

from: Klaus Meyer (1998): *Direct Investment in Economies in Transition*, Cheltenham: Elgar, p. 59-59-79.

3. Determinants of Direct Investment: A Review of the Literature

A comprehensive review of the literature on direct foreign investment (DFI) invariably becomes a *tour d'horizon* of the field of International Business. Disciplines from Economics and Finance to Strategic Management, Marketing and Organizational Behaviour have contributed to the present understanding of DFI. This chapter reviews the literature on DFI with focus on determinants of DFI.¹ It considers DFI at different levels of aggregation, including macroeconomic flows of DFI-capital, industry and firm level analysis, as well as individual decisions by firms.

Early research analysed DFI as a financial flow. As researchers recognize the specific characteristics of direct, rather than portfolio, investment, they focus on three issues: the location of production, the sources of firm-specific advantages, and the reasons for integrating different business units in one firm. John Dunning's OLI paradigm incorporates these three issues. It is now the most common analytical tool for the determinants of DFI and it is applied in this study.

The limitation of the framework is its ability to explain dynamic processes. Therefore, this review pays special attention to recent advances of economic theory addressing the dynamics of DFI. Dynamic models focus on particular types or aspects of DFI and thus are less general than the OLI paradigm. The most familiar dynamic approach is that of the internationalization process models based on the work of the Uppsala school in the 1970s. Recent advances include the rediscovery of economic geography in the work of Paul Krugman and Michael Porter, the integration of MNE into models of international trade by James Markusen, Elhanan Helpman and, again, Paul Krugman, as well as the game-theoretic analysis by the Leuven school. The appendix presents another dynamic approach to DFI, the developmental model. It relates economic development of a country to the characteristics of DFI outflows.

Table 3.1 summarizes economic theories by their level of analysis and whether they focus on static or dynamic analysis. This categorization serves as orientation only. In some cases dynamic theories have developed from static theories, as is the case with the developmental model. In other cases the pairs reflect contrasting views. Internalization and internationalization theories account for some of the most lively encounters at academic conferences. Table 3.2 lists the multitude of theories with their main analytical concepts, original contributors and recent reviews or extensions.

Table 3.1 Theories of direct foreign investment

Unit of analysis	Static analysis	Dynamic analysis
Financial flow of DFI	Capital markets approach, Macroeconometric analysis	Exchange rate analysis
Location of production	Theory of location (L), Institutional analysis	Developmental model, Economic geography
Firms and competition	Resource based view, Ownership advantages (O)	New international trade theory, Game theory
The scope of the firms	Internalization theory (I)	Internationalization process models

DFI AS CAPITAL FLOWS

On an aggregate level, DFI is analysed as a flow of capital between countries. Theoretical research on this level has evolved from financial market analysis and the aggregation of microeconomic theory. Recent research focuses on dynamics of DFI flows which includes exchange rate effects. Available aggregate data permits extensive empirical analysis which is not feasible at lower levels of aggregation.

Capital Markets Approaches

The first response by economists to the emergence of DFI was to observe the new phenomenon 'through the filters least disturbing to reigning paradigms of the profession' [Vernon 1994: 138]. Considering Thomas Kuhn's 'The Structure of Scientific Revolutions' [1962], Vernon [1994] finds it all but surprising that capital market approaches to DFI have been dominant in the 1960s.

The basic premise is that MNEs face differentials in international capital rents and use DFI to overcome barriers to international capital flows. They finance themselves in countries with a relatively high capital endowment and hence lower interest rates. They invest in countries with a relatively low capital endowment and high capital costs. DFI serves as international capital arbitrage. In this framework, international return differentials determine DFI stocks whereas changes in relative return determine DFI flows. As DFI also transfers other resources than capital, these resources also have to yield a higher return abroad to make DFI profitable. This differential rate of return hypothesis has been analysed empirically but has often been insufficient in explaining DFI [see Agarwal 1980 for review]. It can, however, explain DFI in the nineteenth century due to high transaction costs in capital markets at that time [Hennart 1991].

Table 3.2 Core concepts and major contributors

Theory	Early contributors	Recent work and reviews
Capital markets approach	Aliber 1970 Agmon & Lessard 1977	--
Exchange rate analysis	Logue & Willet 1977 Batra & Hadar 1979	Froot & Stein 1991 Kogut & Kulatilaka 1996
Macroeconometric analysis	Scaperlanda & Maurer 1969	Clegg 1995
Theory of location	Mundell 1957	Dunning 1993
Institutional analysis	Kobrin 1987 Guisinger et al. 1985	Stopford & Strange 1991 Loree & Guisinger 1995
Developmental cycles	Vernon 1966 Kojima & Ozawa 1984 Dunning 1986	Ozawa 1992 Narula 1995 <i>see appendix 3.1</i>
Resource based view	Hymer 1960/1976 Kindleberger 1969	Yamin 1991 Dunning 1993
Industrial organisat./ game theory	Knickerbocker 1973 Dixit 1980 Krugman 1983	Markusen 1995
Internalization theory	Caves 1971 McManus 1972 Buckley & Casson 1976	Dunning 1993 Casson 1995 <i>see chapter 4</i>
Economic geography	Marshall 1920 Krugman 1991	Krugman & Venables 1994, Malmberg, Sölvell & Zander 1996
Internationalization process model	Johanson & Wiedersheim-Paul 1975 Luostarinen 1979	Johanson & Vahlne 1990 Nordström 1991 Andersson 1993
Eclectic paradigm	Dunning 1977	Dunning 1993, 1995

Aliber's [1970] widely cited hypothesis of optimal currency areas assumes that MNEs can finance themselves in hard currency countries. They earn a 'currency premium' by utilizing the interest differential between hard currency and weak currency countries because their creditors do not recognize the risk of devaluation associated with DFI in weak currency areas. Therefore they pay lower capital costs than competitors in local markets, whose capital costs are increased by a risk premium to compensate the creditor for the expected devaluation.

