Foreign Direct Investment

Theory, Evidence and Practice

Imad A. Moosa

palgrave

1 Introduction and Overview

WHAT IS FOREIGN DIRECT INVESTMENT?

Foreign direct investment (FDI) is the process whereby residents of one country (the source country) acquire ownership of assets for the purpose of controlling the production, distribution and other activities of a firm in another country (the host country).[1] The International Monetary Fund's *Balance of Payments Manual* defines FDI as 'an investment that is made to acquire a lasting interest in an enterprise operating in an economy other than that of the investor, the investor's purpose being to have an effective voice in the management of the enterprise'. The United Nations 1999 *World Investment Report* (UNCTAD, 1999) defines FDI as 'an investment involving a long-term relationship and reflecting a lasting interest and control of a resident entity in one economy (foreign direct investor or parent enterprise) in an enterprise resident in an economy other than that of the foreign direct investor (FDI enterprise, affiliate enterprise or foreign affiliate)'.[2] The term 'long-term' is used in the last definition in order to distinguish FDI from portfolio investment, the latter characterized by being short-term in nature and involving a high turnover of securities.[3]

The common feature of these definitions lies in terms like 'control' and 'controlling interest', which represent the most important feature that distinguishes FDI from portfolio investment, since a portfolio investor does not seek control or lasting interest. There is no agreement, however, on what constitutes a controlling interest, but most commonly a minimum of 10 per cent shareholding is regarded as allowing the foreign firm to exert a significant influence (potentially or actually exercised) over the key policies of the underlying project. For example, the US Department of Commerce regards a foreign business enterprise as a US foreign 'affiliate' if a single US investor owns at least 10 per cent of the voting securities or the equivalent. Both equity and debt-financed capital transfers to foreign affiliates are included in the US government's estimates of FDI. Sometimes, another qualification is used to pinpoint FDI, which involves transferring capital from a source country to a host country. For this purpose,

investment activities abroad are considered to be FDI when (i) there is control through substantial equity shareholding; and (ii) there is a shift of part of the company's assets, production or sales to the host country. However, this may not be the case, as a project may be financed totally by borrowing in the host country.

Thus, the distinguishing feature of FDI, in comparison with other forms of international investment, is the element of control over management policy and decisions. Razin *et al*. (1999b) argue that the element of control gives direct investors an informational advantage over foreign portfolio investors and over domestic savers. Many firms are unwilling to carry out foreign investment unless they have one hundred per cent equity ownership and control. Others refuse to make such investments unless they have at least majority control (that is, a 51 per cent stake). In recent years, however, there has been a tendency for indulging in FDI co-operative arrangements, where several firms participate and no single party holds majority control (for example, joint ventures).

But what exactly does 'control' mean in the definition of FDI? The term 'control' implies that some degree of discretionary decision-making by the investor is present in management policies and strategy. For example, this control may occur through the ability of the investor to elect or select one or more members on the board of directors of the foreign company or foreign subsidiary. It is even possible to distinguish between the control market for shares and the non-control or portfolio share market as an analogy to the distinction between direct investment and portfolio investment. It may be possible to exercise control via contractual (non-equity) arrangements. The non-equity forms of FDI include, *inter alia*, subcontracting, management contracts, franchising, licensing and product sharing. Lall and Streeten (1977) argue that a majority shareholding is not a necessary condition for exercising control, as it may be achievable with a low equity share and even without an explicit management contract.

So, it is possible (in theory at least) to define and characterize FDI, but measuring FDI in practice is a totally different 'game'. There are inherent problems in measuring FDI, particularly when the investment takes the form of machinery or capitalized technological contributions. There are also gaps in the FDI statistics available from the source and host countries on FDI. Most countries do not publish comprehensive information on the foreign operations of their companies, for reasons of secrecy. Because of these problems, inconsistency

between measures of FDI flows and stocks are the rule rather than the exception.[4] Furthermore, Cantwell and Bellack (1998) argue that the current practice of reporting FDI stocks on a historical cost basis (that is, book value) is unsatisfactory, because it does not take into account the age distribution of stocks, which makes international comparisons of FDI stocks almost impossible.

