MULTINATIONAL ENTERPRISE AND ECONOMIC ANALYSIS

THIRD EDITION

Richard E. Caves

CAMBRIDGE

CAMBRIDGE www.cambridge.org/9780521860130

This page intentionally left blank

MULTINATIONAL ENTERPRISE AND ECONOMIC ANALYSIS
THIRD EDITION

Multinational enterprise is an important subject for students and researchers, both practitioners of business administration and scholars of economics. This highly accessible book surveys the fruits of research from both quarters. It shows how economic analysis can explain multinationals' activity patterns and how economics can shed conceptual light on problems of business policies and managerial decisions arising in practice. It addresses the welfare problems arising from multinationals' activities and the logic of governments' preferences and choices in their dealings with multinationals. The book is readily useful to both researchers and students. This third edition incorporates knowledge about multinationals accumulated over the past decade of research.

Richard E. Caves is the Nathaniel Ropes Professor of Political Economy, Emeritus, at Harvard University. He is the author of many articles on multinational enterprise and other topics in the fields of international economics and industrial organization. He is the co-author of a leading textbook on international economics, *World Trade and Payments*. His most recent books deal with the arts and entertainment industries, *Creative Industries: Contracts between Art and Commerce* and *Switching Channels: Organization and Change in TV Broadcasting*.

CAMBRIDGE SURVEYS OF ECONOMIC LITERATURE

Multinational Enterprise and Economic Analysis

Third Edition
Richard E. Caves, *Harvard University*

Multinational Enterprise and Economic Analysis, surveys the contributions that economic analysis has made to our understanding of why multinational enterprises exist and their consequences for the workings of the national and international economies. Covering both theories and tests of hypotheses, and synthesizing material from social science and applied disciplines, Richard E. Caves develops the logic behind public policies that affects multinational enterprises. He shows the points of harmony and conflict between national policies and sources of discrepancy between national and world welfare. The book summarizes what economists have learned about why multinational firms stretch their activities across national borders and what consequences their presence has for competing local businesses, public policy-makers, and the general public.

Professor Caves has completely updated the book to cover contributions appearing through 2005. Many parts have been rewritten to reflect new ideas and lines of research appearing since the second edition. It remains the only survey volume to draw equally on analytical contributions of both economics and business administration bearing on the multinational firm.

Editor
Professor John Pencavel, Stanford University

The literature of economics is rapidly expanding, and within just a few years, many subjects have changed in recognition. Perceiving the state of knowledge in fast-developing subjects is difficult for students and time consuming for professional economists. This series of books is intended to help with this problem. Each book gives a clear structure to and balanced overview of the topic and is written at an intelligible level for the senior undergraduate. They will be useful for teaching as well as provide a mature, yet compact, presentation of the subject for economists wishing to update their knowledge outside their own specialties.

Other titles in the series
Mark Blaug, *The Methodology of Economics: Or How Economists Explain* (Second Edition)
Robert Clark and Joseph Spengler, *The Economics of Individual and Population Aging*
Thrainn Eggertsson, *Economic Behavior and Institution*
Anthony C. Fisher, *Resource and Environmental Economics*
Helmut Frisch, *Theories of Inflation*
Jack Hirshleifer and John G. Riley, *The Analytics of Uncertainty and Information*
Peter Isard, *Exchange Rate Economics*
Morton I. Kamien and Nancy L. Schwarts, *Market Structure and Innovation*
Anne O. Krueger, *Exchange Rate Determination*
Todd Sandler and Keith Hartley, *The Economics of Defense*
Steven M. Sheffrin, *Rational Expectations*
John B. Shoven and John Whalley, *Applying General Equilibrium*
E. Roy Weintraub, *Microfoundations: The Compatibility of Microeconomics and Macroeconomics*

Multinational Enterprise and Economic Analysis

Third Edition

RICHARD E. CAVES
Harvard University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK
Published in the United States of America by Cambridge University Press, New York
www.cambridge.org
Information on this title: www.cambridge.org/9780521860130

First published in print format 2007

ISBN-13 978-0-511-29281-1 eBook (Adobe Reader)
ISBN-10 0-511-29281-3 eBook (Adobe Reader)

ISBN-13 978-0-521-86013-0 hardback
ISBN-10 0-521-86013-X hardback

ISBN-13 978-0-521-67753-0 paperback
ISBN-10 0-521-67753-X paperback

Contents

Preface

The multinational enterprise (MNE) has attracted much writing, scholarly and otherwise. Treatises bedecked with boxes and arrows instruct business managers on how to run MNEs. Passionate polemics chronicle their alleged misdeeds and call for the regulatory hand of government. Between these poles are found reams of description and comparison. Economic analysis has certainly not neglected the MNE. However, when the first edition of this book was written, the analytical treatments seemed seriously fragmented, as each branch of economic analysis carved its initials into the MNE without worrying much about what other branches made of it. This book's first edition (1982) therefore sought to integrate the research literature in two ways. It characterized the MNE as one form of internalization of transactions, thus placing it in the transaction-cost approach to economic organization, and integrated this core concept with the findings about MNEs reported by each standard functional branch of economic analysis. The second form of integration drew together theory and evidence, the former largely the domain of economics, the latter found adrift in the seas of business administration, political science, and the like. This integrative effort apparently proved useful to readers, which is why this organizational structure has survived two revisions essentially unchanged.

The book was written to reach a rather heterogeneous audience. It aims mainly to serve scholars in economics and business administration. Although it lacks the apparatus of a textbook, it was designed to also provide collateral reading for students in courses that touch on multinational enterprise. That goal accounts for a few patches of undergraduate-level textbook exposition of key concepts and models. The exposition level in general is aimed at users with substantial economics background, but not necessarily much prior acquaintance with every branch of economics that sheds some light on MNEs.

When this book first appeared (1982), its titular subject could be regarded as new. In the 1960s and 1970s, a critical mass emerged consisting of cogent empirical description and insightful applications of economic theory. It seemed feasible (to me, anyhow) to provide a wall-to-wall summary treatment in a book of reasonable length. The critical mass provided an opening for scholars with many interests and research strategies. The ensuing rapid growth of scholarly interest in foreign direct investment warranted occasional revision of this book. A second edition in 1996 followed.

Revising the book presented some difficult choices, however. The stock of knowledge (think of the upthrusting mountain of scholarly pages published on the MNE) was growing apace. Should this summary of that stock grow proportionally? Clearly, its coherence and pedagogical value would decay rapidly – a dismaying outcome for both me and the publisher. Should the book jettison the old and embrace the new (*Multinational Enterprise since Last February*)? Because we are dealing with an accumulating body of knowledge (rather than, say, sartorial fashion), that approach's flaw is immediately apparent. Providing a framework within which innovations can be organized and understood is essential to this survey's usefulness.

That leaves the approach actually taken. The book's account of the basic knowledge stock remains its core, although that core has undergone some compression. The shadow price for reportage of new contributions has been raised somewhat. Where several research papers make overlapping contributions, only one may be summarized. One consequence: a higher shadow price for recent contributions than for early ones. The early contributions and insights (this book's original core) retain their place in the presentation, just as the book itself retains the framework that emerged in the subject's first flowering. For example, the book retains references to early case-study material but barely heeds recent cases (the journal article has become the unit of account). Recent theoretical contributions have probably suffered the most; we attempt to cover important recent contributions to theory but omit the more specialized confections. Also missing from the book is quantitative and descriptive information concerning the MNE population.

Time waits for no survey author. I aimed to cover current contributions up to the beginning of 2006. Apologies are extended to 2005 (and earlier!) authors who were unintentionally omitted. In pursuit of timeliness, I did pay more attention than in previous editions to the current stock of working papers, especially those from the National Bureau of Economic Research and the Centre for Economic Policy Research series. The gain in timeliness from including their nascent contribution surely offsets any substantive gain from waiting out the purifying fires of journals' referee process.