The financial market approach became more comprehensive when the trade-off between yield and risk was incorporated. Firms maximizing yield and minimizing risk diversify their investment portfolio by international investments, as do financial investors. If the systematic risk profiles of the home and foreign markets are less than perfectly correlated, then the risk of an internationally diversified portfolio is lower than that of a purely national portfolio. The capital asset pricing model has been extended to become an international asset pricing model. It can explain diversified DFI as a reaction to barriers and costs pertaining to international portfolio capital flows. By lowering these barriers, MNEs contribute to the integration of international capital markets [Agmon and Lessard 1977, Errunza and Senbet 1984]. For private investors, investing in an MNE becomes an alternative to investing in an international investment fund.

The empirical support for the ability of portfolio investment models to explain direct investment is weak [see Agarwal 1980, Stehn 1992 for reviews]. In recent years, the international liberalization of financial markets has made this motive for DFI increasingly irrelevant with respect to investment between industrialized countries which accounts for most DFI.

Dynamic Macroeconomic Analysis

DFI flows are a function of firms' desired capital stock in given foreign locations, according to their long-term plans. Individual investment decisions determine the timing of a given DFI project. This timing is sensitive to changes, anticipated changes and volatility of major environmental variables, as well as uncertainty. DFI flows react very sensitively because generally they are irreversible. Sunk costs are high due to plant-specific investment, personnel recruitment and training, market research, and negotiations with foreign partners and governments.

Changes in the environment can create temporary cycles of DFI flows as MNEs adjust to new levels of desired foreign holdings. Such changes arise in tax and tariff policies, innovations in corporate finance markets, liberalization of service sectors, or privatization processes. Once the finite number of firms which can make use of new business opportunities have invested, DFI flows drop to their previous level - at a higher level of stock. The reaction to change may involve substantial adjustment costs. Implementation lags may drive a temporary wedge between desired capital stock abroad and the actual capital invested. Also anticipated changes of for example investment incentives, can lead to a rush or delay of DFI and thus cause cycles.

On the other hand, due to so-called 'hysteresis effects' [Dixit 1990, Pindyck 1991], temporary influences such as taxation and exchange rate revaluations, can have long-run implications for the permanent stock of DFI even though the temporary influence has long disappeared. An enterprise investing during an incentive program or a favourable exchange rate constellation will not necessarily withdraw if these cost factors become less favourable, such that the country in question would not attract similar projects again. Divestment decisions depend on expected future cash flow only and ignore sunk costs. Therefore volatile exchange rates and exchange rate expectations will induce investment flows that follow different cyclical paths as increasing investment during a devaluation is not matched by an equivalent divestment during a revaluation. Baldwin and Krugman [1989] made this argument for international trade, showing how entry and exit decisions during a temporary shock can lead to a different equilibrium after the shock although the cause of the shock has been removed. Kogut and Kulatilaka [1996] apply this approach to DFI.

Exchange Rate Analysis

The issues arising from volatility and uncertainty of environmental variables have been discussed in most detail for exchange rates. In perfect capital markets the revaluation of exchange rates does not affect investment flows as both domestic and foreign investors have access to the same financial markets. Three independent lines of arguments have been made why this may not be so.

A devaluation of foreign currency reduces the share of foreign assets in an investor's portfolio. With an unchanged risk evaluation, the share of foreign assets should remain constant. Thus a 'portfolio rebalance effect' induces selling of domestic and buying of foreign assets [Logue and Willet 1977]. Froot and Stein [1991] reach a similar conclusion by assuming imperfect capital markets with information asymmetries: since investors are better informed than bankers on a given project, financing will always require a contribution of an equity share from their own wealth. If a temporary devaluation changes the value of the investors' wealth in foreign currency, this 'wealth effect' will enhance the ability to invest abroad: the same funds buy more foreign assets.²

Third, only changes in real exchange rates affect decisions on location of production. Changes in nominal exchange rates influence only the timing of DFI, not the general trend. A real devaluation of foreign currency reduces relative foreign labour costs and thus leads to more DFI. Local production replaces exports from the home base [Kohlhagen 1977], possibly smoothly, as the adjustment to desired capital stock is lagged [Goldsbrough 1979]. However, this result does not necessarily hold if constant marginal cost functions are not assumed to be constant [Batra and Hadar 1979]. Empirical tests of the influence of the wealth effect versus the labour cost argument on DFI into the US found support

for the wealth effect [Klein and Rosengren 1994].

Exchange rate risk can be hedged in financial markets if it is short- or medium-run and in a commonly traded currency. In this case it 'only' causes transaction costs (TC). To reduce TC and to minimize unhedgeable risk the firm may change its investment decisions. It can reduce 'exposure', that is the net cash flow in foreign currency. Models of MNEs with exchange risk generally find a positive relationship between DFI and exchange risk, as DFI replaces exports. The reason for this relationship is that, when engaged in DFI, only repatriated profits are exposed to exchange risk. As regards exports, all sales revenues are received in foreign currency while costs arise in domestic currency [Itagaki 1981, Batra and Hadar 1979, Calderón-Rossel 1985]. Cushman [1985, 1988] empirically supports this line of argument. These results contradict pure portfolio models [Hartman 1979] that predict a negative correlation.

Exchange risk furthermore induces firms to invest in more locations to be able to react flexibly to changes in real local costs of production [see Kogut and Kulatilaka 1996, Aizenman 1992]. Also, an initial investment can serve as a platform for subsequent entry where the timing of entry is triggered by movements in the real exchange rates [Kogut and Chang 1996].

