptations involve a perceived threat to internal political stability.
3. *Routine as Target: Control, Replication, and Imitation.* Nelson and Winter (1982) note that replication is often a nontrivial exercise. Polanyi (1962) observed the following:

The attempt to analyze scientifically the established arts has everywhere led to similar results. Indeed, even in modern industries the indefinable knowledge is still an essential part of technology. I have myself watched in Hungary a new, imported machine for blowing electric lamp bulbs, the exact counterpart of which was operating successfully in Germany, failing for a whole year to produce a single flawless bulb. (p. 52)

The assumption that perfect replication is possible in evolutionary models is intended primarily to reflect the advantages that favor the going concerns that attempt to do more of the same, as contrasted with the difficulties that they would encounter in doing something else or that others would encounter in trying to copy their success. There are some potential obstacles to replication that may be difficult to overcome even at very high cost. Some employees at the old plant may be exercising complex skills with large tacit components, acquired through years of experience in the firm. Others may have skills of lesser complexity and tacitness but are poor at teaching those skills to someone else—doing and teaching are, after all, different. Some members, for various reasons, may be unwilling to cooperate in the process of transferring their segment of the memory contents to someone else; they may, for example, be unwilling to disclose how easy their job really is or the extent of the shortcuts they take in doing it. Williamson (1975) addresses the question of the incentives of organization members to disclose idiosyncratic information of importance to the organization's functioning under the rubric information impactedness. Nelson and Winter (1982) note that the target routine may involve so much idiosyncratic and impacted tacit knowledge that even successful replication, let alone imitation from a distance, is highly problematic.

4. *Routines and Skills: Parallels.* Nelson and Winter (1982) note that routines are the skills of an organization. Organizations are poor at improvising coordinated responses to novel situations; an individual lacking skills appropriate to the situation may respond awkwardly, but an organization lacking appropriate routines may not respond at all.
5. *Optimal Routines and Optimization Routines.* The heart of Nelson and Winter's (1982) proposal is that the behavior of firms can be explained by the routines that these firms employ. Modeling the behavior of the firm means modeling the routines and how these firms change over time.
6. *Routines, Heuristics, and Innovation.* According to Nelson and Winter (1982), innovation involves change in routine. Similarly, Schumpeter (1934) identified innovation with the “carrying out of new combinations” (pp. 65–66). A *heuristic* is any principle or device that contributes to reduction in the average search to solution. Schumpeter (1950) proposed that sometime during the 20th century the modern corporation routinized innovation.
7. *Routines as Genes.* Nelson and Winter (1982) argue that as a first approximation, firms may be expected to behave in the future according to the routines they have employed in the past. Efforts to understand the functioning of industries and larger systems should come to grips with the fact that highly flexible adaptation to change is not likely to characterize the behavior of individual firms. Evolutionary theory does come to grips with this fact.

Static Selection Equilibrium. Nelson and Winter (1982) note that in Friedman (1953) there is no hint that an evolutionary theory is an alternative to orthodoxy. Rather, the proposition is that selection forces may be the proper explanation of why orthodox theory is a good predictive engine. Alchian (1950) sets forth a perspective regarding firm behavior that resembles Nelson and Winter's in many ways, stressing the element of luck in determining outcomes, the role of learning by trial and feedback and imitation in guiding firms to do better, and of selection forces in molding what firms and industries do. Alchian states the following:

What really counts is the various actions actually tried, for it is from these that success is selected, not from some set of perfect actions. The economist may be pushing his luck too far in arguing that actions in response to changes in environment and changes in satisfaction with the existing state of affairs will converge as a result of adaptation or adoption towards the optimum action that would have been selected if foresight had been perfect. (p. 218)

This statement is not an argument that selection forces provide a reason for adherence to orthodox theory but rather a suggestion that there may be some important differences between an orthodox and an evolutionary perspective. Selection works on what exists, not on the full set of what is theoretically possible (Langlois, 1986; O'Driscoll & Rizzo, 1985).

Competition is viewed as a dynamic process involving uncertainty, struggle, and disequilibrium, not as a tranquil equilibrium state. In evolutionary theory, decision rules are viewed as unresponsive, or inappropriate, to novel situations or situations encountered irregularly and as a legacy from the firm's past and hence appropriate, at best, to the range of circumstances in which the firm customarily finds itself.

The heart of the R&D innovation problem is that reasonable people will disagree about which techniques will be best at which point. Importantly, this uncertainty is a major reason why it makes sense to have R&D largely conducted by competitive business firms who make their own entrepreneurial decisions, rather than place R&D decisions under more centralized control (see Nelson, 1996).

Dynamic Competition and Technical Progress. The market system is (in part) a device for conducting and evaluating experiments in economic behavior and organization. The meaning and merit of competition must be appraised accordingly. In Schumpeter's (1934) terms, competition involves carrying out new combinations. Schumpeter's concept of innovation was a broad one, noting five identified cases:

- The introduction of a new good
- The introduction of a new method of production
- The opening of a new market
- The opening of a new source of supply
- The carrying out of the new organization of any industry, like the creation of a monopoly position (p. 66)

Although Schumpeter (1934) is particularly noteworthy for this emphasis on experimentation, most of the great economists, from Adam Smith (1776/1937) to the onset of the modern period of formalization, gave some weight to the experimental role of competitive markets. An essential aspect of Schumpeterian competition is that firms do not know *ex ante* whether it pays to try to be an innovator or an imitator or what levels of R&D expenditures might be appropriate. Only the course of events over time will determine and reveal which strategies are the better ones. And even the verdict of hindsight may be less than clear.

Normative Organizational Economics From an Evolutionary Perspective. Nelson and Winter (1982) note that the modern advocacy of private enterprise solutions tends to suffer from vagueness or utopianism in its treatment of institutional matters. Three particularly important (and closely interrelated) ones involve the treatment of property rights, contracts, and law enforcement. In almost all formalized economic theories, property rights and contractual obligations are assumed to be costlessly delineated in unambiguous terms, and enforcement of the civil and criminal law is perfect and costless. By virtue of the combined force of these assumptions of clarity, perfection, and zero transaction costs, the problem of providing the basic institutional underpinnings of a system of voluntary exchange is assumed away. It is then not too surprising that voluntary exchange can be shown to be a largely effective economic solution to such problems as are left.

A legal system that could approach the theoretical standards of clarity and perfection in the delineation and enforcement of entitlements would be an elaborate and expensive system indeed. This is particularly obvious if the system of entitlements is supposed to be so sophisticated as to bring within its scope all of the externality problems that economists sometimes treat as merely problems in the definition and enforcement of property rights—for example, the question of whether a chemical plant is entitled to dispose of its hazardous wastes in ways that contaminate the groundwater or whether neighboring

property owners are entitled to uncontaminated groundwater. If the anatomy of market failure is a function of institutional structure, institutional structure itself evolves in part in response to perceived problems with the status quo.

Nelson and Winter (1982) conclude that the attempt to optimize and accordingly to control technological advance will, according to evolutionary theory, lead not to efficiency but to inefficiency. In terms of empirical testing of evolutionary organizational economics, Nelson and Winter note that organizations that operate many very similar establishments—for example, retail and fast-food chains—provide a natural laboratory for studying the problems of control and replication. Students with interest in the area of resources and organizational capabilities should see Foss (1997), Langlois and Robertson (1995), and Nelson and Winter (2002) for an update on recent research literature on dynamic capabilities and evolutionary economics.

Theory and Applications

Resource-based theory addresses some of the fundamental issues in strategy (Rumelt, Schendel, & Teece, 1994; Teece, 2000). Taking 1982 (when Nelson and Winter, 1982, was published) as the starting point, I now discuss some seminal contributions to resource-based theory:

- Lippman and Rumelt (1982): Causal ambiguity inherent in the creation of productive processes is modeled by attaching an irreducible *ex ante* uncertainty to the level of firm efficiency that is achieved by sequential entrants. Without recourse to scale economies or market power, the model generates equilibria in which there are stable interfirm differences in profitability. Sustainable competitive advantage results from the rich connections between uniqueness and causal ambiguity (see also Reed & DeFillippi, 1990; Rumelt, 1984).
- Teece (1982): This article outlines a theory of the multiproduct firm. Important building blocks include excess capacity and its creation, market imperfections, and the characteristics of organizational capabilities, including its fungible and tacit character. Teece both heavily acknowledges and builds on Penrose (1959) and argues that a firm's capabilities are upstream from the end

product—organizational capabilities might well find a variety of end-product applications, as Penrose's (1960) case study of the Hercules Powder Company effectively shows.

- Wernerfelt (1984, 1995): Building on the seminal work of Penrose (1959), these works argue that strategy involves a balance between the use of existing resources and the development of new resources.
- Montgomery and Wernerfelt (1988): According to resource-based theory (Teece, 1982), firms diversify in response to excess capacity of resources that are subject to market frictions. By probing into the heterogeneity of these resources, this article develops the corollary that firms that diversify most widely should expect the lowest average (Ricardian) rents. An empirical test, with Tobin's q as a measure of rents, is consistent with this resource-based theory.
- Dierickx and Cool (1989): This article draws the distinction between tradeable and nontradeable resources (e.g., reputation) and argues for a time-based view of competitive strategy (due, in part, to time compression diseconomies).
- Cohen and Levinthal (1990): The authors argue that prior related knowledge confers an ability to recognize the economic value of new information, assimilate the information, and apply the information to commercial uses. These dynamic capabilities constitute a firm's absorptive capacity. Cross-sectional data on technological opportunity and appropriability conditions in the American manufacturing sector collected for R&D lab managers and the FTC Line-of-Business data indicate that R&D both generates innovation and facilitates learning.
- Henderson and Clark (1990): This article distinguishes between the components of a product and the ways that the components are integrated into the system that is the product architecture. Data were collected during a 2-year, field-based study of the photolithographic alignment equipment industry. The core of the data is a panel data set consisting of research and development costs and sales revenue by product for every product development project

(Continued)

(Continued)

conducted between 1962, when the work on the first commercial product began, and 1986. The concept of architectural innovation provides rich resource-based connections between innovation and organizational capabilities.

- Barney (1991): In this often-cited article, Barney suggests that the search for sources of sustainable competitive advantage must focus on resource heterogeneity and immobility. Barney argues that sustainable competitive advantage is derived from resources that are valuable, rare, imperfectly imitable (due to path-dependence, causal ambiguity, and social complexity), and nonsubstitutable.
- Chatterjee and Wernerfelt (1991): This article theoretically and empirically investigates the resource-based view that firms diversify, in part, to use excess productive resources. In particular, empirical evidence corroborates that excess physical resources and most knowledge-based resources lead to more related diversification.
- Conner (1991): In this article, Conner analyzes resource-based theory as a new theory of the firm and makes insightful connections between resource-based theory and Schumpeterian (1934, 1950) competition.
- Montgomery and Hariharan (1991): Using a sample of 366 firms in the FTC's Line-of-Business database, the research in this article indicates that growth and diversification in large established firms result from a process of matching a firm's lumpy (indivisible) and ever-changing resources with dynamic market opportunities. Overall, this research provides empirical support for Penrose's (1959) theory of diversified entry: Unused productive services of resources are a selective force in determining the direction of firm-level expansion.
- Porter (1991): In this article, Porter argues that firms have accumulated differing resources because of differing strategies and configurations of (value-chain) activities. Resources and activities are, in a sense, duals of each other.
- Williamson (1991): This article suggests the possibility that the dynamic capabilities and resource-based perspectives will play

out in combination. Williamson argues that in the long run, the best strategy for firms is to organize and operate efficiently.

- Leonard-Barton (1992): This article considers core organizational capabilities in terms of employee knowledge and skills, technical systems, managerial systems, and values and norms. Leonard-Barton maintains that managers of new product and process development projects should take advantage of core capabilities while mitigating core rigidities. Twenty case studies of new product and process development projects in five firms (e.g., Chaparral Steel, Ford Motor Company, and Hewlett Packard) provide illustrative data. (For students who find this topic of interest, Leonard-Barton [1995], is an exemplar research book.)
- Mahoney (1992c): In this article, I argue for an integrated organizational economic approach to strategic management based on the behavioral theory of the firm, transaction costs theory, property rights theory, agency theory, and resource-based theory/dynamic capabilities. Essentially, this article outlines the structure of this book.
- Mahoney and Pandian (1992): Following Rumelt (1984), the authors of this paper argue that absent government intervention, isolating mechanisms (e.g., resource position barriers, invisible assets) exist because of asset specificity and bounded rationality.
- Amit and Schoemaker (1993): This article adds behavioral decision-making biases and organizational implementation aspects as further impediments to the transferability or imitability of a firm's resources and capabilities.
- Mosakowski (1993): Using a longitudinal data set, a sample of 86 entrepreneurial firms in the computer software industry that completed an IPO in 1984 is examined. Empirical findings suggest that strategies that represent rare, inimitable and nonsubstitutable resources are a source of competitive advantage.
- Peteraf (1993): This article elucidates the organizational economics logic that is the foundation for the resource-based

(Continued)

(Continued)

theory of Ricardian rents (Ricardo, 1817) and sustainable competitive advantage. The essence of the framework developed here is that four conditions must be met for achieving sustainable competitive advantage: (1) superior resources (firm heterogeneity within an industry), (2) ex post limits to competition (i.e., isolating mechanisms), (3) imperfect resource mobility (e.g., nontradeable assets and cospecialized assets), and (4) ex ante limits to competition.

- Chi (1994): In this article, Chi develops a theoretical framework for analyzing the exchange structure in the trading of imperfectly imitable and imperfectly mobile firm resources. The article first explores the conditions for such resources to be gainfully traded between firms and then investigates the interconnections between barriers to imitation and impediments to trading. A major part of the article is devoted to developing a parsimonious and yet integrative (agency, property rights, and transaction costs) model for assessing the exchange structure between firms that are involved in the trading of strategic resources in the face of significant transaction cost problems, such as adverse selection, moral hazard, contractual cheating, and hold-up problems that are due to information asymmetry, imperfect measurement, imperfect enforcement, and resource interdependencies.
- Farjoun (1994): This article provides empirical support that unused productive services derived from human capital drive the diversification process. Unused productive services from existing human resources present a jigsaw puzzle for balancing processes.
- Henderson and Cockburn (1994): Using both qualitative and quantitative data drawn from both public sources and from the internal records of 10 major European and American pharmaceutical firms, this article attempts to measure the importance of heterogeneous, organizational capabilities. Component and architectural capabilities together explain a significant fraction of the variance in research productivity across firms.
- Godfrey and Hill (1995): This article persuasively espouses the realist philosophy of science, which states that we cannot

reject theories just because they contain key constructs that are unobservable.³ It is not enough to state that the unobservability of utility dooms agency theory, that transaction costs theory is untestable because some transaction costs cannot be measured, or that resource-based theory is invalid because key resources (e.g., invisible assets) are unobservable. To reject a theory one must be able to show that the predictions of observable phenomena that are derived from the theory do not hold up under empirical testing.

- Mahoney (1995): In this article, I argue that the resource-based approach of deductive economics, the dynamic capabilities approach of strategy process, and organization theory research on organizational learning (e.g., Argyris & Schon, 1978; Fiol & Lyles, 1985) need to be joined in the next generation of resource-based research.
- Zander and Kogut (1995): Based on their developed questionnaire distributed to project engineers knowledgeable of the history of 44 major innovations in 20 firms, the authors conclude that the transfer of manufacturing capabilities is influenced by the degree to which capabilities may be codified and taught. Empirical evidence corroborates the view that the nature of dynamic capabilities and the nature of competitive positioning matter.
- Foss (1996): The author argues that there are complementarities between a contractual approach (e.g., transaction costs theory and property rights theory) and a knowledge-based approach (e.g., resource-based theory and knowledge-based theory) to strategic management. These complementarities are argued to be particularly fruitful for analyzing the strategic issues of the boundary and internal organization of the firm.

³In addition to Godfrey and Hill's (1995) lucid discussion on realist philosophy, there are a number of works that cover various issues in philosophy of science and research methodology that are relevant to strategic management research, including Blaug (1980); Caldwell (1984); Camerer (1985); Evered and Louis (1981); Huff (1981, 2000); Kaplan (1964); Kuhn (1970); Ladd (1987); Machlup, (1967); MacKinlay (1997); Mahoney (1993); Mahoney and Sanchez (1997, 2004); McCloskey (1983, 1998); McCloskey and Ziliak (1996); Montgomery, Wernerfelt, and Balakrishnan (1989); Redman (1993); Seth and Zinkhan (1991); and Whetten (1989).

(Continued)

(Continued)

- Grant (1996): In this article, Grant argues that organizational capabilities are the outcome of knowledge integration: complex, team-based productive activities that cohesively integrate the knowledge of many individual specialists. Research in cross-functional capabilities in the context of new product development (Clark & Fujimoto, 1991) would be an exemplar.
- Miller and Shamsie (1996): This article empirically tests resource-based theory in the context of the seven major United States film studios (i.e., MGM, Twentieth Century-Fox, Warner Brothers, Paramount, United Artists, Universal, and Columbia) from 1936 through 1965. The authors find that property-based resources in the form of exclusive long-term contracts with celebrities and theaters helped financial performance in the stable environment from 1936 to 1950. In contrast, knowledge-based resources in the form of production and coordination talent boosted financial performance in the more uncertain posttelevision environment.
- Mowery, Oxley, and Silverman (1996): Examining cross-citation rates for 792 partners in bilateral alliances that involved at least one U.S. firm and were established during 1985 and 1986, this article provides empirical support for the importance of gaining capabilities through alliances. The empirical results bolster the argument that experience in related technological areas is an important determinant of absorptive capacity.
- Spender (1996): Building on Nelson and Winter (1982) and Nonaka and Takeuchi (1995), this article views the firm as a dynamic knowledge-based activity system. The author's arguments are consistent with Penrose's (1959) view of knowledge as the skilled process of leveraging resources, where that knowledge is embedded in the organization.
- Szulanski (1996): Based on 271 observations of 122 best-practice transfers in eight companies, the major barriers to internal knowledge transfer are found to be knowledge-related factors, such as the recipient's lack of absorptive capacity, causal ambiguity, and an arduous relationship between the source and the recipient.
- Helfat (1997): This empirical investigation of dynamic R&D capabilities examines the role of complementary know-how and

other resources in the context of changing conditions in the U.S. petroleum industry during the 1970s and early 1980s. The empirical analysis indicates that in response to rising oil prices, firms with larger amounts of complementary technological knowledge and physical resources also undertook larger amounts of R&D on coal conversion (a synthetic fuel process).

