1 Statistics, Stylized Facts, and Basic Concepts

1.1 Introduction

The early 1980s saw majors developments in international economics as industrial organization aspects of international trade were integrated into more formal trade theory. Prior to the 1980s, two rather distinct literatures existed. There was general-equilibrium trade theory, which relied almost exclusively on the twin assumptions of constant returns to scale and perfect competition in production. Second, there was a partial-equilibrium literature that considered industrial organization effects of trade, such as the effect of trade barriers on concentration, competition, and production efficiency.

The so-called new trade theory and more recently the literature on "geography and trade" enriched our portfolio of theory by integrating these literature streams. Elements of increasing returns to scale, imperfect competition, and product differentiation were added to the more traditional comparative advantage bases for trade in general-equilibrium models. This new theory complements traditional comparative advantage models, in which trade and gains from trade arise as a consequence of differences between countries. In the new trade theory, trade and gains from trade can arise independently of any pattern of comparative advantage as firms exploit economies of scale and pursue strategies of product differentiation. The literature on geography and trade is a natural extension of this line of research, focusing on how industry agglomeration and regional differentiation can arise endogenously as a consequence of transport costs, market sizes, and the trade policy regime.

The new industrial organization (IO) models were an important step, but they nevertheless remained disjoint from any theory of the

multinational enterprise. In the trade-IO models, a firm is generally synonymous with a plant or production facility; that is, a firm is a nationally owned organization that produces one good in one location. Multiplant production is generally excluded from the analysis. This is potentially troubling. After all, industries characterized by scale economies and imperfect competition are often dominated by multinationals. As a result, the policy and normative analysis that comes out of the new trade theory may be significantly off base. For example, conclusions of the "strategic trade policy" literature are fundamentally bound up with the notion of clearly defined national firms competing via trade with the national champions of other countries. Substantial foreign ownership of domestic production facilities radically alters the policy implications.

The purpose of this book, as I noted earlier, is to incorporate multinational firms into the general-equilibrium theory of trade. The purpose of this chapter is to survey some of the empirical evidence that we have accumulated to date. I have three separate objectives for this statistical abstract. The first is to convince you that the topic is too quantitatively important to disregard in trade theory, as has largely happened to date. The second is to convince you that the old view of direct investment as not fundamentally different from the theory of portfolio capital movements is completely wrong. The third objective is to provide motivation and support for assumptions employed in the theoretical models throughout the book. Section 1.4 introduces a general conceptual framework that is motivated by and draws on the empirical evidence.

For those interested in related theoretical treatments, complementary surveys of the individual elements I am combining here can be found in Beckman and Thisse (1986: location theory), Markusen (1995, 1998a: integrating multinationals into the IO theory of trade), and Fujita, Krugman, and Venables (1999: location theory, trade-IO models and "economic geography"). Caves (1996) and Ekholm (1995) are other sources for extensive references and literature reviews of multinational firms.

In the next two sections, I examine statistical and other data to create a general impression as to what key aspects of multinationals we need to capture in formal models. Section 1.4 presents the outline of a general conceptual framework that I refer to as the knowledge-capital model of the multinational, building on the earlier conceptual framework of Dunning (1977, 1981).

1.2 An Empirical Background

A preliminary note about terminology may be useful. Multinational enterprises (MNE) are firms that engage in foreign direct investment (FDI), defined as investments in which the firm acquires a substantial controlling interest in a foreign firm or sets up a subsidiary in a foreign country. I use the abbreviations MNE and FDI fairly interchangeably.

Horizontal direct investments refers to the foreign production of products and services roughly similar to those the firm produces for its home market. Vertical investments refer to those that geographically fragment the production process by stages of production. This terminology is not clearcut. All horizontal investments generally have some vertical element, in that services such as management, engineering, marketing, and finance are often supplied in one direction, from parents to subsidiaries. But the terms are convenient and in widespread use. So when I refer to horizontal investments or horizontal multinationals, I am referring to firms producing roughly the same final products in multiple countries as just noted, even though foreign plants are supplied with headquarters services. Vertical firms generally produce outputs not produced by the parent-country operation. A parent firm may ship designs and/or intermediate inputs to a foreign assembly plant, for example, and export the final output back to the parent-country market. While the horizontal-vertical distinction is not always clear empirically, it is well defined in the theoretical models that follow throughout the book.