Macroeconometric Analysis

The diversity of DFI is a major obstacle to macro-level analysis, both to modelling and to empirical research. Macroeconometric research combines financial data, such as those discussed above, with microeconomic determinants of DFI derived from firm-level theoretical work. The most commonly tested variables include market size, market growth, factor costs, trade barriers, as well as interest rate differentials and exchange rate movements. This research has used three different kinds of data set: cross-country inflows of DFI from a source country, cross-industry data of DFI inflows, or time-series of DFI into a particular country. The time-series approach is most popular as capital flow data is readily available and econometric techniques are well-developed. The dependent variable is DFI as measured in the balance of payments or changes in recorded DFI stock.

Market size and growth are considered in all these studies, as penetration of foreign markets is a major motive for DFI. GDP and the change in GDP are the most common proxies and generally significant.³ Low labour costs are generally presumed to attract DFI as they reduce costs of production. However, it is difficult to show this empirically because low labour costs are associated with low income and thus low local demand.⁴

Exchange rate effects are tested but empirical results are on the whole as inconclusive as the theoretical research [for example Cushman 1985, Stevens 1993, Clegg 1995]. Relative interest rates reflecting the cost of borrowing were however frequently significant [for example Cushman 1985, 1988, Pain 1993 and Clegg 1995, but not Culem 1988]. Institutional aspects can often only be captured by dummy variables that indicate the presence of condition. For instance, dummies for tariff discrimination and the time of membership in the European Union were significant [Culem 1988 and Bajo-Rubio and Sosvilla-Rivero 1994]. Economic risk of macroeconomic instability is related to the inflation rate and reduces investment [Bajo-Rubio and Sosvilla-Rivero 1994].

THE LOCATION OF PRODUCTION

Theory of Location

The traditional basis for analysis of international economic activity in the real (rather than monetary) sector is the neoclassical theory of international trade. However, it provides however no framework for explaining the existence or development of DFI. It explains international trade in terms of comparative advantages of the participating countries based on the assumption of perfect competition. Certain resources or factors are immobile, production functions and consumer preferences are identical, and specialization is incomplete. Countries

specialize in production which uses factors of production that are relatively abundant. Trade leads to an equalization of factor prices (Heckscher-Ohlin Theorem). In this model the assumption of perfect competition eliminates MNEs.

Early attempts to model DFI use a modified factor endowments-based model of international trade. An early popular hypothesis was Mundell's [1957] 'factor endowment theory' that showed that, under certain assumptions, capital flows can substitute trade if barriers prevent the free flow of goods. This type of DFI allocates factors of production in a trade-reducing and inefficient way as it does not utilize comparative factor-cost advantages.

Trade theory suggests that location of international production is based on comparative advantages of factor costs. If firms use DFI to minimize costs, it will move to the location where production costs are lowest. However, empirical evidence shows that trade barriers and labour costs are a very incomplete framework to analyse the location of DFI. The concept of 'locational advantages', as reviewed by Caves [1982] and Dunning [1993], covers many influences. While popular debate still focuses on production costs, research suggests that the attractiveness of the local markets is at least as important.

Production cost advantages are an important component of locational decisions in industries with low transportation costs. Their DFI depends on costs of production in alternative countries, in particular productivity-adjusted labour costs. Thus factors influencing productivity are determinants of DFI. This includes transportation and telecommunications infrastructure, quality of the human capital, for example education and employee motivation,⁵ and quality, reliability and costs of local supplies. Facilities processing natural resources naturally depends on the existence of natural resources.

Market-related advantages are increasingly replacing factor costs as the prime determinant of DFI. Proximity of the production to the market becomes the overriding consideration in any of the following situations:

- C Protectionism (tariffs, quotas, administrative barriers to trade) can be bypassed by DFI. It can jump tariff barriers and obtain or maintain market access, and even extract rents generated by trade barriers. Also, DFI can be a means to prevent or preempt anticipated protectionist measures and to establish presence in a trading bloc.⁶
- C Transportation costs are a natural barrier to trade. They are diminishing in relevance due to modern transportation technology but are still relevant for bulky goods and fresh food.
- C Production and sales activities may be indivisible, especially in service industries (hotels, banking, trade, consulting). DFI in the service sector is of increasing importance, but has often been neglected by academic research focused on manufacturing [McCulloch 1988].
- C The interaction between production and sales activities may require local production. This includes cooperation with downstream firms, such as just-in-time delivery or long-term reliable supplies. Local production can improve performance by increasing flexibility, after-sales service, or access to market information or technological know-how, which in turn influence innovation, product design or marketing.
- C Investment in distribution channels may complement exports to the host economy. Acquisition of existing distribution networks from local competitors or adaptation of established local brand names are fast market penetration strategies [Sölvell 1987].

Investment of these types is becoming more important along with modern management in production and marketing. It depends primarily on the potential market, that is market size and growth, plus costs of local production. This provides a theoretical rationale for empirically established positive effects of host country market size and growth on DFI.⁷ However, according to Dunning [1993: 142] the potential loss of a market is the paramount driving force behind market-oriented investment and is of greater importance than gaining entry into new markets.

Developmental Model

Comparative advantages of nations evolve with the process of economic development. On this basis, stages models relate the product cycle [Vernon 1966, 1979] or the economic development of the source country [Dunning 1986, Ozawa 1992, Narula 1995] to outward DFI. Simultaneously, research has focused on differences of DFI within East-Asia and other regions of the world [for example Kojima 1978, Lee 1990, Ramstatter 1991]. This stream of literature is reviewed and extended in appendix 3.1 with the objective of developing a model of the interaction between changing comparative advantages and inward and outward DFI flows.