- Powell and Dent-Micallef (1997): This article examines the information technology literature, develops an integrative resource-based theoretical framework, and presents results from an empirical study of the retail industry. The empirical results support the view that information technology creates economic value by leveraging and using complementary human and physical resources.
- Teece, Pisano, and Shuen (1997): This article views the dynamic capabilities perspective as building on Schumpeter (1934, 1950), Nelson and Winter (1982), and Teece (1982). Focal concerns are resource accumulation, replicability, and inimitability of organizational capabilities.
- Tripsas (1997): This article analyzes the technological and competitive history of the global typesetter industry from 1886 to 1990. Key success factors include investment, technical capabilities, and appropriability through specialized complementary assets.
- Bogner, Mahoney, and Thomas (1998): In this article, following Machlup (1967), the authors argue that resource-based theory needs to move beyond (1) theoretical construction that abstracts from historical time, (2) theory that focuses only on the stationary state, (3) theory where taxonomic and tautological arguments are made, (4) theory that focuses exclusively on the conditions for establishing equilibrium, and (5) theory that omits time as an independent variable.
- Farjoun (1998): This article examines empirically the joint effect of skill-based and physical-based related diversification on accounting and financial measures of performance. For a sample of 158 large diversified manufacturing firms, the joint effort of skill-based and physical-based related diversification had a strong

(Continued)

(Continued)

positive effect on most indicators of performance. This finding corroborates resource-based theory that related diversification that builds on both skill-based and physical-based resources allows firms to create economic value by sharing and transferring these resources and to use activities and routines in which these resources interact.

- Lieberman and Montgomery (1998): Building on Lieberman (1987) and Lieberman and Montgomery (1988), the authors of this article argue that resource-based theory and first-mover (dis)advantage are related conceptual frameworks that can benefit from closer linkages.
- Argote (1999): This book presents evidence that organizations vary tremendously in the rate at which they learn. Argote argues that differences in patterns of knowledge creation, retention, and transfer contribute to differences in the rates at which organizations learn.
- Brush and Artz (1999): Using a sample of 193 veterinary practices, this article investigates contingencies among resources, capabilities, and performance in veterinary medicine. Empirical evidence supports the view that the economic value of resources and capabilities depends on the information asymmetry characteristics of the product market.
- Silverman (1999): This article considers how a firm's resource base affects the choice of industries into which the firm diversifies and offers two main extensions of prior resource-based research. First, the paper operationalizes technological resources at a more fine-grained level than in prior empirical studies, thereby enabling a more detailed analysis concerning the direction of diversification. This analysis indicates that the predictive power of resource-based theory is greatly improved when resources are measured at a more fine-grained level. Second, the article integrates transaction costs theory and resource-based theory to provide more detailed predictions concerning diversification. Empirical evidence suggests circumstances where resources (that have high asset specificity) can be and are used through contracting rather than through becoming a diversified firm.

- Williamson (1999): This article suggests that one way of looking at research opportunities in strategic management is to view transaction costs theory as feeding into the organizational capabilities perspective. Both transaction costs theory and resource-based theory are viewed as needed in our efforts to understand complex business phenomena as we build a science of organization.
- Yeoh and Roth (1999): This article empirically examines the impact of firm resources and capabilities using a sample of 20 pharmaceutical firms that operated as separate entrepreneurs between 1971 and 1989. The empirical results indicate that R&D and sales force expenditures have direct and indirect effects on sustainable competitive advantage.
- Ahuja and Katila (2001): Using a sample of acquisition and patent activities of 72 leading firms from the global chemicals industry from 1980 to 1991, the relatedness of acquired and acquiring knowledge-based resources has a nonlinear impact on innovation output. In particular, acquisition of firms with high levels of both relatedness and unrelatedness prove inferior to acquiring firms with moderate levels of knowledge-based relatedness.
- Bowman and Helfat (2001): This article examines the resource-based theory that there is a significant role for corporate strategy based on the use of common resources by related businesses within a firm (Peteraf, 1993; Teece, 1982). Based on an analysis of the variance decomposition research literature, Bowman and Helfat conclude that corporate strategy (Andrews, 1980; Ansoff, 1965), in fact, does matter for economic performance.
- Makadok (2001): This article provides a mathematical model synthesizing resource-based and dynamic capabilities views of economic value creation. Resource picking (emphasized by resource-based theory) and capability building (emphasized by the dynamic capabilities approach) for the purpose of achieving economic rent creation are shown to be complementary in some business circumstances but are shown to be substitutes in other business circumstances.

(Continued)

(Continued)

- Mahoney (2001): In this article, I argue that resource-based theory is primarily a theory of economic rents, whereas transaction costs theory is primarily a theory of the existence of the firm. These two theories are complementary and are connected in the following way: Resource-based theory seeks to delineate the set of market frictions that would lead to firm growth and sustainable economic rents (via isolating mechanisms), whereas transaction costs theory seeks to delineate the set of market frictions that explain the existence of the firm. The article submits that the set of market frictions that explain sustainable firm rents (in resource-based theory) will be sufficient market frictions to explain the existence of the firm (in transaction costs theory). I also argue that the resource-based theory of the strategic (rent-generating and rent-sustaining) firm cannot assume away opportunism.
- Afuah (2002): This article provides a model for mapping firm capabilities into competitive advantage. Using a sample of 78 observations for cholesterol drugs in the market from 1988 to 1994, the author illustrates how the model can be used to estimate competitive advantage from technological capabilities.
- Coff (2002): Empirical results from a sample of 324 acquisitions that closed or failed to close in the years 1988 and 1989 offer evidence in support of the hypothesis that related human capital expertise between the acquirer and acquired enterprise can mitigate opportunism hazards associated with human capital asset specificity (Becker, 1964). In this business setting, related knowledge-based resources, in the form of related human expertise, increases the probability that a given transaction will close.
- Madhok (2002): This article maintains that a strategic theory of the firm should not only address the decision with respect to hierarchical governance or market governance but should also take into account how a firm's resources and capabilities can best be developed and deployed in the search for competitive advantage. Or, put differently, transaction costs theory should be coupled with resource-based theory.
- Thomke and Kuemmerle (2002): Using a combination of field research, discovery data from nine pharmaceutical firms, and

data on 218 alliances involving new technologies for experimentation and testing, several causes affecting resource accumulation are identified and described. The article provides empirical support that the difficulty of imitating a particular resource is affected by the interdependencies with other resources.

- Adner and Helfat (2003): This article adds to the study of competitive heterogeneity by measuring the economic effect of specific corporate-level managerial decisions, driven by dynamic managerial capabilities, on the variance of economic performance among U.S. energy companies. The empirical results also strongly suggest that corporate managers matter.
- Helfat and Peteraf (2003): This article introduces the capability life cycle, which identifies general patterns and paths in the evolution of organizational capabilities over time. The framework is intended to provide a theoretical structure for a more comprehensive approach to dynamic resource-based theory.
- Hoopes, Madsen, and Walker (2003): This article maintains that the resource-based view's accomplishments are clearer when seen as part of a larger theory of competitive heterogeneity. Combining economics, organization theory, and traditional business policy, the resource-based view suggests how, in a competitive environment, firms maintain unique and sustainable positions.
- Knott (2003a): The author of this article finds that franchising routines are both valuable and can lead to sustainable competitive advantage. The upshot of this empirical research is that tacit knowledge is not necessary for having an isolating mechanism.
- Knott (2003b): This article outlines a theory of sustainable innovation fueled by persistent heterogeneity. Knott shows that there exist conditions that generate persistent heterogeneity and sustainable innovation with each firm behaving optimally, taking other firms' behaviors into account.
- Lippman and Rumelt (2003): This article critiques the microfoundations of neoclassical theory and develops further the

(Continued)

(Continued)

concept of rent. The article also provides insights on rent sensitivity analysis and a payments perspective of strategic management.

- Makadok (2003): This article models mathematically the joint impact of two determinants of profitable resource advantages: the accuracy of managers' expectations about the future economic value of a resource and the severity of agency problems that cause managers' interests to diverge from those of shareholders. The conclusion is that future research on the origins of competitive advantage should examine agency and governance issues along with, not apart from, resource-based issues.
- Szulanski (2003): This research book on sticky knowledge addresses an important question for managers: Why don't best practices spread within organizations? Szulanski explores the effect of motivational and knowledge barriers on knowledge transfer and presents the empirical results of statistical analyses that stem from data collected through a two-step questionnaire survey. The research relies on 271 surveys studying the transfer of 38 (technical and administrative) practices in eight companies. Szulanski finds that knowledge barriers to transfer have a larger effect on the stickiness of knowledge than motivational barriers, and the two barriers jointly explain nearly 75% of the variance in stickiness.

To conclude this chapter, I focus on a particularly important set of dynamic capabilities that are embedded in real options in strategic decision making. Trigeorgis (1996) provides both rigor and relevance concerning strategic (real) options.

Real Options: Managerial Flexibility and Strategy in Resource Allocation (*Trigeorgis, 1996*)

Financial theory, properly applied, is critical to managing in an increasingly complex and risky business climate. . . . Option analysis provides a more flexible approach to valuing our research investments. . . . To me all kinds of business decisions are options. (Judy Lewent, as cited in Nichols, 1994, and reprinted in Trigeorgis 1996, p. xv)

Trigeorgis (1996) deals with the classical subject of resource allocation or project appraisal under uncertainty, particularly with the economic valuation of managerial operating flexibility and strategic actions as corporate real options. Similar to options on financial securities, real options involve property rights (with no obligations) to acquire or exchange an asset for a specified alternative price. The ability to value options (e.g., to defer, abandon, and grow) has brought a revolution to modern corporate finance theory on resource allocation.

Corporate value creation and competitive positioning are critically determined by corporate resource allocation and by the proper evaluation of investment alternatives. Trigeorgis (1996) argues that traditional quantitative techniques such as discounted cash flow (DCF) analysis (that consider the size, timing, and uncertainty of cash flows) have failed in business practice because these techniques traditionally have not properly captured managerial flexibility to adapt and revise later decisions in response to unexpected market developments. Moreover, these techniques traditionally neither capture the strategic value resulting from proving a technology viable nor capture the economic impact of project interdependencies and competitive interactions. In the Nelson and Winter (1982) sense, organizational capabilities that enhance adaptability and strategic positioning provide the infrastructure for the creation, preservation, and exercise of corporate real options that can have significant economic value.

Trigeorgis (1996) notes that, in practice, managers have often been willing to overrule traditional investment criteria to accommodate operating flexibility and other strategic decisions that managers consider just as valuable as direct cash flows. It is now widely recognized, for example, that traditional DCF approaches to the appraisal of capital investment projects, such as the standard net-present-value (NPV) rule of accepting positive NPV projects, do not properly capture management's flexibility to adapt and revise later decisions in response to unexpected market developments. Or, put differently, a theoretically accurate NPV analysis would include real options values.

Trigeorgis (1996) argues that in the business marketplace, which is characterized by change, uncertainty, and competitive interactions, the realization of cash flows will probably differ from what management expected at the outset. As new information arrives and uncertainties about market conditions and future cash flows are gradually resolved, management may have valuable flexibility to alter its initial operating plan to capitalize on favorable future opportunities or to react so as to mitigate economic losses. For example, management may be able to

defer, expand, contract, abandon, or otherwise alter a project at various stages of the project's useful operating life.

This managerial operating flexibility is likened to financial options. An American call option of an asset (with current value V) gives the right, with no obligation, to acquire the underlying asset by paying a prespecified price (the exercise price, I) on or before a given maturity. Similarly, an American put option gives the right to sell (or exchange) the underlying asset and receive the exercise price. *The asymmetry derived from having the right but not the obligation to exercise an option is at the heart of the option's value.*

Trigeorgis (1996) notes that as with options on financial securities, management's flexibility to adapt its future actions in response to altered future market conditions and competitive reactions expands a capital investment opportunity's value by improving its upside potential while limiting its downside economic losses relative to the initial expectations of a passive management. The resulting asymmetry calls for a strategic investment criterion, reflecting both value components: the traditional static NPV of direct cash flows and the real option value of operating flexibility and strategic interactions.

Trigeorgis (1996) argues that a real options approach to capital budgeting has the potential to conceptualize and quantify the value of options from active management and strategic interactions. This economic value is typically manifest as a collection of real options embedded in capital investment opportunities, having as the underlying asset the gross project value of discounted expected operating cash inflows. Many of these real options (e.g., to defer, contract, shut down, or abandon a capital investment) occur naturally; other real options may be planned and built in at some extra cost from the outset (e.g., to expand capacity, to build growth options, to default when investment is staged sequentially, or to switch between alternative inputs or outputs). Let us now consider various real options.

1. *Option to Defer Investment.* The real option to defer an investment decision is analogous to an American call option on the gross present value of the completed project's expected operating cash flows, V , with an exercise price equal to the required outlay, I . Management holds a lease on (or an option to buy) valuable land or resources. Management can wait x years to see if output prices justify constructing a building or a plant or developing a field. The option to wait is particularly valuable in natural-resource extraction industries, farming, paper products, and real estate development.

Consider the following example from Dixit and Pindyck (1994): How should we decide whether or not to enter into a business? If we refer to the literature on finance, the traditional approach is to use cash flow analysis using a net present value criterion. For example, let us imagine a situation in which we are considering entering the business of making widgets. Assume that it costs \$1,600 to build a widget factory and that our current cost of capital is 10%. In addition, we sell only one widget per year, and the current price of a widget is \$200. Although we know the current price for widgets, we are somewhat uncertain about the future prices. Forecasts indicate that there is a 50% chance that prices will go up to \$300 next period (and remain there forever); however, there is also a 50% chance that prices will go down to \$100. This forecast implies that the expected price of widgets in the future is $\$200 = (0.5 \times \$300 + 0.5 \times \$100)$.

Using these numbers, we can evaluate this project with a standard cash flow analysis. The expected cash flow from entering the widget business appears in the first column of Table 5.1. In period 0, we build the plant ($-\$1,600$) and begin production, receiving \$200 in revenues ($-\$1,600 + \$200 = -\$1,400$). From that period on, we have positive expected cash flow of \$200. We can use this cash flow series to arrive at the NPV for the project, which is \$600. (Because the value at T_0 of a perpetuity cash flow [CF] beginning at T_1 with a discount rate r equals CF/r then here it is: $\$200/.1 = \$2,000$; then we take \$2,000 and subtract \$1,400 to arrive at \$600.) We would then proceed with the project because the NPV of \$600 is greater than zero.

However, what if we wait a period to find out whether the price goes up or down? That is, what if we choose to keep our options open and remain flexible in our decision? Two different scenarios could occur. The first possibility is that the price goes up to \$300, in which case we would experience the cash flow under Scenario 1 in Table 5.1. The second possibility is that the price goes down to \$100, in which case we obtain the cash flows under Scenario 2. Now, one will notice that under Scenario 1, the NPV (in period 0) is positive (i.e., $NPV = \$1,545$); however, under Scenario 2, the NPV is negative (i.e., $NPV = -\$455$). (The present value of the perpetuity is $\$300/.1 = \$3,000$, from which we subtract $\$1,600/.1 = -\$1,455$ to arrive at \$1,545, and for the low demand scenario we have $\$100/.1 = \$1,000$ and subtract $\$1,600/.1 = -\455 .) Thus, if we waited a period and the price went up to \$300, we would proceed with the project, whereas if the price went down, we would not proceed with the project. Thus, under the second scenario, the actual NPV would not be $-\$455$ but would be \$0; that is, we would not invest

Table 5.1 Calculations of Discounted Cash Flows

Time	Expected Cash Flow (Traditional NPV)	Expected Cash Flow (Scenario 1)	Expected Cash Flow (Scenario 2)
0	\$ (1,400)	\$ --	\$ --
1	\$ 200	\$ (1,300)	\$ (1,500)
2	\$ 200	\$ 300	\$ 100
3	\$ 200	\$ 300	\$ 100
4	\$ 200	\$ 300	\$ 100
5	\$ 200	\$ 300	\$ 100
6	\$ 200	\$ 300	\$ 100
7	\$ 200	\$ 300	\$ 100
8	\$ 200	\$ 300	\$ 100
Inf.	\$ 200	\$ 300	\$ 00
NPV	\$ 600	\$ 1,545	\$ (455)

in a negative NPV project. What does this tell us about the value of waiting and remaining strategically flexible?

One way of answering this question is to reframe our cash flow analysis. Instead of taking the NPV of the expected cash flows, let us calculate the expected NPV of the two scenarios combined. That is, we have a 50% chance of the price going up and getting an NPV of \$1,545 and a 50% chance of the price going down and getting \$0. The expected combined NPV is therefore approximately \$773 ($= 0.5 \times \$1,545 + 0.5 \times 0$). The NPV where we wait, find out the true price, and then make the decision is larger (by \$173) than going ahead right now. There is (an option) value to waiting of \$173. Thus, we can increase our expected returns by waiting a year and then deciding whether to undertake the sunk-cost investments in a new plant.

Summary. The previous example illustrates that even when the static (positive) NPV calculation suggests a go, when the real options value of flexibility is taken into account, the top-level manager should wait. The option to wait is equivalent to a call option on the investment project. The call is exercised when the firm commits to the project. But often it is better to defer a positive-NPV project to keep the call option alive. Deferral is most attractive when uncertainty is great and immediate project cash flows—which are lost or postponed by waiting—are small.

2. *Option to Default During Staged Construction (Time-to-Build Option).* Each stage of an investment can be viewed as an option on the economic value of subsequent stages by incurring the installment-cost

outlay (e.g., I_1) required to proceed to the next stage and can therefore be valued similar to options on options (or compound options). Staging the investment as a series of outlays creates the real option to abandon the project in midstream if new information is unfavorable. This real option is valuable in R&D-intensive industries (especially pharmaceuticals); in highly uncertain long-development, capital-intensive industries (such as energy-generating plants or large-scale construction); and in venture capital.