In this revision, each chapter and section has undergone significant change. However, over the past decade, new contributions have been concentrated strongly in three areas. First, advances in international trade theory have brought it closer to the distinctive features of MNEs. Second, business administration researchers have swarmed over the question of modal choice: When does the firm entering a new market forgo a wholly owned subsidiary for exporting, a joint venture, or a contractual arrangement? Third, productivity spillovers have received much attention. Does the MNE affiliate spill significant productivity gains onto its neighbors and market competitors? When does it absorb spillage from them?

For encouraging timely revision of this book, I am grateful to Scott Parris of Cambridge University Press and for efficient production assistance to Ann Flack.

1

The Multinational Enterprise as an Economic Organization

The multinational enterprise (MNE) is defined here as an enterprise that controls and manages production establishments – plants – located in at least two countries. It is simply one subspecies of a multiplant firm. We use the term "enterprise" rather than "company" to direct attention to the top level of coordination in the hierarchy of business decisions; a company, itself multinational, might be the controlled subsidiary of another firm. The minimum "plant" abroad needed to make an enterprise multinational is judgmental. The transition from a foreign sales subsidiary or a technology licensee to a producing subsidiary is not always a discrete jump, for good economic reasons. What constitutes "control" over a foreign establishment is another judgmental issue. An MNE sometimes chooses to hold only a minor fraction of the equity of a foreign affiliate. Countries differ in the minimum percentage of equity ownership that they count as a "direct investment" abroad, as distinguished from a "portfolio investment," in their international-payments statistics.

Exact definitions are unimportant for this study because economic analysis emphasizes that at definitional margins decision-makers face close tradeoffs rather than bimodal choices. However, the definition does identify the MNE as essentially a multiplant firm. We are back to Coase's (1937) classic question of why the boundary between the administered allocation of resources within the firm and the market allocation of resources between firms falls where it does. In a market economy, entrepreneurs are free to try their hands at displacing market transactions by increasing the scope of allocations made administratively within their firms. The Darwinian tradition holds that the most profitable pattern of enterprise organization should ultimately prevail. To explain the existence and prevalence of MNEs, we require models that predict where the multiplant firm enjoys advantages from displacing the arm's-length market and where it does not. In fact, the

prevalence of multiplant (multinational) enterprises varies greatly from sector to sector and from country to country, providing opportunities to test models of the MNE.

The models of the multiplant firm potentially relevant to explaining the presence of MNEs are quite numerous and rather diverse in their concerns. It proves convenient to divide them into three groups: (1) One type of multiplant firm produces broadly the same line of goods in each geographic market where it operates. Such firms are common in domestic industries with fragmented local markets such as metal containers, bakeries, and brewing. Similarly, the many MNEs that establish plants in different countries to make the same or similar goods can be called horizontally integrated. (2) Another type of multiplant enterprise produces outputs in some of its plants that serve as inputs to its other activities. Actual physical transfer of intermediate products from one of the firm's plants to another is not required by the definition; it needs only to produce at adjacent stages of a vertically related set of production processes. (3) The third type of multiplant firm is the diversified company whose plants' outputs are neither vertically nor horizontally related to one another. As an international firm it is designated a diversified MNE.

1.1. Horizontal Multiplant Enterprises and the MNE

We start by equating the horizontal MNE to a multiplant firm with plants in several countries. Its existence requires, first, that *locational forces* justify dispersing the world's production so that plants are found in different national markets. Given this dispersion of production, there must be some *governance* or *transaction-cost advantage* to placing the plants (some plants, at least) under common administrative control. This abstract, static approach provides the most general and satisfying avenue to explaining the multinational company. The location question – why plants are spread around the world as they are – is addressed in Chapter 2. We assume at first that plant *A* was located in southeast England because that was the lowest-cost way to serve the market it in fact serves. We also assume that this locational choice was not essentially influenced by whether the plant was built by an MNE, bought by an MNE, or not owned by an MNE. The static approach also puts aside the vital question of why a company grows into MNE status – something more readily explained after the static model is in hand.

The transaction-cost approach asserts, quite simply, that horizontal MNEs will exist only if the plants they control and operate attain lower costs or higher revenue productivity than the same plants under separate

managements. Why should this net-revenue advantage arise? Some of the reasons have to do with minimizing costs of production and associated logistical activities of the firm. The more analytically interesting reasons – and, we shall see, the more important ones empirically – concern the complementary nonproduction activities of the firm.[1]

Proprietary Assets

The most fruitful concept for explaining the nonproduction bases for the MNE is that of assets having these properties: The firm owns or can appropriate the assets or their services; they can differ in productivity from comparable assets possessed by competing firms; the assets or their productivity effects are mobile between national markets; they may be depreciable (or subject to augmentation), but their life spans are not short relative to the horizon of the firm's investment decision.[2] Successful firms in most industries possess one or more types of such assets. An asset might represent knowledge about how to produce a cheaper or better product at given input prices or how to produce a given product at a lower cost than competing firms. The firm could possess special skills in styling or promoting its product that make it such that the buyer differentiates it from those of competitors. Such an asset has revenue productivity for the firm because it signifies the willingness of some buyers to pay more for that firm's product than for a rival firm's comparable variety. Assets of this type are closely akin to product differentiation – the distinctive features of various sellers' outputs cause each competing firm to face its own downward-sloping demand curve. The proprietary asset might take the form of a specific property – a registered trademark or brand – or it might rest in marketing and selling skills shared among the firm's employees. Finally, the distinctiveness of the firm's marketing-oriented assets might rest with the firm's ability to devise frequent innovations; its proprietary asset then might be a patented novelty, or simply some new combination of attributes that its rivals cannot quickly or effectively imitate. This asset might vary greatly in tangibility and specificity. It could take the specific form of a patented

[1] This approach developed through the works of a number of authors, including Hymer (1960, 1968), Eastman and Stykolt (1967), Kindleberger (1969), Johnson (1970), Caves (1971), McManus (1972), Buckley and Casson (1976), Dunning (1977*a*, 1981*b*), Magee (1977*a*), and Hennart (1982).

[2] No single term used in the literature captures all these conditions. "Proprietary assets" seems to come closest, but "intangible assets," "firm-specific assets," and "monopolistic advantages" generally have the same meaning.

process or design, or it might simply rest on know-how shared among employees of the firm. It is important that the proprietary asset, however it creates value, might rest on a set of skills or repertory of routines possessed by the firm's team of human (and other) inputs (Nelson and Winter, 1982, Chapter 5).

The proprietary assets described by these examples evidently share the necessary conditions to support foreign investment. They are things that the firm can use but not necessarily sell or contract upon. Either the firm can hold legal title (patents, trademarks) or the assets are shared among the firm's employees and cannot be easily copied or appropriated (by other firms or by the employees themselves). They possess either the limitless capacities of public goods (the strict intangibles) or the flexible capacities of the firm's repertory of routines. Especially important for the MNE, while the productive use of such an asset is not tightly tied to a single physical site or even nation, arm's-length transfers of them between firms are prone to market failures. These failures deter a successful one-plant firm from selling or renting its proprietary assets to other single-plant firms and thereby foster the existence of multiplant (and multinational) firms. Proprietary assets are subject to a daunting list of infirmities for being detached and transferred by sale or lease:

1. They are, at least to some degree, *public goods.* Once a piece of knowledge has been developed and applied at a certain location, it can be put to work elsewhere at little extra cost and without reducing the capacity available at the original site. From society's point of view, the marginal conditions for efficient allocation of resources then require that the price of the intangible asset be equal to its marginal cost, zero, or approximately zero. However, no one gets rich selling bright ideas for zero. Therefore, intangible assets tend to be underprovided or to be priced inefficiently (at a net price exceeding their marginal cost) or both.
2. Transactions in intangibles suffer from *impactedness* combined with *opportunism.* This problem is best explained by examples: I have a piece of knowledge that I know will be valuable to you. I try to convince you of this value by describing its general nature and character. But I do not reveal the details because then the cat would be out of the bag, and you could use the knowledge without paying for it unless I have a well-established property right. Therefore, you decline to pay me as much as the knowledge would in fact be worth to you because you suspect that I am opportunistic and inflate my claims.

3. A proprietary asset might be diffuse and therefore incapable of an enforceable lease or sale contract. The owning firm might readily contract with a customer to achieve a specific result using some competence that it possesses but be unable to contract to install that competence within another firm. Even with well-defined intangibles, various sources of uncertainty can render contractual transfers infeasible or distort the terms of viable deals.