Economic Geography

The location of economic activity in geographic space has largely been analysed independently of mainstream economics in the field of economic geography. Krugman [1991] highlights the importance of this work for the explanation of regional concentration of economic activity. Alfred Marshall [1890/1916] already points out the causes of economic agglomeration:

- ⊞ the pooling of markets for specialized skilled labour,
- ⊞ the development of subsidiary trade and suppliers of intermediate inputs,
- ⊞ the flow of information, especially technological know-how, between firms.

Krugman's work focuses on modelling the agglomeration process, especially external economies of scale in labour and input markets. Fixed costs in the industry, regional dispersion of the markets and transportation costs determine the cumulative process of concentration. The locational patterns can change very suddenly: once a critical mass of capital and industry-specific infrastructure is accumulated, investment moves to new centres that may evolve by historical accident or temporary protectionism [Krugman 1991].

Krugman [1992] formalizes the tensions between scale-related 'centripetal' and market-related 'centrifugal' forces of locational decisions. In simulations he shows the agglomeration of economic centres with given economies of scale, transportation costs, immobile farmers, and mobile production workers. The same argument applies to DFI. It is the international allocation of mobile capital in the presence of immobile workers and complex barriers to trade. Research extending this approach primarily uses simulation techniques [for example Krugman and Venables 1994, Markusen and Venables 1995].

These arguments imply that the existing industrial structure can be a major determinant of inward DFI. Suppliers of intermediate goods and a technologically specialized labour force are locational advantages for related firms and competitors.⁸ The effects are especially observable for DFI because of specific externalities:

- ⊞ Service industries such as banks and consultants follow their customers [Erramilli and Rao 1990], but once established they provide services and information to other potential investors.
- ⊞ Local individuals and institutions adapt to the needs of foreign MNEs: managers may learn foreign languages, governments set up foreign investment agencies and change the legal framework, and local businesses upgrade their quality standards.
- ⊞ Suppliers of intermediate goods follow their customers, as widely reported in the automobile industry.

Other research focused on the third aspect of externalities: the exchange of knowledge. Innovation processes tend to be localized, and knowledge is highly tacit at early stages of development and tends to stick to the local milieu [Aydalot 1986, Malecki 1991]. Intense innovative activity in an area contributes not only to firms' competitiveness [Porter 1990] and the evolution of multinational firms [Sölvell, Zander and Porter 1991]. It also attracts additional investors who wish to participate in the innovative activity.

Access to localized knowledge is increasingly important for the advancement of technological competence of MNEs [Cantwell 1989]. Thus, especially the

location of R&D activities follows patterns of similar or complementary technological competence in the local environment, leading to a cumulative process. The Swedes Malmberg, Sölvell and Zander [1996] call it the 'Greta-Garbo-effect' after the Swedish actress who was attracted to Hollywood and later herself attracted more business related to the movie industry. This approach can explain apparently paradoxical phenomena such as Korean DFI in the Californian semiconductor industry: by becoming insiders they gain access to the knowledge pool in Silicon Valley for the benefit of their own innovation and development.

Institutional Analysis

The general institutional framework in both source and host countries influences the volume of DFI and its characteristics. This consists of the social environment as well as the legal, institutional and general policy environment ('Ordnungspolitik'). Research has mainly focused on host country policies rather than on countries of origin, presumably because most countries take a neutral attitude towards outward DFI [Meyer, Ambler and Styles 1994]. Exceptions are Japan in the 1970s when outward DFI was supported actively [Ozawa 1979b], and Sweden [Blomström and Kokko 1995].

Empirical evidence suggests that the general policy framework plays a more important role in attracting DFI than fiscal measures specifically designed to attract DFI. This includes the openness of the economy [Li and Guisinger 1992], approval procedures and bureaucracy, tax regime, environmental regulation and other aspects of business law. The nature of regulative environment may become a significant advantage over alternative locations or the home location if the latter is tightly regulated in specific industries, for example by environmental standards. For instance, negative effects of the social or institutional environment can arise with uncooperative bureaucracy, restrictions on foreign ownership and profit remittance, or a high degree of unionization and union bargaining.

At best, specific fiscal incentives and tax allowances geared towards DFI play a marginal role although they may influence the choice of location within a country or region. This evidence is regularly found in studies of DFI into developing countries [Guisinger et al. 1985, Hill 1990] as well as industrial countries [Safarian 1993].⁹ However, specific incentives and requirements for DFI influence the performance of DFI, for example the local content of inputs or the share of exported output.¹⁰ The impact of governmental policy on the competitive structure of markets and DFI, however, differs between industrialized and developing countries [Brewer 1993].

In the increasingly interrelated world economy the relationships among companies, among host and home governments, and between companies and government are increasingly intertwined and complex [Stopford and Strange 1991]. Major investment projects are increasingly subject to individual negotiations between investor and host country agencies, not only in the case of privatization-related DFI. The relative bargaining power of the MNE *vis-à-vis* its host government has been used successfully to explain the organizational form of DFI as weaker MNEs have to accept a JV partner [Kobrin 1987, Gomes-Casseres 1991]. Game-theoretic models have been developed to analyse some of the emerging interaction between institutions and MNE [for instance Vannini 1995, Haapanranta 1996].

A peculiar aspect of the institutional framework is political risk. It arises with potential changes in the legal framework of any of the countries involved which affect the return on investment of DFI. It includes the impact of political violence or revolutions as well as the changes in the structure of taxes, tariffs and the regulatory environment. All these can induce major changes in relative prices. Thus small changes may have a major impact on the profitability of foreign investments.