3. *Option to Expand, Contract, Shut Down, or Restart Operations.* If market conditions are more favorable than expected, the firm can expand the scale of production or accelerate resource use. Conversely, if conditions are less favorable than expected, the firm can reduce the scale of operations. In extreme cases, production may be halted and restarted. Applications can be found in natural-resource industries (e.g., mining), facilities planning and construction in cyclical industries, fashion apparel, consumer goods, and commercial real estate.

4. *Option to Abandon for Salvage Value.* Management may have a valuable real option to abandon a project in exchange for its salvage value. Naturally, more general-purpose capital assets would have a higher salvage value and abandonment option value than special-purpose assets. Valuable abandonment options are generally found in capital-intensive industries (such as airlines and railroads), in financial services, and in new-product introductions in uncertain markets. Abandonment should not be exercised lightly if it might lead to eventual erosion of valuable expertise and other crucial organizational capabilities that could be applied elsewhere in the business or that could prevent the firm from participating in future technological developments. Moreover, abandonment may lead to the loss of goodwill from customers.

5. *Option to Switch Use (e.g., Inputs or Outputs).* Generally, process flexibility can be achieved not only via technology (e.g., by building a flexible facility that can switch among alternative energy inputs) but also by maintaining relationships with a variety of suppliers and switching among them as their relative prices change. Process flexibility is valuable in feedstock-dependent facilities, such as oil, electric power, chemicals, and crop switching. Product flexibility—enabling the firm to switch among alternative outputs—is more valuable in industries such as machine parts, automobiles, consumer electronics, toys, specialty paper, and pharmaceuticals, where product differentiation and diversity

are important or product demand is volatile. In such business cases it might be worthwhile to install a more costly flexible capacity to acquire the dynamic capability to alter product mix or production scale in response to changing market conditions.

6. *Corporate Growth Options.* Corporate growth options that set the path of future opportunities are of considerable strategic importance. Although in isolation a proposed facility may appear unattractive, such a facility may be only the first in a series of similar facilities if the process is successfully developed and commercialized, and it may even lead to entirely new by-products. Many early investments (e.g., in R&D) can be seen as prerequisites or links in a chain of interrelated projects. The value of the early projects derives not so much from their expected directly measurable cash flows as from the future growth opportunities they may unlock (e.g., access to a new market or strengthening of the firm's core capabilities and its strategic positioning). An opportunity to invest in a first-generation high-tech product, for example, is analogous to an option on options (an interproject compound option). Despite a negative static NPV, the infrastructure, experience, and potential by-products generated during the development of the first-generation product may serve as springboards for developing lower cost or higher quality future generations of that product, or even for generating entirely new applications. But unless the firm makes the initial investment, subsequent generations or other business applications will not even be feasible. The infrastructure and experience gained, if maintained as proprietary knowledge, can place the firm at a competitive advantage, which may even reinforce itself when learning-cost-curve effects are present. Growth options are found in all infrastructure-based or strategic industries—especially in high technology, R&D, industries with multiple product generations or applications (e.g., semiconductors, computers, pharmaceuticals), multinational operations, and strategic acquisitions.

7. *Multiple Interacting Options.* Real-life projects often involve a collection of various options. Upward-potential-enhancing and downward-protection options are present in combination. Their combined economic value may differ from the sum of their separate values (i.e., they interact). They may also interact with financial flexibility options. Applications include most industries listed previously.

Trigeorgis (1996) argues that real options have the potential to make a significant difference in strategic management. Sustainable competitive advantages resulting from proprietary technologies, scale, ownership of

valuable natural resources, managerial capital, reputation, brand name, or patents (Andersen, 2001; Arora, Fosfuri, & Gambardella, 2001) empower companies with valuable real options to grow through future profitable investments and to more effectively respond to unexpected adversities or opportunities in a changing technological, competitive, or general business environment. Students studying the economics of organization have ample opportunity to supplement real options analysis (i.e., often decision-theoretic) with game-theoretic tools capable of incorporating strategic competitive responses, and this research area promises to be an important and challenging direction for strategic management and corporate finance research.

Applications of the Real Options Perspective

- In 1984 the W. R. Grace Corporation made an investment in a new technology for automotive catalytic converters. Although the technology proved uncompetitive on price in the automotive market, new applications arose in cogeneration plants and in utility emission control as a result of the U.S. Clean Air Act.
- In research and development, many high-technology companies invest heavily in technologies that may result in a wide range of possible outcomes and new potential markets but with a high probability of technical or market failure. In the pharmaceutical industry, for example, on average, it costs \$360 million and takes a decade to bring a new drug to the market. Only 1 explored chemical in 10,000 becomes a prescription drug, and once a drug reaches the market it faces a 70% chance of failing to earn the cost of invested capital. Such investments are hard to sell to top management on financial grounds; their benefits are remote and hard to quantify, even though intuitively their growth potential seems promising. Instead of ignoring these technologies, a company can make a capital commitment in stages, effectively taking a call option on the underlying technology (or future applications). The initial outlay is not made so much for its own cash flows as for its growth-option value.
- Merck and Company embarked on extensive automation, starting with a drug packaging and distribution plant, even though

(Continued)

(Continued)

technical success was uncertain and projected labor savings did not seem to justify the investment. Operations valuation allowed engineers to articulate the whole range of outcomes and their benefits. Indeed, building and using real options-based planning methods were viewed as having created a valuable new capability for Merck (Nichols, 1994). In fact, the more uncertain the technology or the future market demand, the higher the value of the real option.

- The case of the adoption of the thin-slab caster by Nucor involved consideration of sunk-cost commitments and real options (Ghemawat, 1997):
 - *The Option to Wait (and Learn) Before Investing.* In the Nucor case, it was very unlikely that another firm would be willing to be the pioneering firm to deploy this new technology. Thus, if Nucor were to wait, the reduction in uncertainty would have been small.
 - *The Option to Make Follow-On Investments if the Immediate Investment Project Succeeds.* In the Nucor case, the first plant appeared to have a slightly negative NPV on a stand-alone basis across a majority of likely scenarios. However, the experience gained in building the first plant would substantially improve the economics of subsequent plants. Thus, the first plant could merely be the price of admission representing a necessary learning curve. Thus, even though the narrow (negative) NPV calculation for Nucor suggested a no-go, the growth options tipped the scale to go. Nucor took into account the strategic value of taking on this negative-NPV project. A close look at Nucor's payoffs reveals a call option on follow-on projects in addition to the immediate project's cash flows. Today's investments can generate tomorrow's opportunities.
 - *The Option to Abandon the Project.* Even if the compact strip production (CSP) fails, that component of the mill could be potentially replaced by another technology; the bulk of the mill, such as electric arc furnace and rolling mills, may be useable even with another thin-slab technology. Thus, when the narrow (negative) NPV calculation suggests a no-go, a high options value of abandonment (i.e., low switching

costs) may tip the scale to go. The option to abandon a project provides partial insurance against failure. This is a put option; the put option's exercise price is the value of the project's assets if sold or shifted to a more valuable use.

The flexibility provided by flexible manufacturing systems, flexible production technology, or other machinery having multiple uses can be analyzed from the real options perspective. Recently, the flexibility created by modular design that connects components of a larger system through standardized interfaces (and its impact on organization design) has captured attention in strategic management (Baldwin & Clark, 2000; Bowman & Kogut, 1995; Garud & Kumaraswamy, 1995; Garud, Kumaraswamy, & Langlois, 2003; Langlois, 2002; Sanchez & Mahoney, 1996, 2001; Schilling, 2000; Worren, Moore, & Cardona, 2002). Students studying the economics of organization have an opportunity to evaluate such flexibility using the real options framework.

In conclusion, the current academic research literature in corporate finance has largely framed real-options problems as decision theoretic. However, we now need to move on to considerations where the timing of investments also depends on how other players will respond. Thus, strategic management must take into account both decision-theoretic problems and game-theoretic problems in the next generation of real options research.⁴

Concluding Comments. The resource-based, dynamic capabilities, and real options literatures are potentially highly synergistic for theory development, empirical testing, and business applications. Students with research interests in business history (e.g., Chandler, 1990), evolutionary theory and organizational capabilities (e.g., Nelson & Winter, 1982), corporate finance (e.g., Trigeorgis, 1996), strategic human resource management (Baron & Kreps, 1999), and entrepreneurship (Penrose, 1959) are anticipated to contribute to the evolving science of organization.

⁴For strategic management contributions to the real options perspective, see Bowman and Hurry (1993), Chi (2000), Folta (1998), Folta and Miller (2002), Kogut (1991), McGrath (1997, 1999), Miller (2002), Miller and Folta (2002), and Sanchez (1993, 2003). Schwartz and Trigeorgis (2001) provide a number of classical readings and recent contributions on real options and investment under uncertainty.





6

The Theoretic Building Blocks of Organizational Economics

In this research book, I have maintained that an integrated organizational economics research program is a feasible, challenging, and rewarding endeavor in pursuing the evolving science of organization. Toward that end, five major theories of the firm were reviewed and interrelated: (1) the behavioral theory of the firm, (2) transaction costs theory, (3) property rights theory, (4) agency theory, and (5) dynamic resource-based theory. In my judgment, content research (e.g., deductive agency and transaction costs economics) and process research (e.g., research on cognitive psychology and the behavioral theory of the firm) need to be joined in the next generation of organizational economics research.

My major philosophical position in this book is that for claims that theories are incommensurable (Kuhn, 1970) the burden of proof should be placed on the scholars making such claims. Thus, I concur with Popper's (1970) strong dissent of Kuhn's incommensurability thesis. In my judgment, bridges can and should be built between contested intellectual terrains.

Although this research book provides a rudimentary first step toward intertwining the various strands of the organizational economics research literature into one cord, I anticipate that the next generation of students will go well beyond the mere comprehension of the various theories of the firm and will successfully apply the organizational economics perspective to emerging business phenomena (e.g., in research studies of information technology and its impact on organizational structures and to the study of transition economies).

In fact, the growing interest in the organizational economics view of strategic management is arguably due to the fruitfulness of organizational economics theory in contributing to our understanding of

currently observed changes at the organizational and institutional levels of analysis. The reason for such a knowledge claim is not hard to find. Even though technological, organizational, and institutional changes advance breathlessly, the organizational economics principles provided in this research book are durable principles that have stood, and will continue to stand, the test of time. Moreover, I also anticipate that the next generation of students pursuing the evolving science of organization will be largely successful in supplying a more unified organizational economics approach to strategic management.

If the benefits of combining content and process are so transparently obvious, as I claim that they are in this concluding chapter, then one might reasonably ask, why didn't industrial organization economics take this path toward combining content and process a long time ago? My own view is most of the answer resides in history: Many of the early influential economists in academia had an expressed desire to build a deductive economic science that would be built from a few premises. Simon's (1957) comments are apposite:

The reluctance of economic theory to relinquish its classical model of economic man is understandable. When even a small concession has been made in the direction of admitting the fallibility of economic man, his psychological properties are no longer irrelevant. Deductive reasoning then no longer suffices for the unique prediction of his behavior without constant assistance from empirical observation. (p. 198)

Once the cognitive limitations of decision makers are admitted, then developing a useful science of organization may begin.

In the behavioral theory of the firm (in Chapter 1), I show that those researchers operating from the behavioral perspective tend to view the organization as a more efficient information processor than the individual. The firm, for example, is considered to be a functionally rational institutional response to uncertainty and bounded rationality (Cyert & March, 1963; March & Simon, 1958). Indeed, Thompson (1967) notes, "Uncertainty appears as the fundamental problem for complex organizations, and coping with uncertainty, (is) the essence of the administrative process" (p. 159). Organizations are structures of mutual expectation that reduce equivocality.

Transaction costs theory (in Chapter 2) builds a theory of the firm based on a core premise of the behavioral theory of the firm—bounded rationality. To make information flow considerations even more problematic, transaction costs theory holds that the human limitations for

processing information are aggravated further by the potential for deliberate nondisclosure of information, deliberate obfuscation, and the making of self-disbelieved statements—all of which are forms of opportunism. The firm is then viewed as a mechanism not only to take account of bounded rationality (in common with the behavioral theory of the firm) but also to attenuate opportunistic behavior. The firm may be preferred to the market due to the firm's adaptability, monitoring, dispute settling, and reward refining attributes.

A theory related to the transaction costs theory is property rights theory (in Chapter 3). Indeed, two of the more important historical figures in transaction costs theory have seminal works in property rights theory (Coase, 1960; Commons, 1924). In the property rights approach, the corporation is viewed as “a method of property tenure” (Berle & Means, 1932, p. 1).

The three criteria for efficiency of property rights are (1) universality—all scarce resources are owned by someone; (2) exclusivity—property rights are exclusive rights; and (3) transferability—to ensure that resources can be allocated from lower to higher yield uses. In neoclassical micro-economic theory, all three criteria are assumed to hold. However, in the real world of positive transaction costs, such as measurement costs (Barzel, 1989), none of these criteria is ensured.

Alchian and Demsetz (1972) argue that the structure of property rights has a significant impact on agency and transaction costs and that property rights establish an institutional context within which transactions are negotiated. Agency theory, transaction costs theory, and property rights theory, thus, are highly intertwined. Indeed, the significance of the study of property rights theory results from the fact that positive transaction costs are present (Coase, 1960).

There are several commonalities between transaction costs theory and agency theory (in Chapter 4)—especially with positive agency theory (Jensen, 1983; Jensen & Meckling, 1976). Both theories assume bounded rationality and opportunism (e.g., moral hazard). The consequences of bounded rationality and opportunism in both theories are incomplete contracting and the potential for contractual hazards. Agency theory emphasizes *ex ante* incentive alignment, whereas transaction costs theory emphasizes *ex post* governance issues (e.g., dispute resolution and maladaptation costs).

Dynamic resource-based theory (Chapter 5) is closely tied to all four previous theories of the firm. Resource-based theory is linked to the behavioral theory of the firm, if superior heuristics lead to sustainable

competitive advantage (Schoemaker, 1990). Moreover, the more evolutionary economics component of dynamic resource-based theory as developed by Nelson and Winter (1982) draws heavily from the Carnegie School behavioral theory of the firm (Cyert & March, 1963; March & Simon, 1958) with its emphasis on routines and search. Resource-based theory is linked to transaction costs theory because resource combinations are influenced by transaction costs considerations (Teece, 1982). Resource-based theory is linked to property rights theory because well-delineated property rights make resources economically valuable, and, as resources become more economically valuable, property rights become more precise (Libecap, 1989). Finally, resource-based theory is linked to agency theory because economic incentives highly influence the deployment of resources and the development of organizational capabilities (Castanias & Helfat, 1991).

This concluding chapter emphasizes the complementarities among these organizational economics theories of the firm and the potential for integration in the evolving science of organization (Langlois, Yu, & Robertson, 2002; Tan & Mahoney, 2003). I hasten to add, however, that each of these five theories has distinctive features. The behavioral theory focuses almost exclusively on the bounded rationality problem, whereas transaction costs theory focuses primarily on the attenuation of opportunism (albeit acknowledging the bounded rationality problem).

Property rights theory, like transaction costs theory, holds the premises of both bounded rationality and opportunism, but property rights theory is at a more aggregate unit of analysis. Property rights theory focuses more on the institutional level of analysis, whereas transaction costs theory focuses at a more microanalytic level of firm governance and the transaction. This difference in unit of analysis helps explain why property rights theory (e.g., Libecap, 1989; North, 1990) emphasizes inertia and the lack of adaptation, whereas transaction costs theory (Williamson, 1975, 1996) emphasizes the firm-level's capabilities to adapt over time to achieve relatively superior economic efficiencies.

Agency theory, in the form of the mathematical principal-agent model, which was highlighted in Chapter 4, assumes perfect rationality (i.e., no bounded rationality problems are assumed to exist). Clearly, such a conceptual lens may magnify and clarify economic incentives issues, but the theory will also inevitably blur some managerial problems and indeed will even neglect managerial problems of information-processing limitations.

Finally, I recently argued (Mahoney, 2001) that although resource-based theory holds the same premises as transaction costs theory—bounded rationality and opportunism—the theories are distinguished by the phenomena to be explained. Resource-based theory is a theory of firm-level economic rents, and transaction costs theory is a theory of the existence of the firm. I also make the conjecture that the set of market frictions to explain sustainable firm-level rents (in dynamic resource-based theory) will be sufficient market frictions to explain the existence of the firm (in transaction costs theory). Thus, these five theories of the firm are unique. The argument advanced here is that the five theories developed in this research book can be thought of as modular building blocks to develop new intellectual combinations of thought in the evolving science of organization.¹

Conclusions. The study of the economics of organization offers a wide range of challenges for students. In my judgment, at a general level, content (e.g., deductive economics) and process (e.g., behavioral and cognitive) approaches need to be joined in the next generation of organizational economics research. Moreover, the economics of organization requires far greater attention to the interactions between the institutional environment (e.g., North, 1990) and organizational-level governance (e.g., Williamson, 1996).

I provide the central messages from the economics of organization, followed by some key questions for the strategy management field. The core ideas of the organizational economics approach are as follows:

1. Behavioral theory of the firm—get the routines right to operate effectively
2. Transaction costs theory—get the governance right to minimize costs
3. Property rights theory—get the property rights well defined and correctly allocated to create wealth
4. Agency theory—get the incentives right to minimize the agency loss
5. Resource-based theory—choose the right resources (i.e., valuable rare, inimitable, nonsubstitutable) to generate and sustain (Ricardian) rents

¹I thank Anne Huff for suggesting the idea that the structure of this research book can be usefully described as having a modular architecture, where doctoral students may use various combinations of these theories of the firm depending on the problem at hand. In addition, Anne also points out that it is worth noting that each of the five theories has its own precise vocabulary and specific phenomena of interest and thus doctoral students may continue to contribute analytically by specializing in one of these theories in addition to the possibilities of synthesis, which I have emphasized throughout this book.