Interest in FDI, which has motivated attempts to come up with theories that explain its causes and effects, is attributed to the following reasons.[5] The first reason is the rapid growth in FDI and the change in its pattern, particularly since the 1980s. In the 1990s, FDI accounted for about a quarter of international capital outflows, having grown relative to other forms of international investment since the 1970s. The rapid growth of FDI has resulted from global competition as well as from the tendency to free up financial, goods and factor markets. It has been observed that FDI flows continue to expand even when world trade slows down. For example, when the growth of trade is retarded by trade barriers, FDI may increase as firms attempt to circumvent the barriers (see for example, Jeon, 1992; and Moore, 1993). It has also been observed that even when portfolio investment dried up in Asian countries as a result of the crisis of the 1990s, FDI flows were not affected significantly. Lipsey (1999) argues that FDI has been the least volatile source of international investment for host countries, with the notable exception of the USA. The latest available OECD figures show the following: FDI inflows to OECD countries increased from US$249 billion in 1996 to US$684 billion in 1999, whereas FDI outflows increased from US$341 to US$768 during the same period. This growth is rather dramatic (we shall examine the relevant statistics in more detail later).

The second reason for interest in FDI is the concern it raises about the causes and consequences of foreign ownership. The views on this issue are so diverse, falling between the extreme of regarding FDI as symbolizing new colonialism or imperialism, and the other extreme of viewing it as something without which the host country cannot survive. Most countries show an ambivalent attitude towards FDI. Inward FDI is said to have negative employment effects, retard home-grown technological progress, and worsen the trade balance. A substantial foreign ownership often gives rise to concern about the loss of sovereignty and compromise over national security. Outward FDI is sometimes blamed for the export of employment, and for giving foreigners access to domestic technology.

The third reason for studying FDI is that it offers the possibility for channelling resources to developing countries. According to this argument, FDI is becoming an important source of funds at a time when access to other means of financing is dwindling, particularly in the aftermath of the international debt crisis that emerged in the early 1980s. Lipsey (1999) argues that FDI has been the most dependable source of foreign investment for developing countries. Moreover, FDI is (or can be) important in this sense not only because it entails the movement of financial capital but also because it is normally associated with the provision of technology as well as managerial, technical and marketing skills. But it has to be emphasized here that FDI does not necessarily involve the movement of financial capital, as the investor may try to raise funds by borrowing from financial institutions in the host country. Moreover, the other benefits of FDI may not materialize, or they may materialize at a very high cost for the host country. All of these issues will be examined in the following chapters.

Finally, FDI is thought to play a potentially vital role in the transformation of the former Communist countries. This is because FDI complements domestic saving and contributes to total investment in the (host) economy. It is also because FDI brings with it advanced technology, management skills and access to export markets. Again, these positive effects may not arise, or they may arise simultaneously with some adverse effects.

TYPES OF FDI

FDI can be classified from the perspective of the investor (the source country) and from the perspective of the host country. From the perspective of the investor, Caves (1971) distinguishes between horizontal FDI, vertical FDI and conglomerate FDI. Horizontal FDI is undertaken for the purpose of horizontal expansion to produce the same or similar kinds of goods abroad (in the host country) as in the home country. Hence, product differentiation is the critical element of market structure for horizontal FDI. More generally, horizontal FDI is undertaken to exploit more fully certain monopolistic or oligopolistic advantages, such as patents or differentiated products, particularly if expansion at home were to violate anti-trust laws. Vertical FDI, on the other hand, is undertaken for the purpose of exploiting raw materials (backward vertical FDI) or to be nearer to the consumers through the acquisition of distribution outlets (forward vertical FDI).