Political risk is generally reported as a major deterrent of DFI in survey-based studies. The econometric evidence is weak, however, mainly due to problems defining political risk and finding appropriate proxies. Studies using indices based on the quantity and intensity of political events are not very successful

in explaining DFI flows. They show the relevance of political variables, but their impact on DFI was small compared with economic variables: Nigh [1985] finds significant effects of inter-country conflict and cooperative events and - in developing host countries - of internal political events. Schneider and Frey [1985] compare various models and obtained best performance of a model that included several both economic and political variables. Edwards [1990] uses indices by Cukierman, Edwards and Tabellini [1992] for the probability of change in government and for political violence but found significant impact on DFI only by the former. Chase, Kuhle and Walther [1988] use commercial risk indices to proxy political risk and find no support for the hypothesis that country risk is compensated by a higher return on investment.

INDUSTRIAL ORGANIZATION

DFI is most prominent in industries with large economies of scale, intangible assets, high product differentiation and worldwide oligopolistic market structures. This is mainly horizontal DFI among high income economies rather than vertical DFI that would take advantage of factor cost differentials [Markusen 1995]. Oligopolistic competition strongly suggests that at least the short-term dynamics of DFI and the timing of investment, if not the location decision as such, are influenced by strategic motives. Multinationals consider their strategic positions *vis-à-vis* their main rivals in their most important markets to decide on market entry and investment projects. This section considers which advantages may induce DFI, how strategic interaction affects DFI in oligopolistic markets, and why firms internalize international business.

Sources of Competitive Advantage

A major school of thought views incomplete markets as the main reason for DFI. Foreign investors have a competitive disadvantage relative to local competitors due to lack of information on local market conditions and higher costs of communication and transportation. To overcome these disadvantages and to operate profitably in foreign markets, they must have some kind of firm-specific advantage. This explanation of DFI as a function of firm-specific or 'ownership advantages' is related to the 'resource-based view' of the firm [Penrose 1959/1995, Wernerfelt 1984, Conner 1991] in the management literature.

Since Hymer [1960/1976] and Kindleberger [1969] many sources of firm-specific advantages have been analysed. In order to induce DFI, the advantage has to be both transferable within the MNE and specific to the firm. Thus the firms have to possess some degree of monopolist power.¹¹ In addition - as argued in the transaction cost approach - internal transfer has to be superior to an external transfer [Caves 1971]. In this framework multinational firms are mainly exporters of the services of firm-specific assets [Markusen 1991, 1995]. Dunning [1993] distinguishes three firm-specific, or 'ownership', advantages:

- C resources based on the assets of the firm, including property rights and intangible assets,
- C advantages of common governance of the established firm over a *de novo* entrant,
- C advantages of common governance arising because of multinationality.

Relevant corporate assets include physical assets, intellectual property rights and intangible assets embodied in the human capital of the firm, such as management, engineering, marketing and financial capabilities. In terms of Prahalad and Hamel [1990], competitive advantages arise from 'core competencies' such as technological know-how, and 'value-creating activities' such as total quality control and just-in-time manufacturing systems. In other cases, firms may possess assets that arise from the regulatory environment, for example preferred access to natural resources.

Advantages of common governance arise from economies of scale on firm level rather than on plant-level. This includes centralized R&D and marketing or

favourable access to resources. Advantages of multinationality arise from market power, worldwide accumulation of technology, and business contacts and knowledge of managing a worldwide network of activities. Operating in a variety of environments exposes MNEs to many challenges and innovations which stimulate the development of specific competencies and learning opportunities which are not available to purely national firms [Bartlett and Ghoshal 1989]. Therefore ownership advantages become increasingly specific to the firm and independent of the asset base and economic structure of the home economy [Narula 1995]. Advantages of multinationality can have a reinforcing 'experience effect'. Firms established internationally are best positioned for further expansion because acquisition of knowledge is a cumulative process of interaction between the creation of technology and its application in production [Pavitt 1987, Cantwell 1989].

Empirical studies have focused on the identification of relevant firm-specific advantages. The review of this research by Dunning [1993: 142-3, 148-53, 160-4] suggests that the most important advantages are technology-related, including capabilities of generating technological know-how, as well as brand names and marketing knowledge. The effect of multi-plant industries conferring back advantages to their owners, receives some support, while the experience effect is empirically difficult to separate from other effects. However, the empirically significant firm-specific advantages vary widely across source countries. For instance technology¹² and marketing assets were of great significance for US firms, but not for Japanese [Hennart and Park 1994].

Strategic Competition

The analysis of oligopolistic competition among MNEs has for a long time considered two effects: Graham's 'exchange of threats' hypothesis, and Knickerbocker's 'follow the leader' hypothesis. More recently game-theoretic models consider DFI as a strategic move within oligopolistic competition in order to obtain first-mover advantages. Formal models have also been developed by international trade economists introducing market imperfections and firm-level economies of scale into their models.

Graham [1975, 1978] models intra-industry DFI resulting from the 'exchange of threats' between rivals. In his model firms finding their domestic market invaded by a foreigner will retaliate by attacking the monopolistic position of the rival in his home market. This strategy is particularly relevant for capital-intensive production processes with significant economies of scale. The basis of the argument is a model of Cournot-type competition between two firms, both enjoying monopolies in their home market but with different marginal costs. Graham identifies conditions that trigger an entry into the rival's market. The argument is further refined by considering experience curve effects that reduce marginal cost as the volume of production increases, creating incentives to increase output at an early stage in order to slide down the experience curve sooner than the competitors. He concludes that the exchange of threats maintains competition, but in a less cut-throat form than between enterprises with large production facilities and low marginal costs. Graham [1985] extends the argument by suggesting that this cross-investment would accelerate new product development and make collusion less likely.

Knickerbocker [1973] suggests that dominated firms in an oligopoly imitate the strategy of the leader to prevent him from gaining an early lead advantage by establishing a position in the market and factually raising entry barriers. Scharfstein and Stein [1990] model this 'follow the leader' pattern: with managers of the follower being assessed in their performance through comparison with the leader. With an imitation strategy their downside risk is to miss a major opportunity. Investing in a project similar to the leader may be more risky in absolute terms but not relative to the position of the leader.