6. Dynamic capabilities theory—build technological and organizational capabilities for (Schumpeterian) rents
7. Real options theory—include real options values to analyze the strategic value of investments under uncertainty

I conclude with the following list of questions for further research using an organizational economics approach to strategic management:

- What are routines? What role do routines play in organizational survival and sustainable competitive advantage? What are the refutable implications?
- What are institutions? How do they differ? To what purpose and effect? Where does the action reside? What are the mechanisms? What are the refutable implications? What are the public policy ramifications? What do the data support?
- How can we describe and explain (and perhaps predict) how institutional (property rights) environments and governance (transaction costs and agency costs) play out?
- How does the interaction of managers and other human resources influence a firm's growth and competitive advantage?
- Why do firms differ in resources, capabilities, and economic performance?
- How can networks and strategic alliances be formed and organized to share or codevelop capabilities?
- What are the implications for firm growth of developing managers internally rather than hiring managers externally?
- To what extent can firms gain and sustain competitive advantage (via both competition and cooperation) by strategic commitment?
- How can firms gain and sustain competitive advantage by strategic flexibility? How can real options research help strategic management to place an economic value on these real options?
- How will the economic surplus generated by the firm be allocated among its members?
- If the many stakeholders who comprise the nexus of contracts are residual claimants (e.g., workers who have invested firm-specific human capital), then why are shareholders necessarily the ones affected the most by the firm's decisions? Even if they are, are these shareholders the party that benefits the most economically from the additional protection granted by the decision rights (Zingales, 2000)?

These are challenging questions for further inquiry. I anticipate that this generation of students studying the economics of organization will contribute greatly to the evolving science of organization. We can and will do better.



References

- Adner, R., & Helfat, C. E. (2003). Corporate effects and dynamic managerial capabilities. *Strategic Management Journal*, 24, 1011-1025.
- Afuah, A. (2001). Dynamic boundaries of the firm: Are firms better off being vertically integrated in the face of a technological change? *Academy of Management Journal*, 44(6), 1211-1228.
- Afuah, A. (2002). Mapping technological capabilities into product markets and competitive advantage: The case of cholesterol drugs. *Strategic Management Journal*, 23, 171-179.
- Ahuja, G., & Katila, R. (2001). Technological acquisitions and the innovation performance by acquiring firms: A longitudinal study. *Strategic Management Journal*, 22, 197-220.
- Akerlof, G. A. (1970). The market for "lemons": Qualitative uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84, 488-500.
- Alchian, A. A. (1950). Uncertainty, evolution, and economic theory. *Journal of Political Economy*, 58, 211-222.
- Alchian, A. A. (1965). Some economics of property rights. *Il Politico*, 30, 816-829.
- Alchian, A. A., & Demsetz, H. (1972). Production, information costs, and economic organization. *American Economic Review*, 62, 777-795.
- Alchian, A. A., & Demsetz, H. (1973). The property right paradigm. *Journal of Economic History*, 33, 16-27.
- Allison, G. T. (1971). *Essence of decision*. Boston: Little, Brown.
- Amit, R., & Shoemaker, P. J. H. (1993). Strategic assets and organizational rent. *Strategic Management Journal*, 14, 33-46.
- Anand, J., & Delios, A. (2002). Absolute and relative resources as determinants of international acquisitions. *Strategic Management Journal*, 23, 119-134.
- Anand, J., & Singh, H. (1997). Asset redeployment, acquisitions and corporate strategy in declining industries. *Strategic Management Journal*, 18, 99-118.
- Andersen, B. (2001). *Technological change and the evolution of corporate innovation: The structure of patenting 1890–1990*. Northampton, MA: Edward Elgar.
- Anderson, E. (1985). The salesperson as outside agent or employee: A transaction cost analysis. *Marketing Science*, 4(3), 234-254.
- Anderson, E., & Schmittlein, D. (1984). Integration of the sales force: An empirical examination. *Rand Journal of Economics*, 15, 385-395.
- Andrews, K. R. (1980). *The concept of corporate strategy*. Homewood, IL: Irwin.
- Ansoff, H. I. (1965). *Corporate strategy: An analytical approach to business policy for growth and expansion*. New York: McGraw-Hill.
- Argote, L. (1999). *Organizational learning: Creating, retaining and transferring knowledge*. Boston: Kluwer.
- Argyres, N. S. (1996). Evidence on the role of firm capabilities in vertical integration decisions. *Strategic Management Journal*, 17, 129-150.
- Argyres, N. S., & Liebeskind, J. P. (1999). Contractual commitments, bargaining power, and governance inseparability: Incorporating history into transaction cost theory. *Academy of Management Review*, 24(1), 49-63.

- Argyris, C., & Schon, D. (1978). *Organizational learning: A theory of action perspective*. Reading, MA: Addison-Wesley.
- Armour, H. O., & Teece, D. J. (1978). Organizational structure and economic performance: A test of the multidivisional hypothesis. *Bell Journal of Economics*, 9, 106-122.
- Arora, A., Fosfuri, A., & Gambardella, A. (2001). *Markets for technology: The economics of innovation and corporate strategy*. Cambridge, MA: MIT Press.
- Arrow, K. J. (1971). *Essays in the theory of risk-bearing*. New York: North-Holland.
- Arrow, K. J. (1974). *The limits of organization*. New York: W. W. Norton.
- Arrow, K. J. (1985). The economics of agency. In J. W. Pratt & R. J. Zeckhauser (Eds.), *Principals and agents: The structure of business* (pp. 37-51). Cambridge, MA: Harvard Business School Press.
- Arthur, W. B. (1994). *Increasing returns and path dependence in the economy*. Ann Arbor: University of Michigan Press.
- Bain, J. S. (1968). *Industrial organization*. New York: John Wiley & Sons.
- Balakrishnan, S., & Wernerfelt, B. (1986). Technical change, competition and vertical integration. *Strategic Management Journal*, 7, 347-359.
- Baldwin, C. Y., & Clark, K. B. (2000). *Design rules: The power of modularity*. Cambridge, MA: MIT Press.
- Barnard, C. I. (1938). *The functions of the executive*. Cambridge, MA: Harvard University Press.
- Barnard, C. I. (1948). *Organization and management: Selected papers*. Cambridge, MA: Harvard University Press.
- Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99-120.
- Barney, J. B., & Ouchi, W. G. (Eds.). (1986). *Organizational economics: A new paradigm for understanding and studying organizations*. San Francisco: Jossey-Bass.
- Baron, J. N., & Kreps, D. M. (1999). *Strategic human resources: Frameworks for general managers*. New York: John Wiley & Sons.
- Barzel, Y. (1982). Measurement cost and the organization of markets. *Journal of Law and Economics*, 25, 27-48.
- Barzel, Y. (1989). *An economic analysis of property rights*. Cambridge, UK: Cambridge University Press.
- Baumol, W. J., Panzar, J. C., & Willig, R. D. (1982). *Contestable markets and the theory of industrial structure*. New York: Harcourt Brace Jovanovich.
- Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. New York: Columbia University Press.
- Bensaou, B. M., & Anderson, E. (1999). Buyer-supplier relations in industrial markets: When do buyers risk making idiosyncratic investments. *Organization Science*, 10(4), 460-481.
- Berle, A. A., & Means, G. C. (1932). *The modern corporation and private property*. New York: Macmillan.
- Blair, R. D., & Kaserman, D. L. (1983). *Law and economics of vertical integration and control*. New York: Academic Press.
- Blaug, M. (1980). *The methodology of economics*. Cambridge, UK: Cambridge University Press.
- Bogner, W. C., Mahoney, J. T., & Thomas, H. (1998). Paradigm shift: The parallel origin, evolution, and function of strategic group analysis with the resource-based theory of the firm. *Advances in Strategic Management*, 15, 63-102.
- Boulding, K. E. (1950). *A reconstruction of economics*. New York: John Wiley & Sons.
- Bowman, E. H., & Helfat, C. E. (2001). Does corporate strategy matter? *Strategic Management Journal*, 22, 1-23.
- Bowman, E. H., & Hurry, D. (1993). Strategy through the options lens: An integrated view of resource investments and the incremental-choice process. *Academy of Management Review*, 18(4), 760-782.
- Bowman, E. H., & Kogut, B. M. (Eds.). (1995). *Redesigning the firm*. New York: Oxford University Press.

- Brush, T. H., & Artz, K. W. (1999). Toward a contingent resource-based theory: The impact of information asymmetry on the value of capabilities in veterinary medicine. *Strategic Management Journal*, 20, 223-250.
- Brynjolfsson, E. (1994). Information assets, technology, and organization. *Management Science*, 40(12), 1645-1662.
- Caldwell, B. (1984). *Appraisal and criticism in economics*. Boston: Allen and Unwin.
- Camerer, C. (1985). Redirecting research in business policy and strategy. *Strategic Management Journal*, 6, 1-15.
- Carney, M., & Gedajlovic, E. (1991). Vertical integration in franchise systems: Agency theory and resource explanations. *Strategic Management Journal*, 12, 607-629.
- Castanias, R. P., & Helfat, C. E. (1991). Managerial resources and rents. *Journal of Management*, 17, 155-171.
- Caves, R. E., & Bradburd, R. M. (1988). The empirical determinants of vertical integration. *Journal of Economic Behavior and Organization*, 9, 265-279.
- Caves, R. E., & Murphy, W. F. (1976). Franchising, firms, markets, and intangible assets. *Southern Economic Journal*, 42, 572-586.
- Chandler, A. D. (1962). *Strategy and structure*. Cambridge, MA: MIT Press.
- Chandler, A. D. (1977). *The visible hand: The managerial revolution in American business*. Cambridge, MA: Harvard University Press.
- Chandler, A. D. (1990). *Scale and scope: The dynamics of capitalism*. Cambridge, MA: Harvard University Press.
- Chandler, A. D. (1992). Organizational capabilities and the economic history of the industrial enterprise. *Journal of Economic Perspectives*, 6(3), 79-100.
- Chatterjee, S., & Wernerfelt, B. (1991). The link between resources and type of diversification: Theory and evidence. *Strategic Management Journal*, 12, 33-48.
- Cheung, S. S. (1969). *The theory of share tenancy*. Chicago: Chicago University Press.
- Cheung, S. S. (1973). The fable of the bees: An economic investigation. *Journal of Law and Economics*, 16, 11-34.
- Chi, T. (1994). Trading in strategic resources: Necessary conditions, transaction cost problems, and choice of exchange structure. *Strategic Management Journal*, 15, 271-290.
- Chi, T. (2000). Option to acquire or divest a joint venture. *Strategic Management Journal*, 21, 665-687.
- Clark, K. B., & Fujimoto, T. (1991). *Product development performance: Strategy, organization, and management in the world auto industry*. Cambridge, MA: Harvard Business School Press.
- Coase, R. H. (1937). The nature of the firm. *Economica*, 4, 386-405.
- Coase, R. H. (1960). The problem of social cost. *Journal of Law and Economics*, 3, 1-44.
- Coase, R. H. (1988). *The firm, the market and the law*. Chicago: University of Chicago Press.
- Coff, R. W. (2002). Human capital, shared expertise, and the likelihood of impasse in corporate acquisitions. *Journal of Management*, 28(1), 107-128.
- Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 35, 128-152.
- Coles, J. W., & Hesterly, W. S. (1998). The impact of firm-specific assets and the interaction of uncertainty: An examination of make or buy decisions on public and private hospitals. *Journal of Economic Behavior and Organization*, 36(3), 383-409.
- Collis, D. J. (1994). How valuable are organizational capabilities? *Strategic Management Journal*, 15, 143-152.
- Commons, J. R. (1924). *Legal foundations of capitalism*. New York: Macmillan.
- Commons, J. R. (1934). *Institutional economics: Its place in political economy*. New York: Macmillan.
- Conner, K. R. (1991). A historical comparison of resource-based theory and five schools of thought within industrial organization economics: Do we have a new theory of the firm? *Journal of Management*, 17(1), 121-154.
- Crocker, K. J., & Masten, S. E. (1988). Mitigating contractual hazards: Unilateral options and contract length. *Rand Journal of Economics*, 19, 327-343.

- Crocker, K. J., & Reynolds, K. J. (1993). The efficiency of incomplete contracts: An empirical analysis of air force engine procurement. *Rand Journal of Economics*, 24(1), 126-146.
- Cyert, R. M., & March, J. G. (1963). *A behavioral theory of the firm*. Englewood Cliffs, NJ: Prentice Hall.
- Davis, L. E., & North, D. C. (1971). *Institutional change and American economic growth*. Cambridge, UK: Cambridge University Press.
- Demsetz, H. (1967). Toward a theory of property rights. *American Economic Review*, 57, 347-359.
- Demsetz, H. (1973). Industry structure, market rivalry, and public policy. *Journal of Law and Economics*, 16, 1-9.
- Demsetz, H. (1988). *Ownership, control, and the firm*. Oxford, UK: Basil Blackwell.
- Demsetz, H. (1995). *The economics of the business firm*. New York: Cambridge University Press.
- Dierickx, I., & Cool, K. (1989). Asset stock accumulation and sustainability of competitive advantage. *Management Science*, 35(12), 1504-1511.
- Dixit, A., & Nalebuff, B. J. (1991). *Thinking strategically*. New York: W. W. Norton.
- Dixit, A., & Pindyck, R. S. (1994). *Investment under uncertainty*. Princeton, NJ: Princeton University Press.
- Dosi, G., Nelson, R. R., & Winter, S. G. (Eds.). (2000). *The nature and dynamics of organizational capabilities*. New York: Oxford University Press.
- Duhame, I. M., & Schwenk, C. R. (1985). Conjectures on cognitive simplification in acquisition and divestment decision making. *Academy of Management Review*, 10, 287-295.
- Dyer, J. H. (1996). Specialized supplier networks as a source of competitive advantage: Evidence from the auto industry. *Strategic Management Journal*, 17, 271-291.
- Dyer, J. H. (1997). Effective interfirm collaboration: How firms minimize transaction costs and maximize transaction value. *Strategic Management Journal*, 18(7), 535-556.
- Earl, P. E. (2001). *The legacy of Herbert Simon in economic analysis*. Cheltenham, UK: Edward Elgar.
- Eggertsson, T. (1990). *Economic behavior and institutions*. Cambridge, UK: Cambridge University Press.
- Eisenhardt, K. M. (1985). Control: Organizational and economic approaches. *Management Science*, 31(2), 134-149.
- Eisenhardt, K. M. (1988). Agency- and institutional-theory explanations: The case of retail sales compensation. *Academy of Management Journal*, 31, 488-511.
- Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *Academy of Management Review*, 14(1), 57-74.
- Evered, R., & Louis, M. R. (1981). Alternative perspectives in the organizational sciences: "Inquiry from the inside" and "inquiry from the outside." *Academy of Management Review*, 6, 385-395.
- Fama, E. F. (1980). Agency problems and the theory of the firm. *Journal of Political Economy*, 88(2), 288-307.
- Fama, E. F., & Jensen, M. C. (1983a). Separation of ownership and control. *Journal of Law and Economics*, 26, 301-325.
- Fama, E. F., & Jensen, M. C. (1983b). Agency problems and residual claims. *Journal of Law and Economics*, 26, 327-349.
- Farjoun, M. (1994). Beyond industry boundaries: Human expertise, diversification and resource-related industry groups. *Organization Science*, 5(2), 185-199.
- Farjoun, M. (1998). The independent and joint effects of the skill and physical bases of relatedness in diversification. *Strategic Management Journal*, 19, 611-630.
- Fiol, C. M., & Lyles, M. A. (1985). Organizational learning. *Academy of Management Review*, 10, 803-813.
- Folta, T. B. (1998). Governance and uncertainty: The trade-off between administrative control and commitment. *Strategic Management Journal*, 19, 1007-1028.
- Folta, T. B., & Miller, K. D. (2002). Real options in equity partnerships. *Strategic Management Journal*, 23, 77-88.
- Foss, N. J. (1996). Knowledge-based approaches to the theory of the firm: Some critical comments. *Organization Science*, 7(5), 470-476.

- Foss, N. J. (Ed.). (1997). *Resources, firms and strategies: A reader in the resource-based perspective*. New York: Oxford University Press.
- Friedman, M. (1953). *Essays in positive economics*. Chicago: University of Chicago Press.
- Fromm, E. (1941). *Escape from freedom*. New York: Rinehart.
- Furubotn, E., & Pejovich, S. (1972). Property rights and economic theory: A survey of recent literature. *Journal of Economic Literature*, 10, 1137-1162.
- Garud, R., & Kumaraswamy, A. (1995). Technological and organizational designs for realizing economies of substitution. *Strategic Management Journal*, 16, 93-109.
- Garud, R., Kumaraswamy, A., & Langlois, R. N. (Eds.). (2003). *Managing in the Modular age: Architectures, networks and organizations*. New York: Basil Blackwell.
- Ghemawat, P. (1991). *Commitment: The dynamic of strategy*. New York: Free Press.
- Ghemawat, P. (1997). *Games businesses play: Cases and models*. Cambridge, MA: MIT Press.
- Gimeno, J., Folta, T. B., Cooper, A. C., & Woo, C. Y. (1997). Survival of the fittest? Entrepreneurial human capital and the persistence of underperforming firms. *Administrative Science Quarterly*, 42, 750-783.
- Godfrey, P. C., & Hill, C. W. (1995). The problem of unobservables in strategic management research. *Strategic Management Journal*, 16, 519-533.
- Goldberg, V. P., & Erickson, J. R. (1987). Quantity and price adjustment in long-term contracts: A case study of petroleum coke. *Journal of Law and Economics*, 30, 369-398.
- Gort, M. (1962). *Diversification and integration in American industry*. Princeton, NJ: Princeton University Press.
- Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology*, 3, 481-510.
- Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, 17, 109-122.
- Grossman, S. J., & Hart, O. D. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy*, 94, 691-719.
- Harrigan, K. R. (1984). Formulating vertical integration strategies. *Academy of Management Review*, 9(4), 638-652.
- Hart, O. (1989). An economist's perspective on the theory of the firm. *Columbia Law Review*, 89(7), 1757-1774.
- Hart, O. (1995). *Firms, contracts, and financial structure*. Oxford, UK: Clarendon Press.
- Hart, O., & Moore, J. (1990). Property rights and the nature of the firm. *Journal of Political Economy*, 98, 1119-1158.
- Hayek, F. A. (1945). The uses of knowledge in society. *American Economic Review*, 35, 519-530.
- Hayek, F. A. (1948). *Individualism and economic order*. Chicago: Henry.
- Hayek, F. A. (1978). *New studies in philosophy, politics, economics and the history of ideas*. London, UK: Routledge & Kegan Paul.
- Heene, A., & Sanchez, R. (Eds.). (1997). *Competence-based strategic management*. New York: John Wiley & Sons.
- Heide, J. B., & Miner, M. S. (1992). The shadow of the future: Effects of anticipated interaction and frequency of contract on buyer-supplier cooperation. *Academy of Management Journal*, 35(2), 265-291.
- Helfat, C. E. (1997). Know-how and asset complementarity and dynamic capability accumulation: The case of R&D. *Strategic Management Journal*, 18(5), 339-360.
- Helfat, C. E., & Peteraf, M. A. (2003). The dynamic resource-based view: Capability lifecycles. *Strategic Management Journal*, 24, 997-1010.
- Henderson, R. M., & Clark, K. (1990). Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly*, 35, 9-30.
- Henderson, R. M., & Cockburn, I. (1994). Measuring competence? Exploring firm effects in pharmaceutical research. *Strategic Management Journal*, 15, 63-84.