Before plunging into the theory, I offer a much-needed background of stylized facts to provide a context within which to evaluate the theory and, indeed, to understand its origins. It is my view that much of the recent theory is fairly closely tied to the evidence, or at least consistent with it. Consider first some factors identified in aggregate data, which generally have to do with country characteristics, and then consider results found in analyses of industry- and firm-level data. The former have to do with characteristics of parent and host countries, while the latter generally refer to characteristics of multinational versus nonmultinational firms.

Country Characteristics

1. Direct foreign investment has grown rapidly throughout the world, with particularly strong surges in the late 1980s and late 1990s.[1]

2. The developed countries not only account for the overwhelming proportion of outward FDI but are also the major recipients of FDI.

3. Two-way FDI flows are common between pairs of developed countries, even at the industry level.[2]

4. Most FDI appears to be horizontal, at least insofar as most of the output of foreign affiliates is sold in the foreign country.[3]

5. A significant percentage of world trade is now intrafirm trade (about 30%). There is some evidence of complementarity between trade and investment.[4]

6. Little evidence exists that FDI is positively related to differences in capital endowments across countries, or alternatively to differences in the general return to capital. Skilled-labor endowments are strongly and positively related to outward direct investment.

7. Political risk and instability seems to be an important deterrent to inward FDI. Taxes appear to be of second-order importance (e.g., if a given U.S. company is going to invest in Europe, taxes help determine which location is chosen). I am not sufficiently knowledgeable to evaluate conflicting results in the international taxation literature.[5]

In summary, direct investment has been growing rapidly, and the bulk of it is horizontal direct investment among the high-income developed countries.

Firm and Industry Characteristics

1. Large differences exist across industries in the degree to which production and sales are accounted for by multinational firms.[6]

2. Multinationals tend to be important in industries and firms that (a) have high levels of R&D relative to sales, (b) employ large numbers of professional and technical workers as a percentage of their total workforces, (c) produce new and/or technically complex products, and (d) have high levels of product differentiation and advertising.[7]

3. Multinationals tend to be firms in which the value of the firms' intangible assets (roughly, market value less the value of tangible assets such as plant and equipment) is large relative to its market value.[8]

4. Limited evidence suggests that plant-level scale economies are negatively associated with multinationality.[9]

5. There seems to be a threshold size for multinationals, but above that level corporate size is not important. Corporate age is highly correlated with multinationality.[10]

6. There is evidence that FDI is positively related to the existence of trade barriers. Evidence by Brainard (1997) demonstrates that the *share* of foreign affiliate sales in the sum of exports and affiliate sales is positively related to trade barriers and transport costs.[11]

Thus trade barriers and transport costs do cause a substitution effect toward direct investment, although they may reduce the levels of both investment and trade. Carr, Markusen, and Maskus (2001) and Markusen and Maskus (2001, 2002) show that it depends on which country, parent, or host has the trade barriers. Host-country barriers clearly encourage FDI while parent-country barriers (weakly) discourage it. Distance between countries clearly discourages FDI.

In summary, multinationals are important in industries in which intangible, firm-specific assets are important. These assets can generally be characterized as "knowledge capital," ranging from proprietary product or process know-how to reputations and trademarks. Direct investment increases relative to trade (but not necessarily absolutely) as host-country trade barriers increase, but decreases with distance.

1.3 Key Statistics

Tables 1.1–1.8 present some statistics to back up and reinforce many of the points of the previous section. I refer to many of these at various points throughout the book to support certain theoretical assumptions.

Table 1.1 presents statistics on the growth of multinational activity over a fourteen-year period and compares them to figures on the growth of GDP, fixed capital, and trade in goods and nonfactor services. The top five rows show different measures of growth in multinational activity. Royalties and fees probably include payments between unaffiliated firms, which is not clear from the World Investment Report, but in any case they are payments for producer services much like that which is transferred within multinationals.