This application of modern transaction-cost analysis underlies a framework widely used in research on the MNE. It layers a third necessary condition for horizontal MNEs atop the two already asserted – the efficiency of dispersed *location* of production and the efficiency of common *ownership* of the dispersed facilities. The third condition, *internalization*, holds that the decentralized application of the proprietary asset is more efficiently managed within the owning firm than by renting it at arm's length to another firm. This framework, developed mainly in Dunning's (e.g., 1981*b*) writings, is commonly called the OLI (ownership location internalization) paradigm. It is controversial only as to its sufficiency to explain all MNEs' operations; it clearly lacks that sufficiency, as it does not apply to the cases of vertical and diversified MNEs (Rugman, 1985; Teece, 1986).

Some Extensions

The proprietary-assets approach embraces certain extensions and variants. Although the standard exposition contemplates a goods-producing firm, it evidently applies as well to MNEs in the services sector.[3] The site of production of a service is sometimes indefinite, and accordingly, it is not subject to the clear dichotomy between exporting and foreign production that is applicable to a good. Although a hotel chain serves customers at the site of the service's consumption, a consulting firm does not (Boddewyn, Halbrich, and Perry, 1986; Enderwick and Associates, 1989; UNCTC, 1989). The proprietary assets of service multinationals seldom result from research investments, but they commonly rest on information and capabilities that effectively yield economies of scale and scope and support goodwill assets. Also, some service MNEs (but not only they) possess an important special type of proprietary asset that is transaction specific. In transaction-cost economics,

[3] The value of foreign investments in services probably accounts for 40 percent of the capital invested in foreign subsidiaries according to the United Nations Centre on Transnational Corporations (hereafter UNCTC, 1989), but the research literature is locked into a goods-production mind-set.

a transaction-specific asset exists in some resources, facilities, or knowledge. It may exist simply in each party's cumulated trust that the owner will not cheat in their mutual dealings. The switching costs that they incur if they change transaction partners support a persistent supplier-customer relation that can deter either party from taking temporary advantage of the other. As empirical evidence subsequently demonstrates, the proprietary assets that drive foreign investment in some business services seem to be strongly transaction specific, with service MNEs emerging to preserve and benefit from the parent's ties to customers who themselves have become MNEs.

Another extension pertains to the longevity of proprietary assets. The standard approach is one of comparative statics: A domestic firm is assigned some fixed proprietary asset, and its profitable exploitation through foreign direct investment is deduced. Proprietary assets can be enlarged or improved through investment, however, and such investment decisions should themselves depend on the firm's opportunities to undertake foreign investments. Foreign investments might be undertaken to develop or to improve proprietary assets. Such assets are also subject to depreciation and obsolescence, and their deterioration might lead to foreign divestment as a reversal of the foreign-investment process (Boddewyn, 1983). The creation and destruction of such assets and the variance of returns in the investments that firms make in them should be reflected in the longevity and turnover of foreign investments themselves (Caves, 1995).

Studies of domestic multiplant operation (Scherer et al., 1975) indicate a number of economies directly relating to the firm's production activities, and these can apply to the MNE if they do not stop at the national boundary. There can be transaction-cost economies in the procurement of raw materials that go beyond the input needs of the single plant. Economies can arise in the transportation network for outbound shipments of finished goods that extend beyond the single plant's output. Localized demand or cost fluctuations can warrant coordinated use of plants' capacities, so that several plants' outputs can be flexibly shipped from the temporarily favored site (de Meza and van der Ploeg, 1987; Kogut and Kulatilaka, 1994). If the industry's output consists of a line of diverse goods, each plant might efficiently specialize in some items rather than each producing the whole array. It is an empirical question how fully these economies are available to a multiplant firm operating across national boundaries because they depend on the cost of moving goods (inputs or outputs) among plants or the effectiveness of managerial coordination of distant activities.

Another asset of the ongoing firm is its capacity to generate investible funds beyond what it can profitably use for expanding its current activities.

One view of the ongoing firm's financial decisions holds that it attaches different opportunity costs to funds from various sources. Externally secured funds – debt and new equity – are costly because of transaction costs and moral-hazard problems and the reduced independence they entail for the managers, as well as the direct cost of paying additional interest or dividends. Internally generated funds – profits not paid out to current shareholders – have a lower opportunity cost, and managers will put them to work in a new activity with an expected profit rate (internal rate of return) lower than what would be needed to warrant external borrowing. Thus, excess capacity in internally generated funds can also motivate foreign investment.[4] Indeed, this point generalizes further to the advantage an established company might have in entering a foreign market simply because excess profits can be earned there, and the firm stands near the front of the queue of potential entrants in terms of its ability to overcome whatever entry barriers sustain the excess profits. The implications of this point for the MNE as a market competitor are discussed in Chapter 4, and empirical evidence appears in Section 9.3's discussion of MNEs originating in less-developed countries.

Finally, the firm's choice of foreign investment for maximizing the returns to its proprietary assets in foreign markets is made against an array of alternative arrangements involving arm's-length deals with other firms. When the proprietary asset is a patent, trademark, or well-defined technology, licensing or franchising it to other firms might be the owner's preferred strategy (technology licensing is reviewed in Chapter 7). When a value-creating activity requires proprietary assets that two (or more) firms must contribute, and outright merger of the firms is not efficient, various alliances, cooperative arrangements, and joint ventures can be employed (Dunning, 1984; Oman, 1984; Buckley, 1985; Hennart, 1989). For example, a firm might prefer some contractual arrangement to serve a small foreign market where establishing its own subsidiary requires an otherwise avoidable fixed cost (Anderson and Gatignon, 1986). Other cooperative arrangements and management-services contracts can become instruments of choice when host governments cannot credibly commit to eschew expropriation (or its equivalent in taxation) once the MNE has sunk its foreign investment (see Section 4.4). Evidence on these forms of inter-firm agreement will be noted subsequently because they compete with foreign investment as a way to maximize returns on proprietary assets.

[4] The financial model of the firm that underlies these propositions has less than universal acceptance among economists but agrees with evidence summarized in Section 6.1.

Empirical Evidence: Prevalence of Horizontal Foreign Investment

Hypotheses about horizontal MNEs have received many statistical tests. The usual strategy of research involves relating the prevalence of MNEs in an industry to structural traits of that industry: If attribute *x* promotes the formation of MNEs, and successful firms in industry *A* have a lot of *x*, then MNEs should be prevalent in industry *A*. These tests have been performed on two dependent variables: foreign operations of firms in a source country's industries normalized by their total activity level in those industries (hereafter "outbound" foreign investment) and foreign subsidiaries' share of activity in a host country's markets normalized by total transactions in those markets (hereafter "inbound" foreign investment). The exogenous variables are chosen to represent features of industries' structures that should either promote or deter foreign direct investment. These econometric studies are prone to at least two types of misspecification that have led to certain modified research strategies. First, foreign investment substitutes for other methods (exporting, licensing foreign producers) of maximizing rents on proprietary assets in foreign markets. A given industry's share of foreign investment might be high either because foreign investment works well or because the alternatives work badly. The most attractive way to address this problem is to measure the extent of use of the alternative methods and test the determinants of all of them together (Buckley and Casson, 1998). Second, the extent to which country 1's firms invest abroad depends not only on the absolute properties or qualities of their own proprietary assets but also on the qualities of assets held by firms competing with them in foreign markets. The data requirements for dealing head on with this problem are onerous, but some progress has been made in studies of bilateral foreign-investment patterns.

The number of studies embodying these designs has grown large enough to sustain its own monograph-length survey (UNCTC, 1992*a*). Here the main conclusions will be summarized, with reference only to selected articles. There is considerable agreement on the major results among studies of both outbound and inbound investment, among studies of a given type for each country, and among studies based on different countries. Therefore, we offer here some generalizations about the principal conclusions without referring extensively to the conclusions reached in individual studies or about particular countries. Then we take up extensions and qualifications. Findings about the trade-off between foreign investment and exporting are treated in Chapter 2 and about the trade-off between foreign investment and other forms of association between business units in Chapter 7.