- Hennart, J.-F. (1982). *A theory of the multinational enterprise*. Ann Arbor: University of Michigan Press.
- Hennart, J.-F. (1988a). A transaction costs theory of equity joint ventures. *Strategic Management Journal*, 9, 361-374.
- Hennart, J.-F. (1988b). Upstream vertical integration in the aluminum and tin industries: A comparative study of the choice between market and intrafirm coordination. *Journal of Economic Behavior and Organization*, 9, 281-299.
- Hennart, J.-F. (1993). Explaining the swollen middle: Why most transactions are a mix of "market" and "hierarchy." *Organization Science*, 4(4), 529-547.
- Hirschman, A. A. (1970). *Exit, voice, and loyalty: Responses to decline in firms, organizations, and states*. Cambridge, MA: Harvard University Press.
- Holmstrom, B. (1979). Moral hazard and observability. *Bell Journal of Economics*, 10, 74-91.
- Holmstrom, B., & Roberts, J. (1998). The boundaries of the firm revisited. *Journal of Economic Perspectives*, 12(4), 73-94.
- Hoopes, D. G., Madsen, T. L., & Walker, G. (2003). Why is there a resource-based view? Toward a theory of competitive heterogeneity. *Strategic Management Journal*, 24, 889-902.
- Huff, A. S. (1981). Multilectic methods of inquiry. *Human Systems Management*, 2, 83-94.
- Huff, A. S. (2000). Changes in organizational knowledge production. *Academy of Management Review*, 25(2), 288-293.
- Itami, H., & Roehl, T. L. (1987). *Mobilizing invisible assets*. Cambridge, MA: Harvard University Press.
- Jaffe, J., & Mahoney, J. M. (1999). The performance of investment newsletters. *Journal of Financial Economics*, 53(2), 289-307.
- Jensen, M. (1983). Organization theory and methodology. *Accounting Review*, 50, 319-339.
- Jensen, M., & Meckling, W. (1976). Theory of the firm: Managerial behavior, agency costs, and capital structure. *Journal of Financial Economics*, 3, 305-360.
- John, G., & Weitz, B. A. (1988). Forward integration into distribution: An empirical test of transaction cost analysis. *Journal of Law, Economics and Organization*, 4(2), 337-355.
- Joskow, P. L. (1985). Vertical integration and long-term contracts: The case of coal-burning electric generating plants. *Journal of Law, Economics, and Organization*, 1(1), 33-80.
- Joskow, P. L. (1987). Contract duration and relationship-specific investments: Empirical evidence from coal markets. *American Economic Review*, 77(1), 1-25.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.
- Kaplan, A. (1964). *The conduct of inquiry: Methodology for behavioral science*. San Francisco: Chandler Publishing.
- Kim, J., & Mahoney, J. T. (2002). Resource-based and property rights perspectives on value creation: The case of oil field unitization. *Managerial and Decision Economics*, 23(4), 225-245.
- Klein, B., Crawford, R. A., & Alchian, A. A. (1978). Vertical integration, appropriable rents, and the competitive contracting process. *Journal of Law and Economics*, 21, 297-326.
- Klein, B., & Leffler, K. B. (1981). The role of market forces in assuring contractual performance. *Journal of Political Economy*, 89(4), 615-641.
- Knight, F. H. (1921). *Risk, uncertainty and profit*. Boston: Houghton Mifflin.
- Knott, A. M. (2003a). The organizational routines factor market paradox. *Strategic Management Journal*, 24: 929-943.
- Knott, A. M. (2003b). Persistent heterogeneity and sustainable innovation. *Strategic Management Journal*, (24), 687-705.
- Kogut, B. (1991). Joint ventures and the option to expand and acquire. *Management Science*, 37(1), 19-33.
- Koja, B., & Prescott, J. E. (2002). Strategic alliances as social capital: A multidimensional view. *Strategic Management Journal*, 23, 795-816.
- Kor, Y. Y., & Mahoney, J. T. (2000). Penrose's resource-based approach: The process and product of research activity. *Journal of Management Studies*, 37(1), 109-139.

- Kor, Y. Y., & Mahoney, J. T. (2004). Edith Penrose's (1959) contributions to the resource-based view of strategic management. *Journal of Management Studies*, 41(1), 183-191.
- Kosnik, R. D. (1987). Greenmail: A study of board performance in corporate governance. *Administrative Science Quarterly*, 32, 163-185.
- Kuhn, T. S. (1970). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Ladd, G. W. (1987). *Imagination in research*. Ames: Iowa State Press.
- Lafontaine, F., & Shaw, K. L. (1999). The dynamics of franchise contracting: Evidence from panel data. *Journal of Political Economy*, 107(5), 1041-1080.
- Lajili, K., Barry, P. J., Sonka, S. T., & Mahoney, J. T. (1997). Farmer's preferences for crop contracts. *Journal of Agricultural and Resource Economics*, 22(2), 264-280.
- Langlois, R. N. (1986). *Economics as a process: Essays in the new institutional economics*. Cambridge, UK: Cambridge University Press.
- Langlois, R. N. (2002). Modularity in technology and organization. *Journal of Economic Organization*, 49(1), 19-37.
- Langlois, R. N., & Robertson, P. L. (1995). *Firms, markets and economic change*. London: Routledge.
- Langlois, R. N., Yu, T. F., & Robertson, P. L. (Eds.). (2002). *Alternative theories of the firm*. Northampton, MA: Edward Elgar.
- Leiblein, M. J., & Miller, D. J. (2003). An empirical examination of transaction- and firm-level influences on the vertical boundaries of the firm. *Strategic Management Journal*, 24, 839-859.
- Leiblein, M. J., Reuer, J. J., & Dalsace, F. (2002). Do make or buy decisions matter? The influence of organizational governance on technological performance. *Strategic Management Journal*, 23, 817-833.
- Leonard-Barton, D. (1992). Core capabilities and core rigidities: A paradox in managing new product development. *Strategic Management Journal*, 13, 111-125.
- Leonard-Barton, D. (1995). *Wellsprings of knowledge: Building and sustaining the sources of innovation*. Cambridge, MA: Harvard Business School Press.
- Levinthal, D. (1988). A survey of agency models of organizations. *Journal of Economic Behavior and Organization*, 9, 153-185.
- Levy, D. T. (1985). The transactions cost approach to vertical integration: An empirical examination. *Review of Economics and Statistics*, 67, 438-445.
- Libecap, G. D. (1989). *Contracting for property rights*. Cambridge, UK: Cambridge University Press.
- Libecap, G. D., & Wiggins, S. N. (1985). The influence of private contractual failure on regulation: The case of oil field unitization. *Journal of Political Economy*, 93, 690-714.
- Lieberman, M. B. (1987). The learning curve, diffusion, and competitive strategy. *Strategic Management Journal*, 8, 441-452.
- Lieberman, M. B. (1991). Determinants of vertical integration: An empirical test. *Journal of Industrial Economics*, 39(5), 451-466.
- Lieberman, M. B., & Montgomery, D. B. (1988). First-mover advantages. *Strategic Management Journal*, 9, 41-58.
- Lieberman, M. B., & Montgomery, D. B. (1998). First-mover (dis)advantages: Retrospective and link with the resource-based view. *Strategic Management Journal*, 19, 1111-1125.
- Liebeskind, J. P. (1996). Knowledge, strategy, and the theory of the firm. *Strategic Management Journal*, 17, 93-107.
- Lippman, S. A., & Rumelt, R. P. (1982). Uncertain imitability: An analysis of interfirm differences in efficiency under competition. *Bell Journal of Economics*, 13, 418-438.
- Lippman, S. A., & Rumelt, R. P. (2003). The payments perspective: Micro-foundations of resource analysis. *Strategic Management Journal*, 24, 903-927.
- Lyons, B. R. (1994). Contracts and specific investment: An empirical test of transaction costs theory. *Journal of Economics and Management Strategy*, 3(2), 257-278.
- Machlup, F. (1967). *Essays in economic semantics*. New York: W. W. Norton.
- MacKinlay, A. C. (1997). Event studies in economics and finance. *Journal of Economic Literature*, 35, 13-39.

- Macneil, I. R. (1980). *The new social contract: An inquiry into modern contractual relations*. New Haven, CT: Yale University Press.
- Madhok, A. (2002). Reassessing the fundamentals and beyond: Ronald Coase, the transaction cost and resource-based theories of the firm and the institutional structure of production. *Strategic Management Journal*, 23, 535-550.
- Mahoney, J. M., & Mahoney, J. T. (1993). An empirical investigation of the effect of corporate charter antitakeover amendments on stockholder wealth. *Strategic Management Journal*, 14, 17-31.
- Mahoney, J. M., Sundaramurthy, C., & Mahoney, J. T. (1996). The differential impact on stockholder wealth of various antitakeover provisions. *Managerial and Decision Economics*, 17, 531-549.
- Mahoney, J. M., Sundaramurthy, C., & Mahoney, J. T. (1997). The effects of corporate antitakeover provisions on long-term investment: Empirical evidence. *Managerial and Decision Economics*, 18, 349-365.
- Mahoney, J. T. (1992a). The adoption of the multidivisional form of organization: A contingency model. *Journal of Management Studies*, 29(1), 49-72.
- Mahoney, J. T. (1992b). The choice of organizational form: Vertical financial ownership versus other methods of vertical integration. *Strategic Management Journal*, 13, 559-584.
- Mahoney, J. T. (1992c). Organizational economics within the conversation of strategic management. *Advances in Strategic Management*, 8, 103-155.
- Mahoney, J. T. (1993). Strategic management and determinism: Sustaining the conversation. *Journal of Management Studies*, 30(1), 173-191.
- Mahoney, J. T. (1995). The management of resources and the resource of management. *Journal of Business Research*, 33, 91-101.
- Mahoney, J. T. (2001). A resource-based theory of sustainable rents. *Journal of Management*, 27, 651-660.
- Mahoney, J. T. (2002). The relevance of Chester I. Barnard's teachings to contemporary management education: Communicating the aesthetics of management. *International Journal of Organization Theory and Behavior*, 5(1), 159-172.
- Mahoney, J. T., & Crank, D. A. (1995). Vertical coordination: The choice of organizational form. In K. Coaldrae, S. T. Sonka, D. Sudharshan, & F. W. Winter (Eds.), *New industries and strategic alliances in agriculture: Concepts and cases* (pp. 89-126). Champaign, IL: Stipes.
- Mahoney, J. T., Huff, A. S., & Huff, J. O. (1994a). Management's search for balance. *Journal of Management Inquiry*, 3(2), 173-174.
- Mahoney, J. T., Huff, A. S., & Huff, J. O. (1994b). Toward a new social contract theory in organization science. *Journal of Management Inquiry*, 3(2), 153-168.
- Mahoney, J. T., & Pandian, J. R. (1992). The resource-based view within the conversation of strategic management. *Strategic Management Journal*, 13, 363-380.
- Mahoney, J. T., & Sanchez, R. (1997). Competence theory building: Reconnecting management research and management practice. In A. Heene & R. Sanchez (Eds.), *Competence-based strategic management* (pp. 45-64). Chichester, UK: John Wiley & Sons.
- Mahoney, J. T., & Sanchez, R. (2004). Building new management theory by integrating process and product of thought. *Journal of Management Inquiry*, 13, 34-47.
- Makadok, R. (2001). Toward a synthesis of the resource-based and dynamic-capability views of rent creation. *Strategic Management Journal*, 22, 387-401.
- Makadok, R. (2003). Doing the right thing and knowing the right thing to do: Why the whole is greater than the sum of the parts. *Strategic Management Journal*, 24, 1043-1055.
- Malmgren, H. B. (1961). Information, expectations and the theory of the firm. *Quarterly Journal of Economics*, 75, 399-421.
- March, J. G. (1988). *Decisions and organizations*. Oxford, UK: Basil Blackwell.
- March, J. G. (1999). *The pursuit of organizational intelligence*. Oxford, UK: Basil Blackwell.
- March, J. G., & Simon, H. A. (1958). *Organizations*. New York: John Wiley & Sons.

- Marris, R. (1964). *The economic theory of "managerial" capitalism*. New York: Macmillan.
- Marschak, J. (1968). Economics of inquiring, communicating, deciding. *American Economic Review*, 58, 1-18.
- Marshall, A. (1920). *Principles of economics*. London: Macmillan.
- Masten, S. E. (1984). The organization of production: Evidence from the aerospace industry. *Journal of Law and Economics*, 27, 403-418.
- Masten, S. E. (1988). A legal basis for the firm. *Journal of Law, Economics and Organization*, 4, 181-198.
- Masten, S. E., & Crocker, K. J. (1985). Efficient adaptation in long-term contracts: Take-or-pay provision for natural gas. *American Economic Review*, 75, 1083-1093.
- Masten, S. E., Meehan, J. W., & Snyder, E. A. (1989). Vertical integration in the U.S. auto industry: A note on the influence of transaction specific assets. *Journal of Economic Behavior and Organization*, 12, 265-273.
- Masten, S. E., Meehan, J. W., & Snyder, E. A. (1991). The costs of organization. *Journal of Law, Economics and Organization*, 7(1), 1-25.
- McCloskey, D. (1983). The rhetoric of economics. *Journal of Economic Literature*, 21, 481-517.
- McCloskey, D. (1998). *The rhetoric of economics*. Madison: University of Wisconsin Press.
- McCloskey, D., & Ziliak, S. T. (1996). The standard error of regressions. *Journal of Economic Literature*, 34, 97-114.
- McGrath, R. G. (1997). A real options logic for initiating technology positioning investments. *Academy of Management Review*, 22(4), 974-996.
- McGrath, R. G. (1999). Falling forward: Real options reasoning and entrepreneurial failure. *Academy of Management Review*, 24(1), 13-30.
- McMillan, J. (1992). *Games, strategies, and managers*. New York: Oxford University Press.
- Michael, S. C. (2000). Investments to create bargaining power: The case of franchising. *Strategic Management Journal*, 21, 497-514.
- Milgrom, P., & Roberts, J. (1992). *Economics, organization and management*. Englewood Cliffs, NJ: Prentice Hall.
- Miller, D., & Shamsie, J. (1996). The resource-based view of the firm in two environments: The Hollywood film studios from 1936 to 1965. *Academy of Management Journal*, 39(3), 519-534.
- Miller, G. J. (1992). *Managerial dilemmas: The political economy of hierarchy*. New York: Cambridge University Press.
- Miller, K. D. (2002). Knowledge inventories and managerial myopia. *Strategic Management Journal*, 23, 689-706.
- Miller, K. D., & Folta, T. B. (2002). Option value and entry timing. *Strategic Management Journal*, 23, 655-665.
- Monteverde, K. (1995). Technical dialog as an incentive for vertical integration in the semiconductor industry. *Management Science*, 41(10), 1624-1638.
- Monteverde, K., & Teece, D. (1982). Supplier switching costs and vertical integration in the automobile industry. *Bell Journal of Economics*, 13, 206-213.
- Montgomery, C. A., & Hariharaan, S. (1991). Diversified expansion by large established firms. *Journal of Economic Behavior and Organization*, 15, 71-89.
- Montgomery, C. A., & Wernerfelt, B. (1988). Diversification, Ricardian rents, and Tobin's q. *Rand Journal of Economics*, 19(4), 623-632.
- Montgomery, C. A., Wernerfelt, B., & Balakrishnan, S. (1989). Strategy content and the research process. *Strategic Management Journal*, 10, 189-197.
- Mosakowski, E. (1993). A resource-based perspective on the dynamic strategy-performance relationship: An empirical examination of the focus and differentiation strategies in entrepreneurial firms. *Journal of Management*, 19(4), 819-839.
- Mowery, D. C., Oxley, J. E., & Silverman, B. S. (1996). Strategic alliances and interfirm knowledge transfer. *Strategic Management Journal*, 17, 77-91.
- Nelson, R. R. (1996). *The sources of economic growth*. Cambridge, MA: Harvard University Press.