The hypothesis implies that DFI increases with industry concentration. Knickerbocker [1973] and Flowers [1976] detect a concentration of entry by firms in the same industries. The phenomenon increases with industry concentration, but decreases with very high concentration which they interpret as indicating tacit collusion in very narrow markets. The hypothesis receives further empirical

support by Yu and Ito [1988] who compare DFI in a competitive and an oligopolistic industry, and Li and Guisinger [1992] who analyse service MNEs. However, the phenomenon may be explained in an alternative way: (a) followers may assume that the leaders have undertaken proper market research and his investment thus 'signals' an investment opportunity; (b) the leaders contribute to the local infrastructure by their externalities which make the location more attractive for suppliers, customers, and, subsequently, competitors;¹³ (c) the effect may be spurious as both competitors react to a common external stimulus such as market liberalization.

Porter [1990] focuses on push factors arising from the competitive nature of the home market. He argues that domestic competition strengthens firms' competitive advantages because it creates permanent challenges for improvements. This competitive strength makes firms 'fit' for international competition where they may compete with a competitor from the same region of origin [Porter 1990: 117-22]. Dominated firms in oligopolistic competition may actually lead the move abroad because they face limits to expansion in their domestic markets [Mascarenhaas 1986, Ito and Pucik 1993].

Other researchers focus on the interaction between the foreign investor and local agents. Dixit [1980] presents a game-theoretic model to analyse the interaction between an MNE and a potential competitor. DFI can be a strategic move to deter entry: by choosing the DFI option the MNE can deter a local competitor from emerging, as his post-investment decisions are based on the lower marginal costs of local production. Extending the model, Smith [1987] and Jacquemin [1989] show that DFI can replace exports even in the absence of tariff barriers: DFI simultaneously reduces transaction costs and increases market power through a commitment of sunk costs because the locational decision is irreversible.

Further models of strategic motives inducing DFI have been developed by what should be called the 'Leuven school' of DFI. For instance, Motta [1992] shows how the decision between exports and DFI becomes non-monotonic because of the interaction with the potential local competitor. In Motta [1994], he shows how DFI can crowd out an existing local competitor as well as a competing MNE exporting to the country. Veuglers [1995] presents a model in which firms are induced to become multinational by their domestic rivals' (potential) DFI, because they may incur competitive disadvantages in their home markets. Sels [1996] considers a waiting game between two potential entrants with externalities from the first mover.

Modern International Trade Theory

The neoclassical theory of international trade assumes perfect competition. Only by dropping this assumption can it explain issues related to DFI which include intra-industry trade and locational decisions of MNEs. Models have been developed to illustrate locational decisions of multinational firms and, building on these models, alternative patterns of multinational firms under different policy regimes.

Locational decisions of firms in imperfect markets have been modelled on the basis of intra-industry trade models. Krugman [1983] presents two models of horizontal and vertical MNEs with firm-specific advantages. Horizontal MNEs are modelled as a response to product differentiation. Costs of producing locally are assumed to be higher than at home, but if marginal costs of exporting exceed the marginal costs of local production, the firm shifts its production to the market. Vertical MNEs are explained in a model of a monopsonistic downstream firm that can eliminate the distortions of monopsonistic markets by international backward integration. Along similar lines Helpman [1984, 1985] and Helpman and Krugman [1985] analyse locational decisions for single plants in general equilibrium trade models with increasing returns on the level of the firm and a given non-competitive market structure with differentiated products. Markusen [1984] and Horstmann and Markusen [1987a] consider multiple plants under a single headquarter.

The next generation of models by Horstmann and Markusen [1992] and Brainard [1993] has endogenous market structures in which they show the emergence of MNEs. Both models have firm-level activities with joint inputs across plants, plant-level economies of scale, and tariffs or transportation costs between the two countries in the model. MNEs emerge in equilibrium if firm-level fixed costs are large and tariffs and transportation costs dominate plant-level scale economies. MNEs are more likely to exist if both countries are large and, in Brainard's [1993] model, if the countries have similar relative factor endowments. Horstmann and Markusen [1992] also show that small changes in the underlying locational advantages, for example taxation, can cause major shifts in the market structure because of movements between different Nash equilibria resulting in jumps in prices and output.

Markusen and Venebles [1995] use this model as the basis for simulations showing that, for countries of similar size, multi-plant MNEs displace international trade. Markusen et al. [1996] and Markusen [1997] analyse alternative trade and investment regimes and suggest that, under full liberalization, MNEs may locate one function in each country replacing the multi-plant MNEs. Between countries of similar size and factor endowment, no MNEs emerge. New international trade theory has been combined with game-theoretic models of the Leuven school by Motta and Norman [1996] and Sanna-Randaccio [1996] in order to analyse the effects of economic integration. Their models show how the removal of trade barriers triggers market-seeking DFI.

Similar models have been used to analyse licensing versus DFI decisions. The internalization theory (discussed in Chapter 4) has explored incentives in great detail but international trade economists have provided more formal models. Ethier [1986] considered market failure due to informational asymmetry with respect to the value of the technology being licensed. Horstmann and Markusen [1987b] consider incentives of franchising contracts where the franchisor has to monitor the quality of the local franchisee to protect his reputation. Ethier and Markusen [1996] and Saggi [1996] consider the potential diffusion of knowledge that may create third market competition. Horstmann and Markusen [1996] use an agency model to analyse the incentives between a licensor and a licensee who has superior information on the local market. They show how DFI is motivated by the unwillingness to share rents with a local licensee. Temporary licensing may be preferred if costs of investment mistakes due to unfamiliarity with the market are high.