First, a roster of the main statistical studies of outbound foreign investment includes, for the United States, Horst (1972*a*), Wolf (1977), Pugel (1978, Chapter 4, 1981*a*), Goedde (1978, Chapter 2), and Lall (1980); for Sweden, Swedenborg (1979); and for Japan, Kogut and Chang (1991) and Drake and Caves (1992). The principal studies of inbound foreign investment include, for the United States, Lall and Siddharthan (1982), Caves and Mehra (1986), and Wesson (1993); for Canada, Caves (1974*b*), Baumann (1975), Saunders (1982), and Owen (1982); for Great Britain, Dunning (1973*b*), Caves (1974*b*), Hughes and Oughton (1992) and Giulietti, McCorriston, and Osborne (food sector) (2004); for Germany, Yamawaki (1985); for Australia, Parry (1978) and Ratnayake (1993); and for India, N. Kumar (1990). Their results confirm, first and foremost, the role of proprietary assets inferred from the outlays that firms make to create and maintain these assets. Research and development intensity (R&D sales ratio) is a thoroughly robust predictor. Advertising intensity has proved nearly as robust, even though most studies have lacked an appropriately comprehensive measure of firms' sales-promotion outlays.[5] Researchers also consistently find a significant positive influence for an industry's intensive use of skilled managerial labor; this variable seems to confirm the "repertory of routines" basis for foreign investment, independent of the strictly intangible proprietary assets (Pugel, 1981*a*). (More comprehensive measures of labor skills also exert statistically significant positive effects in some studies, but it is unclear what hypothesis they test.) A third result that also supports a role for the firm's general coordinating capacity is the positive influence of multiplant operation within large countries such as the United States. This hypothesis was advanced and given some statistical support by Eastman and Stykolt (1967, Chapter 4); both Caves (1974*b*) and Saunders (1982) confirmed that multiplant operations in the United States are a significant positive predictor of foreign investment in adjacent Canada, although Caves found that the hypothesis is not confirmed for remote, insular Great Britain.[6] A final result confirms both the role of intangible assets and the transaction costs that arise for protecting property rights in them: An industry's extent of

[5] More and Caves (1994) showed that intra-firm royalty receipts by MNE parents (after controlling for transfer-pricing distortions) behave like cash flows resulting from foreign investments that transplant the MNE's intangible assets. Survey evidence gathered by Bertin and Wyatt (1988, pp. 25–29) showed that MNEs regard technology advantages as their most potent competitive advantage, followed by marketing and managerial assets.

[6] Juhl (1985) confirmed it for Germany. Useful demonstrations of the nature of proprietary assets other than intangibles lie in studies of MNEs based in "unlikely" source countries such as Canada (Rugman, 1987).

foreign investment increases with the proportion that lawyers make up of its total employment (Denekamp, 1995).

Other tests have dealt with sources of entry barriers that might concentrate production in particular locations. Some evidence indicates that extensive scale economies in production deter the dispersion of plant operations and thus retard foreign investment. Also, some investigators have tested the hypothesis that activities requiring (absolutely) large capital investments might favor the multinational activity of existing large enterprises. None of these hypotheses has been supported robustly, although support for the scale-economies hypothesis is noted in Chapter 2. The hypotheses are not finely tuned, and many studies suffer from the inclusion of such variables as an industry's average firm size or the concentration of its producers, which are themselves endogenous, collinear with other exogenous variables, and lead to results that are sensitive to specification choices and generally untrustworthy.

Included in many of these cross-section models are variables seeking to capture the positive influence of tariff protection of the host-country market or (alternatively) the ease or cost advantage with which a host-country market can be served through exports rather than foreign investment. These are discussed in Chapter 2. The important point is that they have rather little explanatory power compared to variables based on proprietary assets, which embody necessary conditions for foreign direct investment.

Several specialized issues do need to be noted here:

1. *Development of proprietary assets.* The cross-section tests summarized so far neglect the development and turnover of stocks of proprietary assets. This process is most easily seen in studies of individual firms, but it does exert some influence at the national level. Drake and Caves (1992) showed how the development of proprietary assets in Japan's manufacturing industries in the 1970s and 1980s led to subsequent increases of Japan's share of foreign investments in U.S. industries. Cantwell (1989, Chapters 2, 6) explored the long-run relationship between nations' stocks of proprietary assets, reflected in patents, and their revealed comparative advantage in gathering rents on world markets. The association is closer for exports and overseas production taken together than it is for exports alone.
2. *Rivalrous relationships between source- and host-country assets.* The relativity of competing companies' proprietary assets can be tested only at a broad national level (see Chapter 2) or through analyzing industry-level flows of investment between pairs of countries. Kogut and Chang

(1991) explored the roles of both Japanese and U.S. R&D expenditures in influencing the rate of Japanese foreign investment. It turns out to be positively related to both flows; there is no positive relationship to the Japan-U.S. differential, as one would expect if the two expenditure flows create adversary proprietary assets. Apparently spillovers and positive externalities are the dominant factor for R&D, but Pugel, Kragas, and Y. Kimura (1996) found that Japanese foreign investment is repelled by the marketing outlays of U.S. competitors. Kim and Lyn (1987) observed a negative relationship between foreign investment in U.S. industries and the market value of U.S. firms with which they compete – specifically, the component of those market values not explained by the U.S. industries' own R&D and advertising levels and their concentration ratios.

3. *Foreign investment to augment proprietary assets.* Related to the result of Kogut and Chang is the hypothesis that (some) foreign investment takes place to draw on host-country assets in order to augment the proprietary assets of the entering MNE. Case-study evidence documents extensively this motive for foreign investment, for example, Japanese foreign investments in research-intensive industries of the United States and Germany (Alsegg, 1971, pp. 218–30; Tsurumi, 1976, pp. 116–17; Yoshida, 1987, pp. 47–48). The United States remains the natural market in which to test the hypothesis. Wesson (1993) argued that a U.S. industry's share of world exports is the best available proxy for intangibles found in the U.S. market that could serve this purpose. He found that foreign investment in the United States increases with several variables that indicate the stock of relevant U.S. assets (such as classes of skilled labor), either directly or interacted with the export-share measure of the U.S. advantage. Researchers have used several approaches to probe the importance of "base-augmenting" (vs. base-exploiting) investments, often stacking it against the traditional hypothesis that foreign investment proceeds from more research-intensive sources to less research-intensive destinations. Neven and Siotis (1996) found that high host-market R&D intensities lure intra-European foreign investment, but Anand and Kogut (1997) found no predominance of either exploiting or augmenting for entrants into a broad range of U.S. manufacturing industries. Ruckman (2005) concluded that pharmaceutical firms making acquisitions in the United States typically bought targets more R&D-intensive than themselves – but this tendency was less prevalent in international than in domestic mergers. Kuemmerle (1999) focused

closely on MNEs' R&D laboratories located abroad, finding that their activities tend to be bifurcated – either mainly base exploiting or base augmenting but seldom divided evenly between them. In the pharmaceutical and electronics companies covered in his study, 38 percent were base augmenters. Augmenting is more prevalent in high-R&D countries and countries with heavy investments in human capital. Cantwell and Mudambi (2005) studied foreign subsidiaries in the U.K. engineering industries, linking their possession of product mandates from their parents to their R&D levels and choices of location.