Trade has grown faster than GDP, which is a well-known statistic quoted by authors in the industrial organization approach to trade literature. Less well known, affiliate activity has grown much faster than GDP, capital stocks, and trade. Affiliate sales have generally grown significantly more slowly than FDI stocks and flows, and that is a

Table 1.1
Growth in FDI

	Annual growth rate (%), all countries		
	1986–1990	1991–1995	1996–1999
FDI inflows	24.7	20.0	31.9
FDI stocks	18.2	9.4	16.2
Sales of foreign affiliates	15.8	10.4	11.5
Gross product of foreign affiliates	16.4	7.1	15.3
Royalties and fees receipts	22.0	14.2	3.9
GDP at factor cost	11.7	6.3	0.6
Gross fixed capital formation	13.5	5.9	−1.4
Exports of goods and nonfactor services	15.0	9.5	1.5

Source: UNCTAD World Investment Report (2000, and earlier years).

puzzle in itself. The models in this book are addressed more closely to affiliate output and sales than to investment stocks. The mid-1990s was a slower period for FDI and sales relative to GDP and trade, but the former rebounded strongly in the late 1990s. The numbers in table 1.1 are key motivating statistics for this book and provide support for the notion that it is important to spend more time and effort trying to understand multinationals than remaining fixated on trade in international microeconomics.

Table 1.2 gives statistics on the sources and recipients of new direct investment flows, dividing the world into Developed, Developing, and Central and Eastern Europe (CEE) (this is the breakdown in the UNCTAD World Investment Report, not my choice). Not surprisingly, the developed countries are the major source of outward ("out") investment, but perhaps less well known, they are the major recipients ("in") as well. There is a period in the mid-1990s where this was less true, but a look at more refined statistics indicates that the boom in investment in the developing countries during the mid-1990s was almost entirely accounted for by the opening of China. The apparent boom in outward investment from developing countries in the same period was almost entirely accounted for by Taiwanese, American, and other firms funneling their investments into China through Hong Kong subsidiaries, so the funds appeared to be coming from a developing country. This investment boom in China significantly weakened in 1998–1999, and the world may return to a more historical pattern

Table 1.2
FDI inflows and outflows, share in total

	Developed		Developing		CEE	
Year	in	out	in	out	in	out
1983–1987	76	95	24	5	0	0
1988–1992	78	93	21	7	1	0
1993	62	85	35	15	3	0
1994	59	83	39	17	3	0
1995	65	85	32	15	4	0
1996	58	85	38	15	3	0
1997	58	86	38	14	4	1
1998	71	95	26	5	3	0
1999	74	91	24	8	1	0

Source: UNCTAD World Investment Report (2000, and earlier years).

in which the bulk of funds flows not only from but to developed countries.

The statistics in table 1.2 are a major challenge to theory that I alluded to in the preface and earlier in this chapter. Theory must be able to explain why so much FDI flows among the high-income developed countries. Clearly, a theoretical model in which FDI only flows from capital rich to capital poor countries should be dismissed out of hand.

Table 1.3 continues on somewhat the same theme. Table 1.3 presents data on stocks rather than flows, however. I do not know why there is a discrepancy between world inward and outward stocks, whether this is just a statistical discrepancy or not. But these stocks have clearly grown steadily and significantly faster than GDP, as the flow data in table 1.1 suggest. Again, note that the inward numbers for the developed countries are almost equal to the averages for the world as a whole.

Table 1.3 notes that developing countries are net recipients of direct investment, which is not surprising and consistent with the intuition one would get from a model of portfolio capital flows. However, table 1.3 also breaks out the "least-developed countries." This is a United Nations definition that includes forty-eight countries. Note that the least-developed countries have inward stocks that are much smaller than the world average. These are the world's most capital-scarce economies and, for whatever reasons, they do not attract much FDI.