THE SCOPE OF THE FIRM

Internalization Theory

Internationalization theory explains the emergence of multinational enterprises from the failure of markets. Its roots are in the transaction cost (TC) approach initiated by Coase [1937] but it has largely been developed independently of the well known work on TC by Williamson [1975, 1981, 1985]. Early contributions are Caves [1971], Buckley and Casson [1976], McManus [1972], Swedenborg [1979], Rugman [1981] and Hennart [1982].

The views of researchers of internalization theory do not differ in substance from those of transaction costs economists, but in emphasis: whereas Williamson's arguments focus primarily on market failure due to lock-in effects arising from asset specificity, internalization theory focuses on market failure in markets for information. Many assets transferred by MNEs to their affiliates partially have a public-good nature such that market transactions fail due to information asymmetries. Chapter 4 reviews the concepts of this literature and develops a synergetic model. Some authors, in particular Rugman [1981, 1985] and Hennart [1995], argue that internalization is a sufficient explanation for the existence of MNEs. This view contrasts with Dunning's OLI paradigm where all three conditions, ownership, location and internalization, are necessary to explain DFI. In this study Dunning's view is adopted.

Transaction cost economics (TCE) treats decisions on engaging in a transaction and its internalization as distinct and is therefore a static approach.

Some dynamic approaches to TC have aspired to overcome this limitation: Buckley [1988, 1990, Buckley and Casson 1985] incorporates dynamic aspects of corporate expansion and strategic actions, which are taken not to overcome market failure, but to create or exploit it. Internalization incentives arising from strategic positioning have been incorporated in one 'internalization theory'. This approach takes an opposite line of reasoning: rather than deriving internalization incentives from market failure, it defines them as the motive that may lead to an internalization decision. Langlois' [1992, 1995] dynamic view of transaction costs sees boundaries of firms entirely determined by capabilities of the firm rather than market failure. He argues that in an uncertain environment, common ownership of multiple stages of production is a superior institutional arrangement for coordinating systemic change.

Kogut and Zander [1993, 1995, Zander and Kogut 1995] depart from the market-failure approach of TC arguing that the transfer of tacit knowledge explains internalization. Markets are not considered to be a feasible alternative because of the need for an organizational mode to transfer tacit knowledge. Thus the creation, accumulation and transfer of tacit know-how determine the evolutionary growth of firms.

Internationalization Process Model

The theories reviewed so far consider DFI as determined by characteristics of the firm and its environment. Researchers on internationalization processes analyse the international business of a firm as a gradual process. Based on the early contributions by Johanson and Wiedersheim-Paul [1975] and Johanson and Vahlne [1977], this research is frequently called the Uppsala school. Other early contributors are Luostarinen [1979] in Finland, and in the American literature Bilkey and Tesar [1977], Cavusgil and Nevin [1981].

The model by Johanson and Vahlne [1977, 1990] is rooted in the behavioural theory of the firm following Cyert and March [1963] and Aharoni [1966] as well as the growth theory of the firm by Penrose [1959/1995]. The gradual increase of firms' international involvement is explained by an interplay between the development of knowledge on foreign locations and operations in the countries, and, on the other hand, an increasing resource commitment. Knowledge on foreign markets is 'experiential knowledge' which cannot be taught. It can only be acquired through experience and active involvement in the country. Such knowledge is essential for resource commitment because it enables recognition of business opportunities and reduces market uncertainty. Therefore, past commitment and accumulated country-specific experience determine current activities as well as future resource commitments and involvement on a higher level. American researchers modelled the process analogous to innovation adaptation as incremental increases of experience and learning over multiple stages [Cavusgil 1980, Reid 1981].

This understanding of the internationalization process has three implications. Firstly, firms will typically follow an 'establishment chain' moving from lower to higher modes of involvement. This has led to a number of stages models: the Swedish school suggests an initial phase with no regular export activities, then exports via independent representatives, then sales subsidiaries, and eventually local manufacturing [Johanson and Wiedersheim-Paul 1975, Johanson and Vahlne 1990, Nordström 1991]. Other models have additionally introduced contractual business such as licensing and JVs as stages of the internationalization process [Root 1987, Young et al. 1989, Kay 1991].

The second implication is that firms enter markets in a sequence starting in countries in close 'psychic distance'. This term, first used by Beckermann [1956], includes not only geographical but also cultural, political and linguistic communalities between the home and the host economy [Johanson and Wiedersheim-Paul 1975]. Traditional business ties also reduce unfamiliarity and thus increase present DFI. Luostarinen [1979] argues that even similarity of the economic conditions and market size favour an early entry.¹⁴ Thirdly, initial investments in a country can serve as a platform for learning about a market or to allow customers to develop brand loyalty. A platform creates an option for further DFI and taking advantage of emerging opportunities [Kogut 1983, Kogut

and Chang 1996].

Case study research frequently found support for the sequential entry pattern.¹⁵ However, the pattern was established primarily for firms at an early stage of internationalization [Forsgren 1989]. The relative importance of psychic distance appears to have declined since the 1970s as economic conditions are becoming more important, for example industry-specific barriers to entry [Sölvell 1987], market potential and industry structure [Nordström 1991]. Also, firms move more rapidly from low to high involvement modes and may even leap-frog some stages of the traditional model [Nordström 1991, Engelhard and Eckert 1994].

The limitations of the internationalization process models are, first, a weak delineation of theoretical boundaries, that is the underlying assumptions and scope of the models; second, weak explanatory power and, third, insufficient congruence between the theoretical and operational level [Andersson 1993].

A SYNTHESIS: THE ECLECTIC PARADIGM

John Dunning [1977] integrates many theories surveyed in this chapter into a general paradigm of international production. He extends the framework repeatedly [1981, 1988, 1993], most recently to explain strategic alliances [Dunning 1995]. The basic premise is that DFI is undertaken if three conditions are met simultaneously. If not, exporting or licensing may be superior strategies. Based on the acronyms of the three components, this approach is commonly known as the 'OLI-paradigm'.