4. *International mobility of proprietary assets.* Some research addresses the international mobility of proprietary assets by explaining why competing firms in an industry differ in their propensities to invest abroad. Horst (1974*a*, Chapters 4 and 5) explored the effects of various corporate assets on the foreign-investment behavior of firms in the U.S. food-processing sector. The proprietary assets held by these firms divide roughly into two classes. Some succeeded on the basis of heavy national advertising, others with extensive and intricate distribution systems for bringing their products to the final consumer in good condition. The latter group has taken part less extensively in foreign investment because these complex distribution systems are a drain on managerial resources and are not readily replicated in foreign markets. The advertisers, on the other hand, are heavy foreign investors. The firms with intensive distribution systems also display less extensive multiplant development within the United States, suggesting that the diseconomies of scale in extending their empires constrain them geographically within the United States as well as internationally.[7] Statistical studies (Horst, 1972*b*; Grubaugh, 1987*a*) have confirmed the role of different endowments of competing firms as predictors of their MNE status. Belderbos and Sleuwaegen (1996) in particular showed that predictions based on Japanese firms' endowments extend even to the destinations of their foreign investments.
5. *Evidence from market valuations of firms.* Another method recently employed to test the proprietary-assets approach is by means of information on the stock market's valuations of MNEs. Morck and Yeung (1991) analyzed ratios of market to book value (Tobin's q) for U.S. MNEs to show that the market's valuation of these firms increases

[7] Similarly, Meredith (1984) demonstrated the pull of foreign direct investment to Canada associated with spillovers of U.S. sales-promotion outlays across the Canadian border.

with their investments (R&D, advertising) in proprietary assets and with the extent of their multinational operations. This influence of foreign operations, however, depends on and operates through these outlays on proprietary assets. (Otherwise, multinationality might be valued by shareholders for diversification or tax advantages that it provides instead of for rents on proprietary assets; see Sections 1.3 and 6.2). Morck and Yeung (1992) studied the stock market's valuations of announcements that a U.S. firm had acquired control of a firm located abroad. Although U.S. shareholders' reactions to domestic mergers tend to be insignificantly positive or even negative for the acquiring firm, Morck and Yeung found a significant positive response to the average foreign acquisition. Further, the response increases with the firm's rate of spending on proprietary assets – R&D (especially for small acquiring firms) and advertising (especially for large ones). Gupta et al. (1991) similarly observed positive responses to joint ventures announced in the People's Republic of China.

6. *Proprietary assets in the macroeconomy*. Recent research has linked the abundance (or paucity) of a country's stock proprietary assets to traits of its macroeconomy. Yeaple (2003*b*) found that the education level of a country's labor force (and consumers) interacts positively with the prevalence of nonproduction workers in its industries, in explaining the extent of its foreign investment. Proprietary assets represent accumulated and encoded knowledge, so the country well suited to develop and export proprietary assets is the one with skills to devise such assets and the propensity to use them heavily alongside tangible inputs. With this link in hand, it becomes feasible to incorporate the geographic and locational forces along with proprietary assets to explain countries' varying levels of involvement in foreign direct investment (Chapter 2). Furthermore, countries can be regarded symmetrically as sources and hosts of foreign investment.

Multinationals in Service Industries

Horizontal MNEs in banking and other services have received increased attention from researchers. The proprietary-assets hypothesis again makes a good showing, especially when extended to the transaction-specific assets of an ongoing quasi-contractual relationship between the service enterprise and its customer. A bank, advertising agency, or accounting firm acquires a good deal of specific knowledge about its client's business, and the parties'

sustained relationship based on trust lowers the cost of contracting and the risks of opportunistic behavior. The service firm enjoying such a quasi-contractual relation with a parent MNE holds a transaction-cost advantage for supplying the same service to the MNE's foreign subsidiaries. If the service must be supplied locally, the service firm goes multinational to follow its customer.

Much casual evidence reveals this transaction-specific asset behind service industries' foreign investments (e.g., Safarian, 1966, p. 210; Behrman, 1969, pp. 3–4), especially in the banking sector (Grubel, 1977, and references cited therein; Pastré, 1981; Yannopoulos, 1983; Enderwick and Associates, 1989, pp. 61–78). Grubel affirmed the transaction-cost model but also cited two other factors. Some banks acquire particular product-differentiating skills analogous to those found in some goods-producing industries; they can explain banks' foreign investments in less-developed countries (Baum, 1974) and in countries with large populations of migrants from the source country. Also, national banking markets commonly appear somewhat non-competitive because of cartelization or regulation or both, and foreign banks are well-equipped potential entrants. The traits of foreign banks' operations in the United States affirm these propositions (Lees, 1976). Some propositions about internalization in banking have been tested statistically. Miller and Parkhe (2002) provided a test of some liabilities of foreign banks due to their foreignness. Nigh, Cho, and Krishnan (1986) found increases in U.S. bank assets abroad to vary significantly by country with increases in the overall book value of the U.S. foreign-investment position, with the openness of the host country's policies controlled. Sagari (1992) confirmed the same proposition for levels of banking and nonbanking foreign investment. Heinkel and Levi (1992) symmetrically showed the prevalence of foreign countries' banks in the United States to increase with the country's exports and with the value of financial claims that the U.S. holds on the source country's capital market. Li and Guisinger (1992) found that the growth of foreign investment by a source country's service (all services sectors) MNEs increases significantly with the source's total initial stocks of foreign investment; the closeness of the relationship declined from the 1970s to the 1980s.

The prominence of transaction-specific assets as a factor driving foreign investment is apparently matched in service industries such as advertising agencies, accounting, and consulting firms (Terpstra and Yu, 1988; Enderwick and Associates, 1989, pp. 79–106). Studies of other multinational service industries, however, bring out different factors. International hotel chains resemble franchise operations in creating centrally a proprietary asset (standardized product image, reservation system) that must be combined

with other inputs at the site of consumption. No sharp economic boundary exists between domestic and international hotel franchises, and Dunning and McQueen (1982) showed that international hotel chains' penetration of various national markets is inversely related to the development of franchise systems in the domestic industry. International construction firms rely on repertories of routines and reputation assets resembling those that commonly support MNEs in manufacturing (Enderwick and Associates, 1989, pp. 132–51).

1.2. Vertically Integrated MNEs

The vertically integrated MNE is readily regarded as a vertically integrated firm whose production units lie in different nations. Theoretical models that explain vertical integration should therefore be directly applicable. Again, we assume that production units are dispersed in different countries because of conventional location pressures – the bauxite mine where the bauxite is found, bauxite converted to alumina at the mine because the process is strongly weight losing, and the smelter that converts alumina into aluminum near a source of low-cost electric power. The question is, Why do they come under common administrative control? The proprietary-assets model is not necessary because neither upstream nor downstream production units need bring any distinctive qualification to the parties' vertical consolidation. Some proprietary advantage of course *could* explain which producer operating at one stage undertakes an international forward or backward vertical integration.

Models of Vertical Integration

Until the rise of transaction-cost economics the economic theory of vertical integration contained a large but unsatisfying inventory of special-case models. Some dealt with the physical integration of production processes: If you make structural shapes out of the metal ingot before it cools, you need not incur the cost of reheating it. Such gains from physical integration explain why sequential processes are grouped in a single plant, but they neither preclude two firms sharing that plant nor explain the common ownership of far-flung plants. Other traditional models regard vertical integration as preferable to a stalemate between a monopolistic seller and a monopsonistic buyer, or to an arm's-length relation between a monopolistic seller and competitive buyers whose activities are distorted because of paying the monopolist's marked-up price for their input. Some models

explain vertical integration as a way around monopolistic distortions, while others explain it as a way to profit by fostering such distortions.

The theory of vertical integration has been much enriched by the same transaction-cost approach that serves to explain horizontal MNEs. Vertical integration occurs, the argument goes, because the parties prefer it to the ex ante contracting costs and ex post monitoring and haggling costs that would mar the alternative state of arm's-length transactions. The vertically integrated firm internalizes a market for an intermediate product, just as the horizontal MNE internalizes markets for proprietary assets.[8] Suppose that pure competition prevails in each intermediate-product market, with large numbers of buyers and sellers, the product homogeneous (or its varied qualities costlessly evaluated by the parties), information about prices and availability in easy access to all parties in the market. Neither seller nor buyer would then have reason to transact repeatedly with any particular party on the other side of the market. When these assumptions do not hold, however, both buyers and sellers acquire motives to make long-term alliances. The two can benefit mutually from investments that each makes suited to special attributes of the other party. Each then incurs a substantial fixed cost upon shifting from one transaction partner to another. Each seller's product could be somewhat different, and the buyer faces significant costs of testing or adapting to new varieties, or merely learning the requirements and organizational routines of new partners. The buyer and seller gain an incentive to enter into some kind of long-term arrangement.