Table 1.3
Ratio of inward and outward FDI stock to gross domestic product

	1980	1985	1990	1995	1998
World					
inward	4.9	6.7	8.6	9.6	13.7
outward	5.4	6.4	8.6	10.2	14.1
Developed countries					
inward	4.7	6.1	8.3	8.8	12.1
outward	6.4	7.5	9.8	11.7	16.4
Developing countries					
inward	5.4	9.1	10.5	13.4	20.0
outward	0.9	1.6	2.6	4.9	6.7
Least developed countries					
inward	1.8	3.4	4.4	6.9	7.4
outward	0.7	2.7	1.0	1.1	1.9
United States					
inward	3.1	4.6	7.1	7.6	9.5
outward	8.1	6.2	7.8	9.9	11.5
The Netherlands					
inward	11.1	19.5	25.9	31.5	48.0
outward	24.5	37.3	38.4	45.4	68.9
Sweden					
inward	2.9	5.0	13.4	19.1	21.5
outward	3.0	10.7	21.5	31.6	41.3
Switzerland					
inward	8.4	10.8	15.0	18.6	26.5
outward	21.1	27.0	28.9	46.3	69.1
United Kingdom					
inward	11.7	14.0	20.8	18.0	23.2
outward	15.0	21.9	23.4	27.4	35.9

Source: UNCTAD World Investment Report (2000). "Least Developed Countries" is a UN definition consisting of 48 countries.

In the bottom half of table 1.3, I present data on some smaller to moderate-sized high-income countries and include the United States for comparison. The Netherlands, Sweden, Switzerland, and the United Kingdom are all major sources of outward direct investment. But these countries are also major recipients of inward direct investment. The United States, by comparison, is moving toward a position in which inward and outward stocks are converging. These statistics figure importantly in the theory chapters that follow.

Table 1.4
Share of inward world FDI stock/share of world GDP

	Developed countries	Developing countries	Least developed countries
1980	0.96	1.10	0.37
1985	0.91	1.36	0.51
1990	0.97	1.22	0.51
1995	0.92	1.40	0.72
1998	0.88	1.46	0.54

Source: UNCTAD World Investment Report (2000); Zhang and Markusen (1999).

Table 1.4 presents data from the top of table 1.3 on inward investment stocks in a somewhat different form. The statistics in table 1.4 give a group's share of world inward FDI stock divided by its share of world GDP. Once again, the developed countries are major recipients or hosts of FDI. Developing countries have a larger share of inward investment than their share of world GDP. But the least-developed countries have a much smaller share of FDI than their share of income, and a much smaller share of FDI relative to income than the developed countries.

There are, of course, some obvious reasons why the least-developed countries might attract so little FDI. These include the absence of all forms of infrastructure, including physical, institutional, and legal, extending to the absence of rule of law. But there are other reasons as well, and these will be discussed in subsequent chapters. They include demand-side reasons, such as the nature of the products produced by multinationals, and cost-side reasons, such as the need for skilled labor in the production process. Both together suggest that it is in the nature of what multinationals do that their products and processes are not well suited to very poor countries.

Table 1.5 presents some data on labor-force composition, relative wages, and inward and outward FDI. Relative wage data is hard to get, but the GTAP data set has figures for a limited set of countries as shown. For the seven developed and fourteen developing countries noted, the former have a much larger proportion of the labor force classified as skilled, and a much lower relative wage for skilled labor. These are crucial stylized facts that are exploited in the theoretical assumptions later in this book. Table 1.5 also shows the FDI stock data as a percentage of GDP for comparison. The skilled-labor-abundant countries are strong outward investors but also very significant

Table 1.5
Selected statistics, unweighted averages, 1995

	Skilled worker as % of labor force	Ratio of skilled to unskilled wage	Inward FDI as % of GDP	Outward FDI as % of GDP
Seven developed countries (GTAP)	26.0	1.81	13.4	16.5
All developed			8.8	11.7
Fourteen developing countries (GTAP)	10.8	3.54	20.5	5.1
All developing			13.4	4.9

Sources: FDI data from UNCTAD World Investment Reports (1996, 1997, and 2000). Labor force and wage data from GTAP data set, 1995.
Note: Seven developed countries are United States, Canada, Japan, Denmark, Germany, Great Britain, and Sweden. Fourteen developing countries are Mexico, Korea, Singapore, Philippines, Malaysia, Thailand, Indonesia, China, Brazil, Chile, Turkey, Venezuela, Columbia, and Sri Lanka.

recipients of FDI. The skilled-labor-scarce countries are significant recipients of investment but small outward investors.