- Nelson, R. R., & Winter, S. G. (1982). *An evolutionary theory of economic change*. Cambridge, MA: Harvard University Press.
- Nelson, R. R., & Winter, S. G. (2002). Evolutionary theorizing in economics. *Journal of Economic Perspectives*, 16(2), 23-46.
- Nichols, N. A. (1994, January/February). Scientific management at Merck: An interview with CFO Judy Lewent. *Harvard Business Review*, 89-99.
- Nickerson, J. A., Hamilton, B. H., & Wada, T. (2001). Market position, resource profile, and governance: Linking Porter and Williamson in the context of international courier and small package services in Japan. *Strategic Management Journal*, 22, 251-273.
- Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company*. Oxford, UK: Oxford University Press.
- North, D. (1981). *Structure and change in economic history*. New York: W. W. Norton.
- North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge, UK: Cambridge University Press.
- North, D. C., & Thomas, R. P. (1973). *The rise of the Western world: A new economic history*. Cambridge, UK: Cambridge University Press.
- Norton, S. W. (1988). Franchising, brand name capital, and the entrepreneurial capacity problem. *Strategic Management Journal*, 9, 105-114.
- Novak, S., & Eppinger, S. D. (2001). Sourcing by design: Product complexity and the supply chain. *Management Science*, 47(1), 189-204.
- O'Driscoll, G. P., & Rizzo, M. J. (1985). *The economics of time and ignorance*. New York: Basil Blackwell.
- Oliver, C. (1997). Sustainable competitive advantage: Combining institutional and resource-based views. *Strategic Management Journal*, 18, 697-713.
- Ouchi, W. G. (1979). A conceptual framework for the design of organizational control mechanisms. *Management Science*, 25(9), 833-848.
- Oviatt, B. M. (1988). Agency and transaction cost perspectives on the manager-shareholder relationship: Incentives for congruent interests. *Academy of Management Review*, 13, 214-225.
- Palay, T. M. (1984). Comparative institutional economics: The governance of rail freight contracting. *Journal of Legal Studies*, 13, 265-287.
- Palay, T. M. (1985). Avoiding regulatory constraints: The use of informal contracts. *Journal of Law, Economics, and Organization*, 1, 155-175.
- Palmer, D., Friedland, R., Jennings, P. D., & Powers, M. E. (1987). The economics and politics of structure: The multidivisional form and the large U.S. corporation. *Administrative Science Quarterly*, 32, 25-48.
- Peng, M. W. (2001). The resource-based view and international business. *Journal of Management*, 27, 803-829.
- Penrose, E. T. (1955). Limits to the growth of the firm. *American Economic Review*, 45(2), 531-543.
- Penrose, E. T. (1959). *The theory of the growth of the firm*. New York: John Wiley & Sons.
- Penrose, E. T. (1960). The growth of the firm: A case study: The Hercules Powder Company. *Business History Review*, 34, 1-23.
- Peteraf, M. (1993). The cornerstones of competitive advantage: A resource-based view. *Strategic Management Journal*, 14, 179-191.
- Pettus, M. (2001). The resource-based view as a developmental growth process: Evidence from the deregulated trucking industry. *Academy of Management Journal*, 44(4), 878-896.
- Phillips, A., & Mahoney, J. (1985). Unreasonable rules and rules of reason: Economic aspects of vertical price-fixing. *Antitrust Bulletin*, 30(1), 99-115.
- Pirrong, S. C. (1993). Contracting practices in bulk shipping markets: A transaction costs explanation. *Journal of Law and Economics*, 36, 937-976.
- Pisano, G. P. (1989). Using equity participation to support exchange: Evidence from the biotechnology industry. *Journal of Law, Economics and Organization*, 5(1), 109-126.
- Pitelis, C. (Ed.) (2002). *The growth of the firm: The legacy of Edith Penrose*. New York: Oxford University Press.

- Polanyi, M. (1962). *Personal knowledge*. Chicago: University of Chicago Press.
- Popper, K. R. (1970). Normal science and its dangers. In L. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 51-58). Cambridge, UK: Cambridge University Press.
- Poppo, L., & Zenger, T. (1998). Testing alternative theories of the firm: Transaction cost, knowledge-based, and measurement explanations for make-or-buy-decisions in information services. *Strategic Management Journal*, 19, 853-877.
- Poppo, L., & Zenger, T. (2002). Do formal contracts and relational governance function as substitutes or complements? *Strategic Management Journal*, 23, 707-725.
- Porter, M. E. (1980). *Competitive strategy: Techniques for analyzing industries and competitors*. New York: Free Press.
- Porter, M. E. (1985). *Competitive advantage: Creating and sustaining superior performance*. New York: Free Press.
- Porter, M. E. (1991). Towards a dynamic theory of strategy. *Strategic Management Journal*, 12, 95-117.
- Powell, T. C., & Dent-Micallef, A. (1997). Information technology as competitive advantage: The role of human, business, and technology resources. *Strategic Management Journal*, 18(5), 375-405.
- Pratt, J. W., & Zeckhauser, R. J. (1985). Principals and agents: An overview. In J. W. Pratt & R. J. Zeckhauser (Eds.), *Principals and agents: The structure of business* (pp. 1-35). Cambridge, MA: Harvard Business School Press.
- Rasmusen, E. (1989). *Games and information*. Cambridge, UK: Basil Blackwell.
- Rediker, K. J., & Seth, A. (1995). Boards of directors and substitution effects of alternative governance mechanisms. *Strategic Management Journal*, 16(2), 85-99.
- Redman, D. A. (1993). *Economics and the philosophy of science*. New York: Oxford University Press.
- Reed, R., & DeFillippi, R. J. (1990). Causal ambiguity, barriers to imitation and sustainable competitive advantage. *Academy of Management Review*, 15(1), 88-102.
- Ricardo, D. (1817). *Principles of political economy and taxation*. London: J. Murray.
- Richardson, G. B. (1960). *Information and investment*. Oxford, UK: Oxford University Press.
- Richardson, G. B. (1972). The organization of industry. *Economic Journal*, 82, 883-896.
- Richardson, J. (1993). Parallel sourcing and supplier performance in the Japanese automobile industry. *Strategic Management Journal*, 14, 339-350.
- Riordan, M., & Williamson, O. E. (1985). Asset specificity and economic organization. *International Journal of Industrial Organization*, 3, 365-378.
- Robertson, T. S., & Gatignon, H. (1998). Technology development mode: A transaction cost conceptualization. *Strategic Management Journal*, 19, 515-531.
- Rubin, P. H. (1973). The expansion of firms. *Journal of Political Economy*, 81, 936-949.
- Rubin, P. H. (1978). The theory of the firm and the structure of the franchise contract. *Journal of Law and Economics*, 21, 223-233.
- Rugman, A. M., & Verbeke, A. (2002). Edith Penrose's contribution to the resource-based view of strategic management. *Strategic Management Journal*, 23, 769-780.
- Rumelt, R. P. (1974). *Strategy, structure, and economic performance*. Cambridge, MA: Harvard Business School Press.
- Rumelt, R. P. (1984). Toward a strategic theory of the firm. In R. Lamb (Ed.), *Competitive strategic management* (pp. 556-570). Englewood Cliffs, NJ: Prentice Hall.
- Rumelt, R. P., Schendel, D. E., & Teece, D. J. (Eds.). (1994). *Fundamental issues in strategy*. Cambridge, MA: Harvard University Press.
- Saloner, G. (1991). Modeling, game theory, and strategic management. *Strategic Management Journal*, 12, 119-136.
- Sanchez, R. (1993). Strategic flexibility, firm organization, and managerial work in dynamic markets: A strategic options perspective. *Advances in Strategic Management*, 9, 251-291.
- Sanchez, R. (2003). Integrating transaction costs theory and real options theory. *Managerial and Decision Economics*, 24, 267-282.

- Sanchez, R., & Mahoney, J. T. (1996). Modularity, flexibility, and knowledge management in product and organization design. *Strategic Management Journal*, 17, 63-76.
- Sanchez, R., & Mahoney, J. T. (2001). Modularity and dynamic capabilities. In H. Volberda & T. Elfring, (Eds.), *Rethinking strategy* (pp. 158-171). Thousand Oaks, CA: Sage.
- Schelling, T. C. (1960). *The strategy of conflict*. Cambridge, MA: Harvard University Press.
- Scherer, F. M., & Ross, D. (1990). *Industrial market structure and economic performance*. Boston: Houghton Mifflin.
- Schilling, M. A. (2000). Toward a general modular systems theory and its application to interfirm product modularity. *Academy of Management Review*, 25(2), 312-334.
- Schoemaker, P. J. H. (1990). Strategy, complexity and economic rent. *Management Science*, 36, 1178-1192.
- Schumpeter, J. A. (1934). *The theory of economic development*. Cambridge, MA: Harvard University Press.
- Schumpeter, J. A. (1950). *Capitalism, socialism, and democracy*. New York: Harper & Brothers.
- Schwartz, E. S., & Trigeorgis, L. (Eds.). (2001). *Real options and investment under uncertainty: Classical readings and recent contributions*. Cambridge, MA: MIT Press.
- Scott, W. R. (1987). *Organizations: Rational, natural, and open systems*. Englewood Cliffs, NJ: Prentice Hall.
- Scott, W. R. (1995). *Institutions and organizations*. Thousand Oaks, CA: Sage.
- Selznick, P. (1957). *Leadership in administration*. Berkeley: University of California Press.
- Seth, A. (1990). Sources of value creation in acquisitions: An empirical investigation. *Strategic Management Journal*, 11, 431-446.
- Seth, A., & Thomas, H. (1994). Theories of the firm: Implications for strategy research. *Journal of Management Studies*, 31(2), 165-191.
- Seth, A., & Zinkhan, G. (1991). Strategy and the research process: A comment. *Strategic Management Journal*, 12, 75-82.
- Shane, S. (1996). Hybrid organizational arrangements and their implications for firm growth and survival: A study of new franchisers. *Academy of Management Journal*, 39(1), 216-234.
- Shapiro, C., & Varian, H. R. (1999). *Information rules: A strategic guide to the network economy*. Cambridge, MA: Harvard Business School Press.
- Silverman, B. S. (1999). Technological resources and the direction of corporate diversification: Toward an integration of the resource-based view and transaction cost economics. *Management Science*, 45, 1109-1124.
- Simon, H. A. (1947). *Administrative behavior*. New York: Macmillan.
- Simon, H. A. (1957). *Models of man: Social and rational*. New York: John Wiley & Sons.
- Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society*, 106, 467-482.
- Simon, H. A. (1982). *Models of bounded rationality: Behavioral economics and business organization*. Cambridge, MA: MIT Press.
- Simon, H. A. (1996). *The sciences of the artificial*. Cambridge, MA: MIT Press.
- Slater, M. (1980). The managerial limitations to the growth of firm. *Economic Journal*, 90, 520-528.
- Smith, A. (1937). *The wealth of nations*. New York: Methuen. (Original work published 1776)
- Sorenson, O., & Sorensen, J. B. (2001). Finding the right mix: Franchising, organizational learning, and chain performance. *Strategic Management Journal*, 22, 713-724.
- Spender, J.-C. (1996). Making knowledge the basis of a dynamic theory of the firm. *Strategic Management Journal*, 17, 45-62.
- Stigler, G. J. (1968). *The organization of industry*. Homewood, IL: Irwin.
- Stuckey, J. (1983). *Vertical integration and joint ventures in the aluminum industry*. Cambridge, MA: Harvard University Press.
- Subramani, M. R., & Venkatraman, N. (2003). Safeguarding investments in asymmetric interorganizational relationships: Theory and evidence. *Academy of Management Journal*, 46(1), 46-62.

- Sundaramurthy, C., & Lewis, M. (2003). Control and collaboration: Paradoxes of governance. *Academy of Management Review*, 28(3), 397-415.
- Sundaramurthy, C., Mahoney, J. M., & Mahoney, J. T. (1997). The effects of corporate antitakeover provisions on long-term investment: Empirical evidence. *Strategic Management Journal*, 18(3), 231-245.
- Sutton, J. (1992). *Sunk costs and market structure: Price competition, advertising, and the evolution of concentration*. Cambridge, MA: MIT Press.
- Szulanski, G. (1996). Exploring internal stickiness: Impediments to the transfer of best practice within the firm. *Strategic Management Journal*, 17, 27-43.
- Szulanski, G. (2003). *Sticky knowledge: Barriers to knowing in the firm*. Thousand Oaks, CA: Sage.
- Tan, D. (2003). The limits to the growth of multinational firms in a foreign market. *Managerial and Decision Economics*, 24, 569-582.
- Tan, D., & Mahoney, J. T. (2003). Explaining the utilization of managerial expatriates from the perspectives of resource-based, agency, and transaction costs theories. *Advances in International Management*, 15, 179-205.
- Tan, D., & Mahoney, J. T. (in press). Examining the Penrose effect in an international business context: The dynamics of Japanese growth in U.S. industries. *Managerial and Decision Economics*.
- Teece, D. J. (1980). Economies of scope and the scope of the enterprise. *Journal of Economic Behavior and Organization*, 1, 223-245.
- Teece, D. J. (1982). Towards an economic theory of the multiproduct firm. *Journal of Economic Behavior and Organization*, 3, 39-63.
- Teece, D. J. (2000). *Managing intellectual capital: Organizational, strategic and policy dimensions*. New York: Oxford University Press.
- Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, 18(7), 509-533.
- Telser, L. G. (1980). A theory of self-enforcing agreements. *Journal of Business*, 53, 27-44.
- Thomke, S., & Kuemmerle, W. (2002). Asset accumulation, interdependence and technological change: Evidence from pharmaceutical drug discovery. *Strategic Management Journal*, 23, 619-635.
- Thompson, J. D. (1967). *Organizations in action*. New York: McGraw-Hill.
- Tirole, J. (1988). *The theory of industrial organization*. Cambridge, MA: MIT Press.
- Trigeorgis, L. (1996). *Real options: Managerial flexibility and strategy in resource allocation*. Cambridge, MA: MIT Press.
- Tripsas, M. (1997). Unraveling the process of creative destruction: Complementary assets and incumbent survival in the typesetter industry. *Strategic Management Journal*, 18, 119-142.
- Volberda, H. W., & Elfring, T. (Eds.). (2001). *Rethinking strategy*. Thousand Oaks, CA: Sage.
- Walker, G., & Poppo, L. (1991). Profit centers, single-source suppliers, and transaction costs. *Administrative Science Quarterly*, 36, 66-87.
- Walker, G., & Weber, D. (1984). A transaction cost approach to make-or-buy decisions. *Administrative Science Quarterly*, 29, 373-391.
- Walker, G., & Weber, D. (1987). Supplier competition, uncertainty and make-or-buy decisions. *Academy of Management Journal*, 30, 589-596.
- Walsh, J. P., & Seward, J. K. (1990). On the efficiency of internal and external corporate control mechanisms. *Academy of Management Review*, 15(3), 421-458.
- Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal*, 5, 171-180.
- Wernerfelt, B. (1995). The resource-based view of the firm: Ten years after. *Strategic Management Journal*, 16, 171-174.
- Whetten, D. A. (1989). What constitutes a theoretical contribution? *Academy of Management Review*, 14(4), 490-495.
- Williamson, O. E. (1971). The vertical integration of production: Market failure considerations. *American Economic Review*, 61, 112-123.

- Williamson, O. E. (1975). *Markets and hierarchies: Analysis and antitrust implications*. New York: Free Press.
- Williamson, O. E. (1979). Transaction cost economics: The governance of contractual relations. *Journal of Law and Economics*, 22, 233-261.
- Williamson, O. E. (1985). *The economic institutions of capitalism: Firms, markets, relational contracting*. New York: Free Press.
- Williamson, O. E. (1991). Strategizing, economizing, and economic organization. *Strategic Management Journal*, 12, 75-94.
- Williamson, O. E. (Ed.). (1995). *Organization theory: From Chester Barnard to the present and beyond*. Oxford, UK: Oxford University Press.
- Williamson, O. E. (1996). *The mechanisms of governance*. New York: Oxford University Press.
- Williamson, O. E. (1999). Strategy research: Governance and competence perspectives. *Strategic Management Journal*, 20, 1087-1108.
- Williamson, O. E. (2000). The new institutional economics: Taking stock and looking ahead. *Journal of Economic Literature*, 38, 596-613.
- Williamson, O. E., & Masten, S. E. (Eds.). (1999). *The economics of transaction costs*. Northampton, MA: Elgar.
- Williamson, O. E., & Winter, S. G. (Eds.). (1991). *The nature of the firm: Origin, evolution, development*. New York: Oxford University Press.
- Wolfson, M. A. (1985). Empirical evidence of incentive problems and their mitigation in oil and gas shelter programs. In J. W. Pratt & R. J. Zeckhauser (Eds.), *Principals and agents* (pp. 101-125). Cambridge, MA: Harvard Business School Press.
- Worren, N., Moore, K., & Cardona, P. (2002). Modularity, strategic flexibility, and firm performance: A study of the home appliance industry. *Strategic Management Journal*, 23, 1123-1140.
- Yao, D. A. (1988). Beyond the reach of the invisible hand: Impediments to economic activity, market failures and profitability. *Strategic Management Journal*, 9, 59-70.
- Yeoh, P.-L., & Roth, K. (1999). An empirical analysis of sustained advantage in the U.S. pharmaceutical industry: Impact of firm resources and capabilities. *Strategic Management Journal*, 20, 637-653.
- Zander, U., & Kogut, B. (1995). Knowledge and the speed of the transfer and imitation of organizational capabilities: An empirical test. *Organization Science*, 6(1), 76-92.
- Zenger, T. R. (1994). Explaining organizational diseconomies of scale in R&D: Agency problems and the allocation of engineering talent, ideas, and effort by firm size. *Management Science*, 40(6), 708-729.
- Zingales, L. (2000). In search of new foundations. *Journal of Finance*, 55, 1623-1653.