If transaction-specific assets deter anonymous spot-market transactions, they leave open the choice between long-term contracts and vertical integration. Contracts, however, encounter the costs of negotiation and of monitoring and haggling previously mentioned. These ex ante and ex post costs trade off against one another – a comprehensive contract can reduce subsequent haggling – but the overall cost remains.[9] The problem is compounded because, even in a market with many participants, unattached alternative transaction partners tend to be few *at any particular time* when a party might wish to recontract. Fewness compounds the problems of governance in arm's-length vertical relationships.

[8] O. E. Williamson (1985) deserves credit for developing and popularizing this approach. For a survey of models of vertical integration, see Perry (1989).

[9] Economists make the point that the uncertainties impelling vertical integration could be averted by resorting to comprehensive forward-contract markets, if they existed (Buckley and Casson, 1976). Because they do not exist for the same reasons that vertical integration emerges, the point lacks operational significance.

One special case of the transaction-cost theory of vertical integration holds promise for explaining MNEs involved in processing natural resources. Vertical integration can occur because of failings in markets for information, as analyzed earlier in the context of proprietary assets. A processing firm must plan its capacity on some assumption about the future price and availability of its key raw material. The producers of that raw material have the cheapest access (perhaps exclusive) to that information, but they have an incentive to overstate availability to the prospective customer: The more capacity customers build, the higher they are likely to bid in the future for any given quantity of the raw material. Therefore, vertical integration could occur to evade problems of impacted information (Arrow, 1975).

To summarize, intermediate-product markets can be organized in a spectrum of ways stretching from anonymous spot-market transactions through a variety of long-term contractual arrangements at arm's length to vertical integration. Switching costs and durable, specialized assets discourage spot transactions and favor one of the other modes. If, in addition, the costs of negotiating and monitoring arm's-length contracts are high, the choice falls on vertical integration (or some less extensive pooling of equity). These empirical predictions address both where vertical MNEs will appear and how they will trade off against contractual relationships.

Empirical Evidence

Much less research has addressed vertically related MNEs than the horizontal ones just reviewed. Much of what does exist addresses not only the causes of vertical linkages but their prevalence. Indeed, it indicates that vertical linkages are widespread and have been increasing. U.S.-domiciled subsidiaries of foreign MNEs obtain about two-thirds of their imported inputs from their corporate affiliates. For individual affiliates, these flows tend to be steady over time rather than being displaced by local production as a subsidiary matures (Zeile, 1998). Indeed, Keane and Feinberg (2005*a*) discovered a large increase in the proportion of total shipments by Canadian subsidiaries to their U.S. parents between 1984 and 1995. Numerous studies have found foreign subsidiaries to outsource more than their host country's domestic firms (e.g., Girma and Görg, 2004). MNEs in industries making heavy use of (nonubiquitous) natural resources tend to place large proportions of their assets abroad (Pugel, 1978; Owen, 1982). Japanese MNEs invest heavily abroad in countries that are large suppliers of imports to Japan – whose imports run heavily to raw materials and fabricated inputs (Farrell,

Gaston, and Sturm, 2004). Various studies associate intra-MNE trade flows with affiliates located in small or low-wage host countries and not too distant from the parent (Hanson et al., 2005; Milner, Reed, and Talerngsri, 2004). These statistical findings do not, however, directly address the similarity or difference between intra-corporate and arm's-length trade flows. A rare exception is Celly, Spekman, and Kamauff (1999), who investigated determinants of relationship-specific investments made by foreign suppliers to U.S. firms. These actually increase with technological uncertainty and with the importance to the buyer of the supplier's responsiveness. The competitiveness of the market in which the supplier operates is not statistically significant. This study suggests that arm's-length contracts in such vertical relationships succeed in evading the theoretical hazards facing them. A near-total gap in this literature is evidence on how arm's-length relationships are sustained; Dyer and Chu (2000) established the important role of trust between auto industry suppliers and assemblers. The correlates of trust seem consistent with the theory of repeated games.

For evidence more relevant to the transaction-cost determinants of vertical organization, we must rely on older research and case studies. McKern (1976) studied a group of extractive industries in which the vertical interface between extractive and processing stages holds central importance. Monopoly/monopsony market structures he found could not explain foreign investment in Australia's extractive industries. Also, he could not assign much importance to the foreign MNEs' motive of ensuring themselves access to supplies because, in many cases, they did not transfer Australian raw materials directly to their own refining facilities but instead sold them on the open market. Accordingly, he argued that an important motive for vertical integration is the use by MNEs of the knowledge they have acquired about the international market for the raw materials in question. This basis for vertical integration in MNEs adds up to a proprietary-assets explanation, analytically similar to the one that proves so fruitful for explaining horizontal MNEs. Case studies reaching this conclusion include Read (1983) and Chalmin (1986).

Much information exists on individual extractive industries in which MNEs operate on a worldwide basis, and this case-study evidence merits a glance in lieu of more broadly based findings. For example, Stuckey (1983) found the international aluminum industry to contain not only MNEs integrated from the mining of bauxite through the fabrication of aluminum products but also a network of long-term contracts and joint ventures. Market participants are particularly unwilling to settle for spot transactions in

bauxite (the raw ore) and alumina (output of the first processing stage). The problem is not so much the small number of market participants worldwide as the extremely high switching costs. Alumina refining facilities need to be located physically close to bauxite mines (to minimize transportation costs), and they are constructed to deal with the properties of specific ores. Likewise, for technical and transportation-cost reasons, aluminum smelters are somewhat tied to particular sources of alumina. Arm's-length markets, therefore, tend to be poisoned by the problems of small numbers and switching costs. And the very large specific and durable investments in facilities also invoke the problems of long-term contracts that were identified earlier. Finally, Stuckey gave some weight to Arrow's model of vertical integration as a route to securing information: Nobody knows more about future bauxite supplies and exploration than an existing bauxite producer.

A good deal of evidence also appears on vertical integration in the oil industry. The ambitious investigations have addressed the U.S. segment of the industry, but there appears to be no strong difference between the forces traditionally affecting vertical integration in national and international oil companies.[10] These studies give considerable emphasis to the costs of supply disruption faced by any nonintegrated firm in petroleum extraction or refining. Refineries normally operate at capacity and require a constant flow of crude-oil inputs. Storing large inventories of input is quite costly, and so backward integration that reduces uncertainty about crude supplies can save the refiner a large investment in storage capacity. It also reduces risks in times of "shortages" and "rationing," when constraints somewhere in the integrated system (crude-oil supplies are only the most familiar constraint) can leave the unintegrated firm out in the cold. The hazard of disrupted flows translates into a financial risk, as vertically integrated firms have been found to be able to borrow long-term funds more cheaply than those with exposure to risk (Greening, 1976, Chapter 1).

Country-based studies of the foreign-investment process have also regarded vertical MNEs as the outcome of failed arm's-length market transactions. Japanese companies became involved with extractive foreign investments only after the experience of having arm's-length suppliers renege on long-term contracts, and they also experimented with low-interest loans to independent foreign suppliers as a way to establish commitment (Tsurumi, 1976, Chapter 2).

[10] By "traditionally" we mean before the OPEC cartel became fully effective in the early 1970s. See Penrose (1968, pp. 46–50, 253–59), Greening (1976), and Teece (1976, Chapter 3).