Table 1.6 continues along similar lines but adds a twist in considering market size along with per capita income. Developing countries are first divided into groups on the basis of per capita income, and then each income class is divided into small and large countries. The right-hand column of table 1.6 then gives FDI per capita. First, clear evidence is visible that richer countries have more inward investment per person. This may be because there is a high income elasticity of demand for the products multinationals produce and/or the need for skilled labor and related factors in production. But there may well be some reverse causality as well, in that more inward FDI may help generate higher per capita incomes. Second, larger countries in a given per capita income class receive significantly more inward investment per capita (except for the very poorest class). This suggests to me that local sales are quite important in overall multinational activity.

Table 1.7 explicitly considers the role of the local market versus production for export and the role of imports from the parent firm. The data is all bilateral with the United States, giving U.S. parents' trade with their foreign affiliates (outward data) and foreign firms' trade with their U.S. affiliates (inward data). The country abbreviations should be obvious, except for OAP which is Other Asia-Pacific, and LAT which is Latin America.

Table 1.6
Inward FDI flows and their links with GDP per capita and national income of developing countries in 1993

Country groups by GDP per capita (US$)	Average FDI per capita (US$)	Country groups by country size in GDP (US$, millions)	Average FDI per capita (US$)
>5000	226.89	>55000	242.20
		<49000	53.83
2500–5000	45.30	>31000	45.73
		<17000	32.30
1200–2500	33.02	>10000	33.43
		<9600	30.60
600–1200	10.06	>10000	10.86
		<9300	2.59
300–600	6.56	>4800	6.91
		<3700	3.68
<300	0.63	>2000	0.34
		<1500	2.47

Sources: FDI data are from International Monetary Fund (1995), *Balance of Payments Statistics Yearbook 1995*. GDP data are from International Monetary Fund (1995), *International Financial Statistics Yearbook 1995*. See also Zhang and Markusen (1999) for more detailed definitions and discussions of these data.

The top panel gives 1997 data, and one sees that, for all countries, numbers in the left-hand column for all U.S. partner countries are quite modest, not exceeding 15 percent of total affiliate sales. Affiliates are, on the whole, clearly not just assembly operations for export back to the home country, nor do they source a major part of their inputs from home-country imports. U.S. imports from the foreign affiliates of U.S. corporations are quite small as a percentage of total affiliate sales, and similarly for U.S. exports by affiliates of foreign corporations to their foreign parents. The numbers for OAP and LAT are somewhat higher. But still, only a relatively small proportion of output is shipped back to the United States. U.S. affiliates in these countries are not primarily assembly or other "vertical" operations producing for export. The largest numbers for 1997 are for U.S. imports from affiliates in OAP (27% of sales) and LAT (26% of sales). This reinforces the point of table 1.6, that the local host-country market is, on the whole, quite important for multinational firms.

In the bottom row of the 1997 and 1987 panels, I present Grubel-Lloyd indices of cross- or intra-industry affiliates activity. The intra-industry affiliate sales index (IIAS) is defined as follows. Let AS_{ij} denote affiliate sales by affiliates in country i of country j parent firms.