Index

- Active decision area, 65
Adaptive behavior
autonomous/Hayekian, 101–102
ceteris paribus and, 32
of Competitive market systems, 73
organizational, 38–40
strategizing, 104
Administrative Behavior, 2–3
Administrative group and theory of the firm, 171–172
Administrative organization. *See* Organizations
Adner, R., 207
Adverse selection
agency theory and, 150–151
insurance and, 62
opportunism and, 75
Aerospace industry and vertical integration, 97–98
Afuah, A., 206
Agency theory
“A Survey of Agency Models of Organizations” and, 156–163
“Principals and Agents: An Overview” and, 145–149
“The Economics of Agency” and, 149–156
“Theory of the Firm: Managerial Behavior, Agency Costs, and Capital Structure” and, 163–165
analysis of, 96–97
anatomy of organizations and, 222–223
design of structures in, 148–149
evaluation of, 153–154
ex ante contracts and, 90–91
general overview of, 139–143
The Modern Corporation and, 143–145
monitoring and, 146, 152–153, 162
property rights and, 130–131, 221
and trust, 103–104
Ahuja, G., 205
Amit, R., 199
Analysis
of agency theory, 96–97
business history as, 94
of discrete structural alternative, 102–103
An Economic Analysis of Property Rights and, 123–125
game-theoretic, 105–107
governance and, 98–99
historian and, 182–183
net-present-value (NPV) rule and, 209
property rights and, 113–115
remediability as, 100–101
Appropriateness and executive aspirations, 5
Approximation, 45–46, 52
Argote, L., 204
Arrow, K. J.
“The Economics of Agency” and, 149–156
The Limits of Organization and, 55, 57–68
Artificial memory, 14
Artz, K. W., 204
Asset specificity
the aerospace industry, 97–98
agency theory and, 163
economic hostages and, 95

- transaction costs theory and, 87–88, 90–93
- Assets. *See also* Resources
Mobilizing Invisible Assets and, 183–185
property rights and nonhuman, 134–137
- Attention
as limiting factor in organizational activity, 21–22
rationality of organizations and, 14
scarcity of, 43–44
- Auditing of organizations
internal, 79, 84
operating decisions and, 83
- Authority
behavioral theory and, 4–5
and formal theory of employment relationship, 41
price relations and, 60–61
rationality of organizations and, 7–8
role of, 16–18
role of nonhuman assets and, 135–137
transaction costs theory and, 55, 66–67
- Autonomous adaptation, 101–102
- Balance and executive aspirations, 5
- Bargaining, 25. *See also* Conflict
- Barnard, C. I. *See The Functions of the Executive*
- Barney, J. B., 198
- Barzel, Y., 110, 123–125
- Bees and externalities, 71
- Behavioral economics
and business firm as organization, 41
decision making theories and, 44–46
formal theory of employment relationship and, 41–42
organizational theories and, 42–43
and scarcity of attention, 43–44
substantive/procedural rationalities and, 48–51
theories of bounded rationality and, 46–48
uncertainty and, 46–48
- Behavioral model of March/Simon, 22–23. *See also* Organizations
- Behavioral patterns. *See* Routines
- Behavioral theory of the firm. *See also* Behavioral economics; Individual behavior
A Behavioral Theory of the Firm, 33–40
- Behavioral theory of the firm
anatomy of organizations and, 20–21, 220, 223
and bargaining outcomes of conflict, 25–26
choice/decision making and, 2
communication and, 18–20
cooperative behavior and, 3–6
criterion of efficiency and, 20
decision making process and, 33–34
efficiency and, 10
information processing and, 21–22
and mechanisms of organizations, 22–23, 24–25
organizational behavior as, 8
organizational choice and, 35
organizational control and, 35–36
organizational expectations and, 35
organizational goals and, 34
problem solving and, 26–29
process of integration of, 14
psychology of organizations and, 12–14
rationality of organizations and, 11–13
reward schemes and, 23–24
- Behavioral uncertainty, 92
- Berle, A. A., 143–145
- Bias
bounded rationality and, 53
problemistic searching and, 37–38
- Bogner, W. T., 203
- Bounded rationality. *See also* Rationality
administrative behavior and, 7–9
analysis/antitrust implications of, 73–76
resource-based theory and, 223
Schumpeterian competition and, 188–189

- Bowman, E. H., 205
Brush, T. H., 204
- Calculated trust, 104
Capitalism, 178–183
Career expectations, 148
Carnegie School framework. *See also* Behavioral theory of the firm
Cuban Missile Crisis and, 39–40
evolutionary theory of economics and, 222
- Cash flow allocations
discounted cash flow (DCF) as, 209
transaction costs theory and, 84–85
- Ceteris paribus, 32
- Chandler, Alfred, D., 168, 178–183
- Chatterjee, S., 198
- Chi, T., 200
- Chicago School, 139, 142
- Choice. *See also* Decision making
behavioral theory and, 35
objective rationality and, 12–13
skills and, 190, 192
structure/boundaries of rationality and, 31
- Clark, K. B., 197–198
- Cliques and behavioral theory, 18–19
- Coase, R. H.
“The Fable of the Bees” and, 71
The Firm, the Market and the Law and, 55, 68–71
- Cockburn, I., 200
- Coff, R. W., 206
- Cognitive processes
decision making theories and, 45
evaluation of actions regarding goals and, 29
- Cohen, W. M., 197
- Common ownership, 89
- Communication
authority and, 4–5
behavioral theory and, 18–20
failures due to misinterpretations, 189
interdependence of organizations and, 31
as irreversible investment, 63–64
language limits and, 75–76
- mass marketing and, 179
rationality of organizations and, 13–14
- theory of the firm and, 170
verbal arguments, 97
- Compact strip production (CSP), 216–217
- Competition
cash flow realization and, 209
experimentation and, 177
property rights’ interest groups and, 114–115
research questions for, 224
Schumpeterian, 185–186, 188
and technical progress, 194–195
- Competitive dynamics and behavioral theory, 39
- Concepts of organizations, 10
- Conflict
bargaining outcomes of, 25
individual/social rationality and, 57–61
internal organizations and, 79
mathematical principal-agent problem and, 140
organizational goals and, 34
principal-agent problems and, 154–156
property rights and political, 110, 115
quasi-resolution of, 36
- Conner, K. R., 198
- Consensus, 66–67
- Constraints and rationality of organizations, 7
- Contracting for Property Rights*, 109–118
- Contracts
“The Fable of the Bees” and, 71
asset specificity and, 91–92
authority and, 66
Contracting for Property Rights and, 109–118
costs of, 74
economic implications of incompleteness of, 132–133
economic incentive and, 90
- Firms, Contracts, and Financial Structure and*, 128–137

- and formal theory of employment relationship, 41
governance and, 98–99
internal organizations and, 79–80
motivations for property rights, 114
opportunism and, 77
private failure and oil field unitization, 117
property rights and, 126
property rights approach to, 133–134
renegotiation of, 131–132
risk neutrality of agents in, 161–162
role of nonhuman assets and, 134–137
tournament, 162
transaction costs theory and, 69, 88
uncertainty and, 61–62
- Cool, K., 197
- Cooperation
behavioral theory and, 3–6
and economic gain from specialization, 119
game-theoretic analysis and, 106
internal organizations and, 78–79
organizations as systems of, 12
- Coordination of activities
and modern industrial enterprises, 181
role of authority and, 18
- Corporations. *See also* Agency theory
effect of separation of control and ownership on, 143–144
growth options for, 214
resource allocation and, 209
- Costs
agency theory and, 164–165
of contracts, 74
ex ante/post transaction, 89–91
knowledge and, 192
social, 70–71
sunk, 32
of transactions in property rights, 123–125, 131–132
- Cuban Missile Crisis and Carnegie framework, 39–40
- Culture
effect on organizations of, 51
property rights and, 123
- Cyert, R. M., 33–40, 53. *See also A Behavioral Theory of the Firm*
- Decentralization and organizational theory, 10
- Decision making
administrative organization and, 9
agency theory and, 162–164
authority and, 7–8
as basis of organization, 6
behavioral theory and, 2
bound rationality and, 45–46
cognitive limits on rationality and, 26
concept of purpose in, 10
design of search process in, 27
division of labor and, 21–22
executive, 5
as function of information, 65
influence of property rights on, 127–128
internal organizations and, 79
M-form organization and, 85–86
operating decisions and, 82
organizational learning and, 38–39
psychology of, 12–14
resources/property rights and, 113–114
role of authority and, 16–18
skills and, 190
stochastic elements of, 187–188
substantive/procedural rationalities and, 48–51
- Decomposition principles of hierarchy, 87
- Default option, 212–213
- Delegating. *See* Task allocation
- Dent-Micallef, A., 203
- Descriptive decision theory, 52–53
- Design and rationality of organizations, 6
- Dierickx, I., 197
- Discounted cash flow (DCF), 209
- Diversification of economies, 175–177

- Division of labor, 174
Divisionalization, 84–85
Dynamic capabilities approach. *See also*
 Resource-based theory
 anatomy of organizations and, 224
 Mobilizing Invisible Assets and,
 183–185
 theory of, 1
 transaction costs theory and, 104
Dynamic programming, 48
- An Economic Analysis of Property Rights, 123–125
Economic Behavior and Institutions, 125–128
Economic evolutionary theory
 decision making and, 187–188
 individual behavior and, 189–192
The Economic Institutions of Capitalism, 22
“The Economics of Agency”, 149–156
Economies of scale and transaction costs theory, 94–95
Economies of scope, 22
Efficiency
 agency theory and, 145
 agency theory over time and, 162
 agenda of organizations and, 64
 boundaries of the firm and, 130
 criterion of, 20
 divisionalization and, 84
 economic implications of incomplete contracts and, 132–133
 and evaluation of alternative organizations, 58
 informational overload and, 67
 of internal organization, 77–79
 organizational theory and, 10–11
 procedural rationality and, 49
 of property rights, 221
 property rights and, 109
 rationality of organizations and, 6
 technology and, 196
 transaction costs theory and, 56, 73–74
Eggertsson, T., 110, 125–128
Employment
 formal theory of, 41–42
 job idiosyncrasy and, 80–81
 nonhuman assets and, 136
 relationship of program/members and, 87
 as voluntary subordination, 72
Entrepreneurs
 contracts and, 69
 necessary investments of, 179–180
 neoclassical economic theory and, 42–43
 and theory of the firm, 171
Environment
 asset specificity and, 92
 decision making theories and, 45–46
 and economic gain from specialization, 119
 evaluation of actions regarding goals and, 29–30, 32–33
 governance and, 99–100
 institutional analysis and, 121–122
 Mobilizing Invisible Assets and, 183
 research questions for, 224
 resources and, 168
Equilibrium of organizations, 15–16
Ethics and executive aspirations, 5
Evolutionary economics
 behavioral theory as building block of, 1
 Carnegie framework and, 222
 dynamic competition and, 194–195
 normative organizational economics and, 195–196
 routine and, 193
 static selection equilibrium and, 194
An Evolutionary Theory of Economic Change, 185–208
Ex ante adverse selection.
 See Adverse selection
Ex ante/post transaction costs, 89–91
Exclusivity, 221. *See also* Efficiency
Executives, 3–6, 14
Expansion without mergers, 171–173
Expectations of authority, 67. *See also*
 Authority
Experience through knowledge, 172–173
Experimentation as part of competition, 177

- Externalities and "The Fable of the Bees", 71
 "The Fable of the Bees", 71
- Factual elements of decisions, 11
- Farjoun, M., 200, 203–204
- Feelings and executive aspirations, 5
- Fels Lecture, 55
- Filter (perception), 46
- Financial options within managerial operations, 210
- The Firm, the Market and the Law*, 55, 68–71
- Firms. *See also* Real options
 as collection of resources, 170
 as response to uncertainty, 220
 theory of the, 169–170, 177
- Firms, Contracts, and Financial Structure*, 128–137
- Fisher Body/General Motors merger, 136–137
- Flexibility
 Real Options: Managerial Flexibility and Startegy in Resource Allocation and, 208–215
 research questions for, 224
- Forbearance law, 103
- Formal organizations
 autonomous/Hayekian adaptation and, 102
 cooperative behavior and, 3–4
 transaction costs theory and, 61–64
- Formal theory of employment
 relationships, 41–42
- Foss, N. J., 201
- Franchises, 95–96
- Freedom in price systems/markets,
 see also Markets and Hierarchies: Analysis and Antitrust Implications, 58–59
- The Functions of the Executive*, 3–6
- Game theory, 105–107
- General Motors/Fisher Body merger, 136–137
- Germany as industrial country, 179, 183
- Goals. *See also* Conflict
- anatomy of organizations and, 20–21
behavioral theory and, 34
evaluation of actions regarding, 29–31
institutions and, 119
M-form organization and, 86
organizational, 16
organizational control and, 35–36
personal, 15
rationality of organizations and, 45
- Godfrey, P. C., 200–201
- Governance
 agency theory and, 139
 efficiency and, 92–93
 firm as, 89
 The Mechanisms of Governance and, 98–104
 research questions for, 224
 transaction costs theory and, 91, 94
- Governmental influence. *See* Political process
- Grant, R. M., 202
- Great Britain as industrial country, 179, 183
- Gresham's law of planning, 32
- Growth of firms
 and expansion without mergers, 171–173
 Hercules Powder Company and, 178
 inherited resources and, 173–175
 lack of historical precedents of, 183
 research questions for, 224
 through time, 176
- Hariharan, S., 198
- Hart, O., 110–111, 128–137
- Hayekian adaptation, 101–102
- Helfat, C. E., 202–203, 205, 207
- Henderson, R. M., 197–198, 200
- Hercules Powder Company, 178
- Hidden information/action, 75, 150–152
- Hierarchy
 decomposition principles of, 87
 environmental factors and, 74
 of modern industrial enterprises, 180–182
 and relationship of program/members, 28

- simple, 80
as substitute of internal organizations, 72
as ultimate contract law, 102
- Hill, C. W., 200–201
- Hoopes, D. G., 207
- Human behavior. *See Behavioral theory of the firm*
- Implementations and rationality of organizations, 6
- Incentives
“The Fable of the Bees” and, 71
agency theory and, 130–131, 147, 156
behavioral theory and, 4
divisionalization and, 84
effect of separation of control and ownership on, 142
ex ante, 91
inherited resources and, 173–174
internal organization auditing and, 79
and limits of firms, 94–95
multiple agents/repeated relations and, 153
oil field contracting and, 115–116
property rights and, 123
property rights as, 111
reward schemes and, 23–24
using new knowledge and, 174–175
- Individual behavior. *See also Personal ideals/interests*
economic evolutionary theory and, 189–192
research questions for, 224
- Informal organizations. *See also Organizations*
autonomous/Hayekian adaptation and, 102
within formal organizations, 4
- Information. *See also Communication; Uncertainty*
cash flow realization and, 209–210
and challenges in agency relationships, 145–146, 148–149
decision making theories and, 45–46
ex ante adverse selection as, 75
invisible resources and, 183
- limits of organization and, 55
manipulative, 84
natural/artificial memory as, 14
opportunism and, 77–78
overload, 67
processing of, 21–22
rationality of organizations and, 6
and scarcity of attention, 43–44
transaction costs theory and, 61–64
transmittance of, 19
trust and, 103–104
- Innovations and behavioral theory, 39
- Input
measurement of, 91–92
options to switch use and, 213–214
transaction costs theory and, 56
- Institutions. *See also Organizations*
Economic Behavior and Institutions
and, 125–128
and economic gain, 118–123
research questions for, 224
- Institutions, Institutional Change and Economic Performance*, 118–123
- Insurance
adverse selection and, 62
agency theory and, 140
hidden information/action and, 150–151
- Intended rationality, 8. *See also Rationality*
- Investments
Chandler, Alfred, D. and, 168
evaluation of alternative, 209
option to defer, 210–212
real options and, 216–217
- Irreversible investment, 64
- Itami, H., 168, 183–185
- Jensen, M., 163–165
- Judgement and executive aspirations, 5
- Katila, R., 205
- Knott, A. M., 207
- Knowledge
change analysis and, 69
forbearance law disputes and, 103

- importance of communication for, 13–14
incentive to using new, 174–175
organizational theory and, 43
routine as organizational memory of, 192
skills and, 190–192
through teaching/experience, 172–173
- Kogut, B., 201
- Kuemmerle, W., 206–207
- Language limits and bounded rationality, 75–76
- Learning
education as information source, 66
and importance of communication, 13–14
knowledge through, 172–173
organizational, 38–40
- Leasing franchises, 95–96
- Leg, 143
- Leonard-Baront, D., 199
- Levinthal, D. A., 156–163, 197
- Libecap, G. D., 109–118
- Lieberman, M. B., 204
- The Limits of Organization*, 55, 57–68
- Linear programming, 48
- Lippman, S. A., 196, 207–208
- Madhok, A., 206
- Madsen, T. L., 207
- Mahoney, J. T., 199, 201, 203, 206
- Makadok, R., 205, 208
- Management and real options, 208–215
- Management information system. *See* Information
- March, James G., 22–23, 33–40, 53
- Markets
adaptive behavior of, 73
analysis of failures of, 73–74
asymmetric information and, 83
diversification strategies and, 175–176
dynamic competition and, 195
freedom in price systems and, 58–59
human attributes and, 73
lack of historical precedents of growth, 183
- M-form organization and, 83, 85–86
and organizational capabilities, 180–182
resources and, 178
and transaction costs theory, 68–69
vertical integration and, 93
- Markets and Hierarchies: Analysis and Antitrust Implications*, 55, 72–87
- Mass marketing, 179
- Means, G. C., 143–145
- The Mechanisms of Governance*, 98–104
- Meckling, W., 163–165
- Memory and behavioral theory, 14
- Merck and Company, 215–216
- Merged firms. *See also* Expansion without mergers
- Merged/acquired firms, 133, 136–137
- M-form organization, 83, 85–86
- Microeconomics
economic evolutionary theory and, 188
economic well-being and, 121
governance and, 98, 100
transaction as ultimate unit of, 72–73
- Miller, K. B., 202
- Minimum wages, 57–58
- Mobilizing Invisible Assets*, 183–185
- Models of agency theory, 156–163
- Models of Bounded Rationality*, 2–3, 40–53. *See also* Rationality
- The Modern Corporation*, 143–145
- Monitored decision area, 65
- Montgomery, C. A., 197–198, 204
- Moral hazard, 150. *See also* Adverse selection
- Mosakowski, E., 199
- Motivation
capital and, 147
institutional analysis and, 121–122
and modern industrial enterprises, 181
problemistic searching and, 37–38
- Motivational constraints. *See also* Incentives and mechanisms of organizations, 24–25

- Mowery, D. C., 202
Mutual sunk cost economics, 89
- Natural memory, 14
Nelson, R. R., 168, 185–208
Neoclassic economy
 “A Survey of Agency Models of Organizations” and, 156–163
 agency theory and, 147
 agenda of organizations and, 64
 analysis assumptions of, 48–49
 and business firm as organization, 41
 choice/decision making and, 26–27,
 52–53
 and cognitive limits on rationality,
 25–26
 difficulties in, 26
 discrete structural alternative
 and, 103
 and formal theory of employment
 relationship, 41–42
 M-form organization and, 86
 organizational theories and, 42–43
 property rights and, 112
 rationality of organizations and, 8
Net-present-value (NPV) rule, 209
Nobel Prize winners
 North, Douglass, 109
 Simon, Herbert, 1
Normative organizational economics,
 195–196
Normative theory of procedural
 rationality, 49, 52–53
North, Douglass, 109–110, 118–123
Nucor, 216–217
- Oil fields
 contracting for, 115–118
 principal-agent problem and,
 154–156
- Opportunism
 organizational failures framework
 and, 79–80
 and resource-based theory, 223
 transaction costs theory
 and, 73–75, 77
as uncertainty, 92
- Optimality. *See Efficiency*
Organizational behavior
 according to Simon, 8
 equilibrium and, 15–16
Organizational capabilities
 behavior and, 192–195
 capitalism and, 180–182
 evolutionary theory of, 186–188
Organizational economics, 139
Organizational genetics, 186
Organizational influence, 15
Organizations, 22–33
Organizations
 advantages of internal, 78–79
 agenda of, 55, 64–66. *See also Goals*
 anatomy of, 20–21
 authority/responsibility and, 66–68
 business firm as, 41
 corporations as, 143
 decision making and, 9, 22–23.
 See also Decision making
 evaluation of alternative, 58
 failures framework of, 73–74, 79–80
 M-form, 83, 85–86
 predictions of behavior of other, 39
 reward schemes of, 23–24
 social functioning and, 67
 structure/boundaries of rationality of,
 31–32
 v. institutions, 119–120
- Output decisions
 behavioral theory and, 39
 options to switch use and, 213–214
 transaction costs theory and, 56
- Ownership. *See also Property rights*
 theory
Fisher Body/General Motors merger
 and, 136–137
separation of control and, 140–141,
 143–144
unified, 97
Oxley, J. E., 202
- P&G (game theory), 105–107
Pandian, J. R., 199
Pareto, Vilfredo. *See Efficiency*
Partitioned property rights, 127

