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