Table 1.7
Parent-affiliate trade as a proportion of total affiliate sales, 1997 and 1987 (foreign affiliates of U.S. firms and U.S. affiliates of foreign firms)

	Countries										
	ALL	CAN	FRA	GER	NET	SWI	UK	AUS	JAP	OAP	LAT
1997 total manufacturing											
Outward data											
U.S. exports to affiliates	0.14	0.41	0.05	0.05	0.07	0.04	0.06	0.09	0.12	0.15	0.23
U.S. imports from affiliates	0.15	0.42	0.04	0.03	0.04	0.06	0.05	0.04	0.04	0.27	0.26
Inward data											
U.S. exports shipped by affiliates of foreign firms	0.10	0.10	0.13	0.13	0.07	0.11	0.10	0.09	0.11	0.12	0.07
U.S. imports shipped by foreign parents to their U.S. affiliates	0.15	0.14	0.16	0.15	0.13	0.08	0.09	0.10	0.23	0.11	0.22
Grubel-Lloyd indices	82.92	52.04	88.00	80.66	82.23	30.68	98.93	71.78	45.95	23.57	16.62
1987 total manufacturing											
Outward data											
U.S. exports to affiliates	0.14	0.38	0.04	0.03	0.08	0.04	0.05	0.08	0.07	0.24	0.18
U.S. imports from affiliates	0.12	0.33	0.03	0.03	0.02	0.04	0.05	na	0.07	na	0.17
Inward data											
U.S. exports shipped by affiliates of foreign firms	0.07	0.09	0.06	0.09	0.05	0.04	0.05	0.03	0.07	0.05	0.09
U.S. imports shipped by foreign parents to their U.S. affiliates	0.11	0.10	0.10	0.14	0.10	0.09	0.07	0.08	0.27	0.09	na
Grubel-Lloyd indices	73.38	72.09	72.48	67.01	94.88	22.73	91.02	47.92	93.51	47.42	28.73

Source: Calculated from Bureau of Economic Analysis data.

The Grubel-Lloyd index applied to cross-country affiliate sales is

$$IIAS_{ij} = \left[1 - \frac{|AS_{ij} - AS_{ji}|}{AS_{ij} + AS_{ji}}\right] * 100.$$

The IIAS index ranges from a low of zero, when affiliate activity is one way only, to a value of one hundred when affiliate sales are perfectly balanced.

The Grubel-Lloyd indices in table 1.7 are high for developed-country partners of the U.S. except for Japan (1997: 46%) and Switzerland (1997: 31%), although even these are moderately high. Cross- or intra-industry penetration of each developed country's firms in the other market is high. The numbers of OAP and LAT are significantly smaller as one might expect. Once again, we see that much direct investment is among the developed countries, rather than a one-way trip from developed to developing countries.

Table 1.8 looks at firm characteristics. Data is very scarce on these issues, but I have assembled some data from the different sources noted. The top numbers give a proxy for skill level, which is compensation per employee. If we accept this as a rough proxy, then parents are skilled-labor-intensive relative to affiliates, but there is not a huge difference for developed-country affiliates. I have included the GTAP figure for the fourteen developing countries from table 1.5 for comparison with the number for affiliates in developing countries. The numbers indicate that developing-country affiliates pay an average wage that is close to the average earnings of skilled workers in developing countries. This in turn suggests that affiliates are skilled-labor-intensive relative to the developing host countries as a whole. I use these stylized facts in assumptions about factor intensities later in the book.

The second set of numbers gives the share of nonproduction workers in total employment of parents and affiliate. While nonproduction workers are at best a crude proxy for skilled or "knowledge workers," they again suggest that skilled or knowledge workers are somewhat, but not completely, concentrated in the parent operations. These numbers will be used to justify some assumptions on factor intensities used later in the book.

Assets per employee, the third set of numbers, serves as a proxy for physical capital and perhaps intangible capital (I am not sure) in parents and affiliates. Parents and affiliates in developed countries are

Table 1.8
Characteristics of U.S. multinational corporations in manufacturing (1989 data unless otherwise indicated)