- Partnerships and principal-agent problems, 154–156
- Passive decision area, 65
- Penrose effect, 172
- Penrose, E. T., 169–178
- Perceptual processes and decision making theories, 45–46
- Performance program, 28–31
reward schemes and, 23–24
- Persistence, 14. *See also* Attention
- Personal ideals/interests
behavioral theory and, 4–5
economic evolutionary theory and, 189–192
- individual rationality, 57–61
- and interference with organizational efficiency, 11
- organizational influences upon, 15
- research questions for, 224
- Peteraf, M., 199–200, 207
- Pisano, G. P., 203
- Planning
Gresham's law of, 32
integration of behavior and, 14
- Polanyi, Michael, 190
- Political process. *See also* Governance
economic power of corporations and, 144–145
property rights and, 110, 112, 114–115, 128
resource allocation and, 59
- Pool losses, 115–118
- Pool resources, 176
- Porter, M. E., 198
- Positive externalities, 71
- Powell, T. C., 203
- Pratt, J. W., 145–149
- Price systems
autonomous/Hayekian adaptation and, 102
freedom and, 58–59
limitations of, 61–62
transaction costs theory and, 69
- Pricing decisions and behavioral theory, 39
- Primary uncertainty, 92
- Principal-agent theory. *See also* Agency theory
“The Economics of Agency” and, 149–156
agency theory and, 140–141
as ex-ante incentives, 91
oil fields and, 154–156
- “Principals and Agents: An Overview”, 145–149
- Prioritization and strategizing, 104
- Private ownership, 111–113. *See also* Property rights theory
- Problem solving. *See also* Efficiency models for, 29
routine responses to, 26–27
searching and, 27–28, 37–38.
See also Decision making
- Procedural planning, 14
- Procedural rationality, 48–51
- Process
of integration of behavior, 14
organizational theory and, 9–10
- Programming, linear/dynamic, 48
- Property rights theory
agency theory and, 163, 221
analytical framework of, 113–115
anatomy of organizations and, 221–223
approach to contracts of, 133–134
Contracting for Property Rights, 109–118
An Economic Analysis of Property Rights and, 123–125
Economic Behavior and Institutions and, 125–128
economic incentive and, 90
efficiency and, 109
enforcement of, 122, 126
Firms, Contracts, and Financial Structure and, 128–137
formal constraints and, 122
historical aspects of, 120–121
institutions and, 118–123
microeconomics and, 100
oil field contracting and, 115–118
ownership and, 129–130
role of authority and, 135–137

- role of nonhuman assets and, 134–137
transaction costs and, 120–121
transactions costs and, 221
Proportion and executive aspirations, 5
Psychology
and neoclassical economics, 48–49
uncertainty and, 50–51
Purpose
importance of, 17–18
organizational theory and, 9–10
Quality dimensions of goods, 126
Rationality
and attention, 14
bounded, 7–9, 73–76, 188–189, 223
and business firm as organization, 41
cognitive limits on, 25–26
consequences of bounded, 53
decision making theories and, 44–46
formal theory of employment relationship and, 41–42
game-theoretic analysis and, 105–107
individual/social, 57–61
limits of organization and, 55
limits to, 11
in organizational behavior, 11–13
organizational failures framework and, 79–80
organizational theories and, 42–43
and scarcity of attention, 43–44
structure/boundaries of, 31–32
substantive/procedural, 48–51
theories of bounded, 46–48
transaction costs theory and, 73–76
uncertainty and, 46–48, 50–51
view of organizations and, 6–7
Real options
and abandonment for salvage value, 213
anatomy of organizations and, 224
applications of, 215–217
corporate growth options and, 214
and expansion/contraction/shut down/restart, 213
inputs/outputs and, 213–214
investment deferral and, 210–212
multiple interacting options and, 214
Real Options: Managerial Flexibility and Strategy in Resource Allocation and, 208–215
time-to-build option and, 212–213
Real Options: Managerial Flexibility and Startegy in Resource Allocation, 208–215
Recommendations of role of authority, 17
Redundancy in information/resources, 44
Renegotiation in contracts, 131–132.
See also Contracts
Replication, 192–193
Reputations of products, 148
Resource-based theory. *See also Evolutionary economics*
anatomy of organizations and, 223
and behavioral theories of the firm, 221–223
contributions to, 196–208
dynamics of capitalism and, 178–183
growth of the firm and, 169–178
invisible resources and, 183–185
Resources
agency theory and, 162–163
behavioral theory and allocation of, 39
categories of, 174
criterion of efficiency and, 20
efficiency approaches based on, 104
and the environment, 168
firm as pool of, 176
governmental influence and, 59
and importance of property rights, 111
inherited, 173–175
invisible, 182–183
limits of natural, 58
market transactions and, 178
nonhuman, 134–137
oil field contracting and, 115–118
options to accelerate, 213
property rights and, 113

- property rights and allocation of, 126–127
real options and, 215
Real Options: Managerial Flexibility and Startegy in Resource Allocation and, 208–215
research questions for, 224
and scarcity of attention, 43
social cost and, 70–71
as source of innovation, 177
and theory of the firm, 170, 177
transaction costs theory and
 allocation of, 61–62
- Responsibility
 executive aspirations and, 5
 operating decisions and, 82
 role of authority and, 17–18
 transaction costs theory and,
 55, 67–68
- Risk
 agency theory and, 156, 161
 as consequence of
 decision-making, 26
 franchises and, 96
 opportunism and, 74
 principal-agent problem and, 152
- Roehl, T. L., 168, 183–185
- Roth, K., 205
- Routines
 behavioral patterns, 187, 189–190
 Carnegie framework and, 222
 functions of, 192–193
 rationality of organizations and, 7
 research questions for, 224
- Rumelt, R. P., 196, 207–208
- Sales contracts, 41–42. *See also* Contracts
Salvage value, 213
Scale and Scope: The Dynamics of Capitalism, 168, 178–183
Schoemaker, P. J. H., 199
Schumpeterian competition, 185–186, 188
Searching and problem solving
 action alternatives/consequences
 and, 31–32
- approximation features
 and, 45–46, 52
behavioral theory and, 37–38
rational decision and, 27–28
- Secondary uncertainty, 92
Selection and behavior, 9
Sense and executive aspirations, 5
Shamsie, J., 202
Shareholders
 agency theory and, 141, 164
 of corporations, 144
 hidden information/action and, 150
 research questions for, 224
- Shuen, A., 203
- Silverman, B. S., 202, 204
- Simon, Herbert A., 1–3, 22–33, 53, 186
- Skills
 individual behavior and, 189–192
 routine and, 193
- Social cost, 70–71
- Society
 declining in economic well-being, 120
 property rights as costly to, 113–114
 property rights as economic incentive
 system in, 111–113
 rationality and, 57–61
- Span of control, 63
- Specialization (division of labor), 174
- Specialization by function, 10
- Spender, J. C., 202
- State of rest, 174
- Strategic management
 agency theory and, 139
 divisionalization and, 84
 future of, 219–220
 real options and, 214–215, 217
 research questions for, 224
 transaction costs theory and, 56, 87
- Subgroup gains and internal
 organizations, 78–79
- Substantive rationality, 48, 50
- Sunk costs
 game-theoretic analysis
 and, 105–107
 inertia and, 32
 Nucor and, 216–217
 property rights and, 112

- "A Survey of Agency Models of Organizations", 156–163
- Systems viewpoint of behavioral theory, 6
- Szulanski, G., 202, 208
- Task allocation, 22
- Technology
- and applications of real options, 215–217
 - dynamic competition and, 194–195
 - entry into new markets and, 175
 - flexible production, 217
 - inefficiency due to, 196
 - intangible assets and, 182
 - mass marketing and, 179
 - organizational form and, 72
 - and resources to define property rights, 123
- Teece, D. J., 196–197, 203
- "Theory of the Firm: Managerial Behavior, Agency Costs, and Capital Structure", 163–165
- The Theory of the Growth of the Firm*, 169–178
- Third world countries and property rights, 123
- Thomas, H., 203
- Thomke, S., 206–207
- Time-to-build optionreal options and, 212–213
- Tournament contracts, 162
- Training, 16. *See also* Learning
- Transaction costs theory
- adaptation and, 101–102
 - agenda of organizations and, 55, 64–66
 - anatomy of organizations and, 220–221, 223
 - asset specificity and, 87–88, 90
 - authority/responsibility and, 66–68
 - behavioral theory as building block of, 1
 - bounded rationality and, 73–76
 - business history and, 94
 - change analysis and, 69
 - comparative efficiency criteria and, 56
- corporate finance/governance and, 103
- An Economic Analysis of Property Rights* and, 123–125
- ex ante/post, 89–91
- firm as governance structure and, 89
- formal organizations and, 61–64
- game-theoretic analysis and, 105–107
- governance and, 91, 98–104
- influential students of, 53
- and limits of firms, 94–95
- limits of organization and, 57–68
- market transactions and, 72
- markets and, 68–69
- measurement and, 91–92
- microeconomics and, 72–73
- multidivisional structure and, 82–84
- rationality and, 55, 57–61
- remediableness and, 100–101
- social cost and, 70–71
- sources and, 131–132
- unilateral/bilateral applications and, 95–96
- vertical integration and, 81–82
- Transactions
- information available to contractual parties in, 146
 - mass marketing and, 179
- Transferability, 221. *See also* Efficiency
- Transportation and mass marketing, 179
- Trigeorgis, L., 208–215
- Tripsas, M., 203
- Trust and agency theory, 103–104
- Uncertainty
- administrative process and, 220
 - avoidance of, 36–37
 - behavioral economics and, 42
 - cash flow realization and, 209
 - categories of, 92
 - organizational failures framework and, 79–80
 - project appraisal under, 209
 - theories of bounded rationality and, 46–48, 50–51
 - transaction costs theory and, 61–62

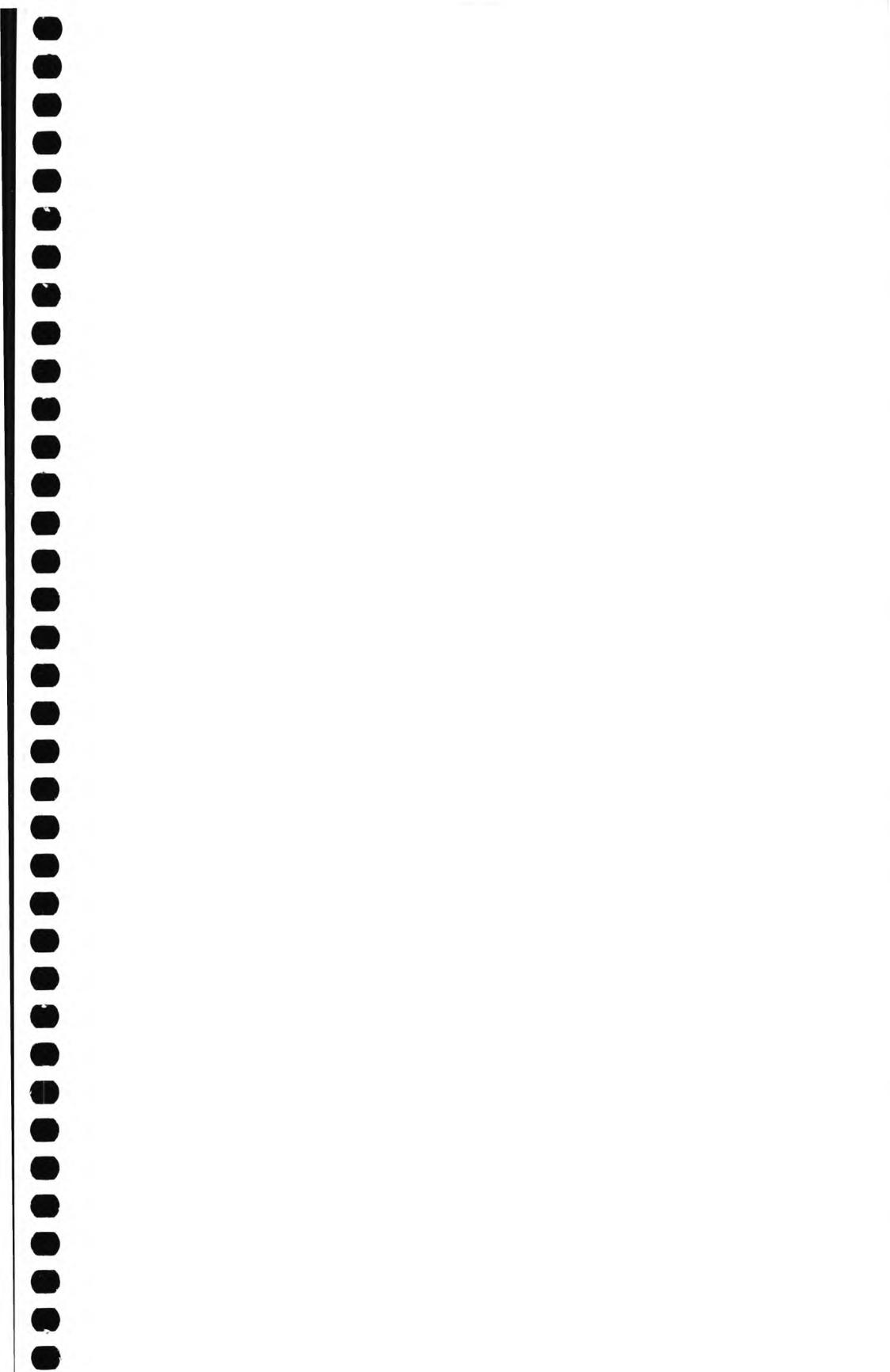
- Unified ownership, 97
- United States as industrial country, 179, 183
- Unity of command, 10
- Universality, 221. *See also* Efficiency
- Value elements of decisions, 11
- Vertical integration
- in the aerospace industry, 97–98
 - differences between market/internal organizations and, 92–93
- Firms, Contracts, and Financial Structure* and, 128–137
- Fisher Body/General Motors merger and, 136–137
- game-theoretic analysis and, 105
- as substitute of internal organizations, 72–73
- transaction costs theory and, 81–82
- Virutous circle, 174
- W. R. Grace Corporation, 215
- Walker, G., 207
- Wal-Mart (game theory), 105–107
- Wenerfelt, B., 197–198
- Williamson, O. E.
- The Economic Institutions of Capitalism* and, 87–98
- Markets and Hierarchies: Analysis and Antitrust Implications* and, 55, 72–87
- The Mechanisms of Governance* and, 98–104
- resource-based theory and, 198–199
- transaction costs theory and, 205
- Winter, S. G., 168, 185–208
- Yeoh, P.-L., 205
- Zander, U., 201
- Zeckhauser, R. J., 145–149

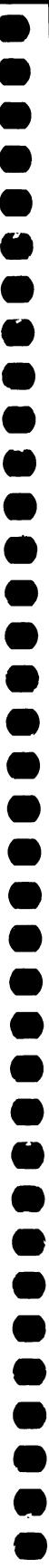


About the Author

Joseph T. Mahoney received his BA in economics at the University of Pennsylvania in 1980, his MA in 1984, and his PhD in 1989 in business economics at the Wharton School of Business at the University of Pennsylvania. Joe joined the faculty of the Department of Business Administration at the College of Business at the University of Illinois at Urbana-Champaign in January 1988. Joe was promoted to Associate Professor in 1995, and he was promoted to Full Professor in 2003. Currently, Joe is on the editorial board of *Academy of Management Review*, *Journal of Management Studies*, and *Strategic Management Journal*. From 2000 to 2002, he was the book review editor of *Academy of Management Review*. He has published in *Academy of Management Review*, *Journal of Business Research*, *Journal of Management*, *Journal of Management Inquiry*, *Journal of Management Studies*, *Managerial and Decision Economics*, and *Strategic Management Journal*, among others. Joe currently serves on the editorial boards of *Academy of Management Review*, *Journal of Management Studies*, and *Strategic Management Journal*.







"Economic Foundations of Strategy uniquely contributes to the strategic management discipline by discussing and integrating the most important theoretical approaches in strategic management research. The book articulates the essence of these theories creatively and eloquently. Currently, our field lacks a scholarly book that brings together these theoretical building blocks of strategic management research. Mahoney's book is a superb guide for doctoral students and scholars who are entering the discipline of strategic management. I conjecture that this book will become a must-have guide for seasoned strategy researchers as well . . . scholars from related disciplines such as industrial organization economics, international business, and organization would also want to have a copy of this phenomenal guide. . . . It was a pleasure to read this book."

—Yasemin Y. Kor, University of Delaware

Economic Foundations of Strategy provides not only the essential basic tenets of strategy, it also shows the interrelationships of five major theories of the firm: behavioral theory; transaction costs theory; property rights theory; agency theory; and dynamic resource-based theory. Even though technological, organizational and institutional change advances breathlessly, the theories of the firm provided in this research book are durable principles that have stood, and the author maintains will continue to stand, the test of time. **Economic Foundations of Strategy** emphasizes the complementarities among these five theories of organization, and the potential for integrating these theories in the evolving science of organization. Applications of these theories to business practice are emphasized throughout the book.

Key Features:

- Provides an integrative approach
- Covers economics (e.g., transaction costs, property rights, and agency theory)
- Covers finance (e.g., real options theory)
- Covers organization theory (e.g., the behavioral theory of the firm)
- Covers modern strategic management theory (e.g., dynamic resource-based theory)
- Connects "know-why" and "know-how"
- Covers the theoretical material deeply enough to be of use to advanced students
- Connects this research to business applications to be of use to those interested in business practice

Visit our Web site at www.sagepublications.com

ISBN 978-1-4129-0543-5



9 0 0 0 0 >