- C The investing firm needs 'ownership advantages', that is specific assets to obtain a competitive advantage over local competitors. They include property rights and intangible assets, named 'Oa advantages', as well as advantages arising from common governance, named 'Ot advantages'. Oa advantages include advantages due to abilities that facilitate the generation of new assets, especially knowledge. Ot advantages are capabilities of organizing Oa advantages with complementary assets. They include (i) those of branch plants of established enterprises over *de novo* firms, and (ii) those arising specifically from multinationality.
- C The host country must possess 'locational advantages' which include factor cost advantages, proximity to the market, the existing economic structure, and the legal, social and political frameworks.
- C 'Internalization incentives' must make it more efficient for the MNE to use its competitive advantage by selling components internally rather than in the market place. These advantages may arise from market failure as discussed in the transaction cost and internalization literature (see Chapter 4), but may also arise because of distortions in the regulatory environment.

This chapter has given a condensed summary of the various streams in the literature that contribute to the explanation of DFI. It sets the scene for two lines of theoretical work that shall be extended in the next chapter: the internalization or transaction cost approach.

NOTES

1. Dunning [1993, chapter 4] gives a comprehensive survey of the theory of DFI in a mainly microeconomic context. Pitelis and Sugden [1991] present a collection of papers on current issues related to the theory of the multinational firm, of which Cantwell [1991] gives a comprehensive discussion of the state of the art. Markusen [1995] reviews modern international trade literature related to DFI. Also see Lizondo [1991] and Meyer and Rühmann [1993]. Agarwal [1980] surveys earlier theoretical and empirical literature.
2. The model assumes that the long-term expected value of foreign assets in home currency is independent of current exchange rates.

3. GDP and/or GDP growth have been tested by, inter alia, Hultman and McGee [1988], Culem [1988], Barrel and Pain [1991], Bajo-Rubio [1991] as well as Bajo-Rubio and Sosvilla-Rivero [1994]. Clegg [1995] found a surprisingly negative effect. O'Sullivan [1993] and Millner and Pentecost [1992] defined the relevant market beyond the host country, by referring to the European Union.
4. This applies to time-series studies [Goldsbrough 1979, Cushman 1988, Barrel and Pain 1991, Pain 1993, 1996, Bajo-Rubio and Sosvilla-Rivera 1994] as well as cross-country studies [Swedenborg 1979, Kravis and Lipsey 1982, Schneider and Frey 1985, Woodward and Rolfe 1993, Yamawaki 1993, Thiran and Yamawaki 1995, Döhrn 1996].
5. Educational and technological infrastructure was shown to be significant in attracting DFI by Swedenborg [1979], Cantwell [1989] and Yamawaki [1993]. For this reason some studies found a positive association between endowments with skilled labour [Svensson 1996], or the wage level, and DFI [Swedenborg 1979, Thiran and Yamawaki 1995].
6. Pain [1996] and Döhrn [1996] estimate the impact of the integration in the European Union on DFI flows and find a significant positive effect on flows to the countries of the union.
7. See cross-country studies by Swedenborg [1979], Kravis and Lipsey [1982], Dunning [1980], Veuglers [1991], and Svensson [1996], time-series studies cited above, and reviews by Lizondo [1991] and Stehn [1992].
8. Svensson [1996] found empirical support for an industry agglomeration index attracting DFI.
9. The impact of general tax legislation and specific tax allowances on DFI has been analysed empirically by Root and Ahmad [1978, 1979], Lim [1983] and Woodward and Rolfe [1993] for DFI in developing countries, and Grubert and Mutti [1991], He and Guisinger [1992] and Loree and Guisinger [1995] for outward DFI from the US.
10. This arises in Guisinger et al. [1985], Wells [1986], Hill [1990], and Loree and Guisinger [1995].
11. The monopolistic nature of firm specific advantages has been of major concern to Hymer [1960/1976], Kindleberger [1969, 1984] and others because of the potential extraction of monopolistic rents from the host economy. For instance, Newfarmer [1985] and Cowling and Sugden [1987] are concerned with collusion between MNEs that may become more likely if they develop similar international structures and more easily can agree on any collusive action. See Yamin [1991] for a review of research following the market-power approach in Hymer's work. Most researchers of MNEs apply a Schumpeterian view of competition where monopolistic advantages are temporary and create incentives for innovation and dissemination of new products.
12. Technological capabilities are commonly measured by the ratio of R&D expenditures in turnover. Empirical research established a positive relation between R&D expenditure and the propensity for DFI [for example Caves 1974, Mansfield, Romeo and Wagner 1979, Grubaugh 1987, Hennart and Park 1994, Wagner and Schnabel 1994, Kogut and Chang 1996, Svensson 1996].
13. The empirical studies use broad industry classifications which cannot distinguish between 'follow the leader' and 'follow the customer' as is commonly the case for automotive suppliers and services.
14. Casson [1995] presents a formal model of entry decision-making and shows which conditions would favour sequential entry, including expected similarities between alternative foreign markets, and their differences to the home market, low costs of deferred entry, and high costs of learning combined with low costs of communicating experience from the first to the second foreign country.
15. See for example Johanson and Wiedersheim-Paul [1975], Luostarinen [1979], Larimo [1985], Buckley, Newbold and Thurwell [1979], Davidson [1980], Veuglers [1991] Jansson [1993], Chang [1995]. Contradictory findings emerge from Hood and Young [1983]. The entry sequence pattern does not seem to apply to service industries such as advertising agencies [Terpstra and Yu 1988]. A possible explanation is that the internationalization of business service firms is driven typically