Skill level (compensation per employee $000, World Investment Report 1993)	
Parents	38.9
Affiliates	25.2
Developed countries	33.3
Developing countries	9.5
Fourteen developing countries from table 1.5	
Average earnings of skilled workers	9.8
(GTAP data set)	
Share of nonproduction employees in total employment 1982 (Slaughter 2000)	
Parents	0.54
Affiliates	0.42
Assets per employee ($000 per employee, World Investment Report 1993)	
Parents	186
Affiliates	114
Developed countries	147
Developing countries	52
R&D expenditures as a percentage of sales (World Investment Report 1993)	
Parents	3.33
Affiliates	1.12
Developed countries	1.27
Developing countries	0.30
All U.S. R&D performing manufacturing firms	3.20
R&D employment as a percentage of total employment (World Investment Report 1993)	
Parents	5.46
Affiliates	2.42
All U.S. R&D performing manufacturing firms	4.90

Sources: UNCTAD World Investment Report (1993), Slaughter (2000), and 1995 GTAP data set converted to 1989 US$. All manufacturing from NSF data, Survey of Industrial Research and Development (1991).

not wildly different, but affiliates in developing countries are much less "asset intensive."

The final two sets of numbers are measures of R&D intensity for parents and affiliates, and I also include figures for all U.S. R&D performing manufacturing firms. Parents are significantly more R&D-intensive than their affiliates, including affiliates in developed countries. Parents are slightly more R&D-intensive than all R&D performing firms in the United States, but I do not know what share of manufacturing does not do R&D, and therefore do not know the R&D intensity of multinational parents relative to all of U.S. manufacturing. Obviously, the inclusion of all manufacturing firms would make the multinationals look considerably more R&D-intensive than just considering R&D-producing firms. In any case, these numbers complement statistics cited in the previous section that multinational firms have a high value of intangible assets, suggesting that multinationals are relatively intensive in knowledge-based assets.

1.4 A Knowledge-Capital Approach

A typical point of departure for theory has been the logical premise that firms incur significant costs of doing business abroad relative to domestic firms in those countries. Therefore, for a firm to become a multinational, it must have offsetting advantages. A limited but very useful organizing framework for inquiring into the nature of these advantages was proposed by John Dunning (1977, 1981). Dunning proposed that three conditions are needed for firms to have a strong incentive to undertake direct foreign investments.

1. Ownership advantage: The firm must have a product or a production process such that the firm enjoys some market power advantage in foreign markets.

2. Location advantage: The firm must have a reason to want to locate production abroad rather than concentrate it in the home country, especially if there are scale economies at the plant level.

3. Internalization advantage: The firm must have a reason to want to exploit its ownership advantage internally, rather than license or sell its product/process to a foreign firm.

An important task of theory is to connect these ideas with the firm (technology) and country characteristics in a consistent way. This is something that was undertaken in a number of papers including

Markusen (1984, 1997), Ethier (1986), Helpman (1984, 1985), Horstmann and Markusen (1987a,b; 1992), Brainard (1993a), Ethier and Markusen (1996), and Markusen and Venables (1998, 2000). I will refer to a synthesis of several approaches as the "knowledge-capital" model, although I note that this is not a widely used term.

Consider first ownership advantages. Evidence indicates that multinationals are related to R&D, marketing, scientific and technical workers, product newness and complexity, and product differentiation. This suggests that multinationals are firms that are intensive in the use of knowledge capital. This is a broad term that includes the human capital of the employees, patents, blueprints, procedures, and other proprietary knowledge, and finally marketing assets such as trademarks, reputations, and brand names.

The crucial question then is why should knowledge capital be associated with multinationals while physical capital is not? I have suggested that the answer lies in three features of knowledge capital, although these should be referred to as assumptions pending econometric support. First, the services of knowledge capital can be easily transported to foreign production facilities, at least relative to the services of physical capital. Engineers and managers can visit multiple production facilities with some ease (although stationing them abroad is costly) and communicate with them in a low-cost fashion via telephone, fax, and electronic mail. This property of knowledge capital is important to firms making either horizontal or vertical investments.

Second, knowledge-based assets are skilled-labor-intensive relative to production. This creates a motive for the geographical fragmentation of production and vertical multinationals. Skilled-labor-intensive "headquarters" activities such as R&D and management should be located where skilled labor is abundant and relatively cheap while production may be located in less-skilled-labor-abundant countries and/or in large markets.