Vertical Integration: Other Manifestations

Some international vertical relationships illustrate not only the problems of contracting but also the payout when it works well. Writers on offshore procurement and the associated international trade commonly refer to the role of foreign investment in transplanting the necessary know-how and managerial coordination (Helleiner, 1973; Sharpston, 1975). Jarrett (1979, Chapters 7 and 8; also see Helleiner, 1979; and Lee, 1986). Jarrett explored statistically both the structural determinants of this type of trade and the role of MNEs in carrying it out. His data pertain to imports under a provision of the U.S. tariff whereby components exported from the United States for additional fabrication abroad can be reimported with duty paid only on the value added abroad. His statistical analysis explains how these activities vary both among U.S. industries and among countries taking part in this trade. His results confirm the expected properties of the industries that make use of vertically disintegrated production: Their outputs have high value per unit of weight, embody reasonably mature technology (so are out of the experimental stage), are produced in the United States under conditions giving rise to high labor costs, and are easily subject to decentralized production.[11] Among overseas countries, U.S. offshore procurement favors those not too far distant (transportation costs) and with low wages and favorable working conditions. With these factors controlled, the component flows increase with the extent of U.S. foreign investment, both among industries and among foreign countries.[12]

A considerable amount of vertical integration is also involved in the "horizontal" foreign investments described in Section 1.1, and the behavior of horizontal MNEs cannot be fully understood without recognizing the complementary vertical aspects of their domestic and foreign operations. Many foreign subsidiaries do not just produce their parents' goods for the local market; they process semifinished units of that good, or package or assemble them according to local specifications. Pharmaceuticals, for example, are prepared in the locally desired formulations using basic ingredients

[11] Jarrett measured this last by the extent of multiplant operation of companies in the United States and by the extent to which U.S. producers depend on inputs purchased from other establishments in the same industry.

[12] If the presence of foreign investment is associated with offshore procurement, it should also be true that the factors influencing the proportion of U.S. imports that come from overseas corporate affiliates should include these same determinants of offshore procurement. This proposition is confirmed in Jarrett's analysis (1979, Chapter 2) of related-party imports to the United States.

imported from the parent. The subsidiary organizes a distribution system in the host-country market, distributing partly its own production, but with its line of goods filled out with imports from its parent or other affiliates.[13] Or the subsidiary integrates forward to provide local distribution. These activities are bound up with the development and maintenance of the enterprise's goodwill asset, as described earlier, through a commitment of resources to the local market. The firm can thereby assure local customers, who are likely to incur fixed investments of their own in shifting their purchases to the MNE, that the company's presence is not transitory. This consideration helps explain foreign investment in some producer-goods industries for which the proprietary-assets hypothesis otherwise seems rather dubious (Tsurumi, 1976, Chapter 4).[14] All of these activities represent types of forward integration by the MNE, whether into final-stage processing of its goods or into ancillary services.

The evidence of this confluence of vertical and horizontal foreign investments mainly takes the form of case studies. It is emphasized in the study of foreign investments by West German enterprises by Fröbel, Heinrichs, and Kreye (1980, Chapter 12). It is implied by the extent of intra-corporate trade among MNE affiliates – flows that would be incompatible with purely horizontal forms of intra-corporate relationships. Imports of finished goods by Dutch subsidiaries from their U.S. parents (Stubenitsky, 1970, p. 102) are high (as percentages of the affiliates' total sales) in just those sectors where imports might complement local production for filling out a sales line – chemicals (24.9 percent), electrical equipment (35.4 percent), and transportation equipment (65.5 percent). The prevalence of intra-corporate trade in engineering industries also suggests the importance of components shipments (U.S. Tariff Commission, 1973, pp. 284, 314–20). The case studies of intra-firm trade in Casson and Associates (1986) showed the importance of this forward integration for innovative and complex manufactured goods.

Statistical evidence on U.S. exports and imports passing between corporate affiliates sheds light on this mixture of vertical and horizontal foreign investment. Lall (1978*b*) analyzed the factors determining the extent of U.S. MNEs' exports to their affiliates (normalized either by their total exports or by their affiliates' total production). He could not discriminate between two

[13] Nicholas (1983) emphasized vertical foreign investment in distribution, following upon failed arm's-length contracts, as a critical step in the development of many British MNEs.

[14] Also, Jarrett (1979, Chapter 3) found that the extent of foreign investment by U.S. industries increases with the percentage of their product lines deemed to require frequent or extensive sales or technical services to customers. This influence is significant with other influences such as advertising and research intensity taken into account.

hypotheses that together have significant force: (1) that trade is internalized where highly innovative and specialized goods are involved and (2) that trade is internalized where the ultimate sales to final buyers must be attended by extensive customer engineering and after-sales services. Jarrett (1979, Chapter 2; also see Helleiner and Lavergne, 1979) confirmed these hypotheses with respect to the importance in U.S. imports of inter-affiliate trade, which in his data includes exports by foreign MNEs to their manufacturing and marketing subsidiaries in the United States as well as imports by U.S. MNEs from their overseas affiliates. Jarrett also found evidence that inter-affiliate trade in manufactures reflects several conventional forms of vertical integration: More of it occurs in industries populated (in the United States) by large plants and companies, capable of realizing the scale economies accessible in the international disintegration of production, and in industries that carry out extensive multiplant operations in the United States.

Sleuwaegen and Yamawaki (1991) showed that the prevalence of Japanese foreign investment in U.S. distribution (relative to manufacturing) is greater for durable and heterogeneous goods that cannot be promoted to buyers simply through media advertising. The productivity of foreign investments in forward integrated distribution activities is shown directly by Yamawaki's (1991) finding that such investments in the U.S. distribution sector contributed substantially to increasing Japanese exports to the United States. As Williamson and Yamawaki (1991) showed, these investments get the foreign MNE over a substantial entry barrier into distribution that provides rents to manufacturers who surmount it.

The entwining of vertical and horizontal relations has important corollaries for the behavior of MNEs that will emerge in later chapters. For example, it suggests why the expansion of output by foreign subsidiaries can coincide with expansion of the parent's production for export to the same market. A purely horizontal relationship between parent and subsidiary implies that their outputs will be substitutes for one another, whereas the confluence of horizontal and vertical relations raises the possibility that they are complementary within the MNE. Evidence lending some support to this proposition will be reviewed in Chapters 2 and 5.

1.3. Portfolio Diversification and the Diversified MNE

This section completes the typology of international multiplant firms by considering those whose international plants have no evident horizontal or vertical relationship. An obvious explanation of this type of MNE (though not the only one, it turns out) lies in the spreading of business risks. Going multinational in any form brings some diversification gains to the enterprise,

and these are increased when the firm diversifies across "product space" as well as geographical space.

Gains from Diversification versus Losses from Uncertainty

Economic analysis normally assumes that individual investors are risk averse and hence seek to compose portfolios of assets so as to eliminate nonsystemic risks associated with particular securities (companies), leaving them to face only system-wide risks. For this purpose, the international diversification of portfolios holds an obvious attraction, although that process might be inhibited by various factors discussed in Chapter 6.

Given the diversification achieved by shareholders, the value-maximizing firm's management selects a risk/return trade-off that values risk at the market price of residual, systemic risk (Greenberg, Marshall, and Yawitz, 1978). It is widely recognized, however, that firms might behave as if averse to risks specific to the enterprise itself. This behavior could result even with optimal principal-agent contracts between the firm's owners and its manager because risks to the firm's survival threaten its employees with large adjustment costs. Also, the firm as a working coalition of heterogeneous inputs – a characterization notably consistent with the standard model of the horizontal MNE – has a substantial organizational investment at hazard of obsolescence. The likely reaction of MNEs to opportunities for international diversification can be viewed against this background. On the one hand, individual foreign investments might be regarded as particularly risky. Risks arise in the behavior of host-country governments that in many ways can disfavor an alien firm lacking local support. Also, information on the host-country market is more costly to the foreign investor than to the native; even after rational investments in information, the MNE settles for incomplete knowledge and hence exposes its investment to a larger variance of expected outcomes. On the other hand, the firm that makes investments in several national markets should enjoy diversification gains, benefiting not only itself but also shareholders, if they are constrained from international diversification (Chapter 6). The larger variance of international projects fights against the lower correlation between the returns to the firm and the "market factor," making it unpredictable whether investors will place a premium or a penalty on the MNE's cash flows.

Empirical Evidence

Now we consider empirical evidence on diversification as a motive for the MNE. Within a national economy, many shocks affect most firms rather

similarly – recessions, major changes in macroeconomic policy. Between countries, such disturbances are less closely correlated. Also, changes in exchange rates and terms of trade tend to favor business profits in one country while worsening them elsewhere.[15] Statistical evidence confirms that MNEs enjoy gains from diversification: The larger the share of foreign operations in total sales, the lower the variability of the firm's rate of return on equity capital (Cohen, 1972; Rugman, 1979, Chapter 3; Miller and Pras, 1980).[16] MNEs also enjoy lower levels of risk in the sense relevant to the stock market – financial risk (beta), according to Hughes, Logue, and Sweeney (1975), Thompson (1985), and Michel and Shaked (1986). Kwok and Reeb (2000) added an interesting perspective by showing that multinationality reduces risk (standard deviation of monthly market returns) for firms based in unstable and uncertain national economies while increasing it for those based in stable economic environments.