The third property of knowledge capital that leads to the association of multinationals with knowledge capital is the fact that knowledge capital often has a joint-input or "public-good" property within the firm. Blueprints, chemical formulae, or even reputation capital may be very costly to produce, but once they are created, they can be supplied at relatively low cost to foreign production facilities without reducing the value or productivity of those assets in existing facilities. The blueprint, for example, can yield a flow of services in multiple locations simultaneously. This property of knowledge capital, which does not

characterize physical capital, is particularly important to horizontal multinationals. But it may be quite important to vertical multinationals as well, insofar as the "blueprint" indicates exactly how the geographically fragmented activities, components, and products must fit and work together. In the knowledge-capital framework, multinationals are then exporters of knowledge-based services: managerial and engineering services, financial services, reputations and trademarks.

The sources of location advantages are varied, primarily because they can differ between horizontal and vertical firms. Consider horizontal firms that produce the same goods and services in each of several locations. Given the existence of plant-level scale economies, there are two principal sources of location advantages in a particular market. The first is the existence of trade costs between that market and the MNE's home country, in the form of transport costs, tariffs, and quotas, and more intangible proximity advantages. Indeed, if trade costs were truly zero, production would be concentrated in a single location (again, assuming plant-level scale economies) with other locations served by exports. That is, some sort of trade cost seems to be a necessary condition for horizontal multinationals to exist. The second source of location advantage, again following from the existence of plant-level scale economies, is a large market in the potential host country. If that market is very small, it does not pay for a firm to establish a local production facility, and the firm will instead service that market by exports.

The sources of location advantage for vertical multinationals are somewhat different. Suppose, for example, that a particular MNE exports the services of its knowledge capital and perhaps other intermediate inputs to a foreign production facility for final assembly and shipment back to the MNE's home country. This type of investment is likely to be encouraged by low trade costs rather than by high trade costs. The most logical situation in which this type of fragmentation arises is when the stages of production have different factor intensities and the countries have different relative factor endowments. Then, for example, skilled-labor-intensive R&D and intermediate goods should be produced in the skilled-labor-abundant country, and less-skilled-labor final assembly should be done in a country with low-wage unskilled labor. Fragmentation arises to exploit factor-price differences across countries.

Internalization advantages are the most abstract of the three. The topic quickly gets into fundamental issues such as what is a firm, and

why and how agency problems might be better solved within a firm rather than through an arm's-length arrangement with a licensee or contractor. Basically, it is my view that internalization advantages often arise from the same joint-input, public-goods property of knowledge that creates ownership advantages. The property of knowledge that makes it easily transferred to foreign locations also makes it easily dissipated. Firms transfer knowledge internally in order to maintain the value of assets and prevent asset dissipation. Licensees can easily absorb the knowledge capital and then defect from the firm or ruin the firm's reputation for short-run profit. Internalization models will be the focus of the last three chapters of the book (again, see Markusen 1995 for a survey).

This section can be summarized as follows.

Ownership advantages: Arise from knowledge capital, which (a) can be easily transported or transferred to foreign production facilities and (b) has a joint-input property across the different production facilities.

Location advantages: For horizontal firms, location advantages arise when the host-country market is large and when trade costs (broadly defined) are moderate to high. For vertical firms, location advantages arise when trade costs are low, stages of production differ in factor intensities, and countries differ significantly in relative factor endowments.

Internalization advantages: Although there are many facets to this issue, internalization advantages can arise from the same joint-input characteristic of knowledge capital that creates ownership advantages. Transferring knowledge-based assets through arm's-length market mechanisms runs the risk of asset dissipation.

1.5 Summary

The stylized facts and statistics presented in this chapter lay down a network of facts that need explanation. The chapter also outlines a rough idea, referred to as the knowledge-capital approach, that provides an organizing theoretical framework. Development of the formal theory is the task for chapters 2–9. Chapter 7 provides a more rigorous statement as to exactly what I mean by the knowledge-capital model, since we will not need the full model until that point. I then subject the theory to formal econometric testing in chapters 10–12.