Other variables related to diversification and risk have also been analyzed. One of those is beta, which embraces both the variance of disturbances and the correlation of the firm's returns with the market factor. No sign is predicted for its relation to multinationality, and indeed opposite-sign results have been reported – negative for Hughes et al. (1975), Thompson (1985), and Michel and Shaked (1986), but positive in the article by Reeb, Kwok, and Baek (1998) which employs a larger data set than its predecessors. Another variable that should reflect the net influence of uncertainty and diversification is the debt ratio (debt/assets). Earlier articles summarized by Chkir and Cosset (2001) found debt ratios negatively related to firms' extent of multinational activity, implying a predominance of variability of disturbances in foreign markets. Burgman (1996) similarly concluded that debt ratios of domestic firms and MNEs are driven by different determinants, and that MNEs' debt ratios are depressed by large uncertainties such as political and exchange-rate risks.

Jacquillat and Solnik (1978) investigated the degree to which large MNEs based in Europe and America can be regarded as "walking mutual funds" that are diversified across national economies. They found that the rates of return on the market values of their firms' equity shares are still quite closely tied

[15] See Rugman (1979), especially Chapters 2 and 4.

[16] Miller and Pras (1980) found that the variability of operating income for U.S. MNEs is negatively related to both their sizes and the numbers of foreign countries in which they have subsidiaries; they also concluded that being diversified among heterogeneous regions offers more stabilization than being in closely similar countries. Oddly, with these influences controlled, they did not find significant stabilization of profits due to the companies' exports and their product-market diversifications in the United States.

to economic conditions in their national home markets, except the MNEs based in the smaller European countries. In general, this evidence supports the hypothesis that the MNE attains appreciable international diversification. However, the diversification might result from investments that were propelled by other motives; whether foreign direct investment yields diversification gains for which shareholders will pay is considered in Chapter 6.

MNEs' productivity and efficiency could be affected by international diversification under congenial conditions, although the counterbalanced forces already noted deny access to any clear predictions. Hitt, Hoskisson, and Kim (1997) is a recent addition to studies of multinationality's relation to accounting profitability (return on assets, equity, and sales). Its authors found that international diversification appears to have an internal optimum, with profits lowered for the firm operating in too many or too few countries. Baek (2004) investigated another performance measure – the firm's productive efficiency (its productivity relative to an estimated frontier defined by the most efficient firms). For a large sample of U.S. firms, he found that efficiency is positively related to international diversification.

Geographic and Product Diversification

Further evidence on MNEs' diversification can be found in specific transactions with potentials for spreading risk. The most diversification should accrue to the MNE that acquires a foreign subsidiary diversified in product line as well as geographical space. If diversification promotes foreign investment, we should find some of this "double diversification" in MNEs' structures. Early surveys (Barlow and Wender, 1955, p. 159) asserted that diversified foreign investment is a rare phenomenon. Caves (1975) and Dubin (1976, Chapter 6) found statistical evidence that MNEs' activities are more diversified among products on their national home ground than in foreign subsidiaries, confirming the impression from surveys (Dunning, 1958, pp. 115–18; Safarian, 1966, p. 211; Saham, 1980, pp. 172–75). Apparently, the extra costs and risks of adding activities abroad look unappetizing to the firm that seeks diversification from whatever source;[17] also, minor related products in the firm's line tend to get made at the home base.

[17] If foreign investment typically had diversification value that offset its specific risks, we should expect MNEs to accept lower expected rates of return on foreign investments than on domestic investments. But survey evidence, such as that of Barlow and Wender (1955, p. 114), points to a higher minimum for foreign investments.

Nonetheless, diversifying in domestic product markets and investing abroad are alternatives for mature companies (Caves, 1975; Wolf, 1977) even though in uncontrolled samples the larger and more mature firms will have expanded in both directions (Pearce, 1993). Also, specifically diversified foreign investments are growing more numerous. Kopits (1979) found that the diversified foreign investment of U.S. MNEs in 1968 was positively related to the extent of R&D activities in the U.S. base industry of the parent (company size and seller concentration were also controlled in this regression analysis but did not prove significant). The result agrees with the hypothesis that a firm's research activities often produce proprietary assets useful outside its base industry; these should lead to international diversification, just as they promote diversification at home (also see Pearce, 1993). In this vein, Hisey and Caves (1985) analyzed a sample of international acquisitions by U.S. companies that could be classified as either related or unrelated diversifications relative to the acquirers' U.S. activities. The unrelated ones are significantly associated with risk-spreading properties, the related ones only weakly with spillovers of proprietary assets among product markets. Kim, Hwang, and Burghers (1993) undertook an elaborate analysis of the means and standard deviations of U.S. MNEs' returns on assets in the 1980s, estimating how each firm's risk/return pattern differs from that of its (U.S.) industry and relating the residuals to the properties of the firms' diversification patterns. They found (consistent with Hisey and Caves) that highly geographically diversified MNEs had apparently located an attractive niche of high returns coupled with low risks. Other groups of MNEs reveal the trade-off normally expected: either low risks and returns (with high unrelated product diversification but low diversification of other types) or high risks and returns (with high related and low unrelated product diversification). Davies, Rondi, and Sembenelli (2001) studied the prevalence of product and geographic diversification of large European Union firms, seeking evidence of either substitution or complementarity in these two types of diversification. Overall, neither could be detected; their most interesting result was that firms based in differentiated product industries tend to diversity in both directions – consistent with the properties of proprietary assets.

Some other hypotheses not covered in this statistical analysis also help to explain MNEs' diversification. MNEs in the United States make a somewhat larger proportion of diversified foreign investments in developing countries than in developed countries. This is probably due to controls imposed by governments on the remittance of profits by MNEs operating within their boundaries; restricted from repatriating its profits, the MNE's best

alternative might be to invest in some diversifying activity within the country. Another explanatory factor is the large wave of conglomerate mergers that took place in the United States in the 1960s and 1970s. Suppose that firm *B*, either a horizontal or a vertical MNE, is acquired by the larger firm *A*. If *A*'s base industry remains the principal activity of the merged firm, *B*'s overseas assets will appear to be a diversified foreign investment of the merged firm. Or if *A* diversifies domestically, whether by merger or otherwise, its diversified domestic division might later sprout a horizontal foreign subsidiary, making the firm as a whole appear (to the statistician) diversified internationally.[18]

1.4. Summary

The existence of the MNE is best explained by identifying it as a multiplant firm that sprawls across national boundaries, then applying the transaction-cost approach to explain why dispersed plants should fall under common ownership and control rather than simply trade with each other (and with other agents) on the open market. This approach is readily applied to the horizontal MNE (its national branches produce largely the same products) because the economies of multiplant operation can be identified with use of the firm's proprietary assets, which suffer many infirmities for trade at arm's length. This hypothesis receives strong support in statistical studies, with regard both to intangible assets and to capabilities possessed by the firm. Foreign investments also take place to augment the investor's proprietary assets via leakage from host countries.

A second major type of MNE is the vertically integrated firm, and several economic models of vertical integration stand ready to explain its existence. Once again, the transaction-cost approach holds a good deal of power because vertical MNEs in the natural-resources sector seem to respond to the difficulties of working out arm's-length contracts in small-numbers situations where each party has a transaction-specific investment at stake. Evading problems of impacted information also seems to explain some vertical foreign investment. The approach also works well to explain the rapid growth of offshore procurement by firms in industrial countries, which involves carrying out labor-intensive stages of production at foreign locations with low

[18] For evidence, see Horst (1974*a*, pp. 110–11). That overseas diversification represents some kind of optimizing global calculation is suggested by Gorecki's finding (1980) that the diversification levels of Canadian domestic firms can be explained by Canadian market variables, whereas the diversification levels in Canada of foreign subsidiaries operating there cannot.