For example, for a long time, US car makers found it difficult to market their products in Japan because most Japanese car dealers have close business relationships with Japanese car makers, thus making them reluctant to promote foreign cars. To overcome this problem, American car dealers embarked on a campaign to establish their own network of dealerships in Japan to market their products. The third type of FDI, conglomerate FDI, involves both horizontal and vertical FDI. In 1999 horizontal, vertical and conglomerate mergers and acquisitions (which is one of two forms of FDI, as we shall see later) accounted for 71.2 per cent, 1.8 per cent and 27 per cent, respectively, of the total value of mergers and acquisitions worldwide.

From the perspective of the host country, FDI can be classified into (i) import-substituting FDI; (ii) export-increasing FDI; and (iii) government-initiated FDI. Import-substituting FDI involves the production of goods previously imported by the host country, necessarily implying that imports by the host country and exports by the investing country will decline. This type of FDI is likely to be determined by the size of the host country's market, transportation costs and trade barriers. Export-increasing FDI, on the other hand, is motivated by the desire to seek new sources of input, such as raw materials and intermediate goods. This kind of FDI is export-increasing in the sense that the host country will increase its exports of raw materials and intermediate products to the investing country and other countries (where the subsidiaries of the multinational corporation are located). Government-initiated FDI may be triggered, for example, when a government offers incentives to foreign investors in an attempt to eliminate a balance of payments deficit. A similar, trade-related classification of FDI is adopted by Kojima (1973, 1975, 1985). According to Kojima's classification, FDI is either trade-orientated FDI (which generates an excess demand for imports and excess supply of exports at the original terms of trade) or anti-trade-orientated FDI, which has an adverse effect on trade.

Finally, FDI may be classified into expansionary and defensive types. Chen and Ku (2000) suggest that expansionary FDI seeks to exploit firm-specific advantages in the host country. This type of FDI has the additional benefit of contributing to sales growth of the investing firm at home and abroad. On the other hand, they suggest that defensive FDI seeks cheap labour in the host country with the objective of reducing the cost of production. Chen and Yang (1999) suggested that a multinomial logit model can be used to identify the determinants of the two types of FDI in the case of Taiwan. Their

empirical results indicated that expansionary FDI is influenced mainly by firm-specific advantages such as scale, R&D intensity, profitability and motives for technology acquisition. Defensive FDI, on the other hand, is shown to be influenced by cost reduction motives and the nexus of production networks. Both types of FDI are affected by the characteristics of the underlying industry.

WHAT ARE MULTINATIONAL CORPORATIONS?

Most FDI is carried out by multinational corporations (MNCs) which have become household names. Examples (without any particular order in mind) are Toyota, IBM, Phillips, Nestlé, Sony, Royal Dutch Shell, IBM, GM, Coca-Cola, McDonald's, Daimler-Benz, and Bayer. It is, however, difficult to pinpoint what constitutes an MNC, and there is not even an agreement on what to call these firms. The literature shows various 'labels' for these firms, consisting of the words 'international', 'transnational', or 'global' followed by any of the words 'corporations', 'companies' and 'enterprises'. What is more important is that there is no single definition for an MNC. For example, the United Nations (1973) lists twenty-one definitions for MNCs, or whatever they may be called (the UNCTAD in fact calls them TNCs).

Sometimes, however, a distinction is made between the terms 'international', 'multinational' and 'transnational'. The term 'multinational firm' has evolved from changes in the nature of international business operations. The term 'international business firm' referred traditionally to the cross-border activity of importing and exporting, where goods are produced in the domestic market and then exported abroad, and vice versa. The financial implications of these transactions pertain to the payment process between buyers and sellers across national frontiers. As international operations expand, the international firm may feel that it is desirable, if possible, to expand in such a way as to be closer to foreign consumers. Production will then be carried out both at home and abroad. Thus, a multinational firm carries out some of its production activity abroad by establishing a presence in foreign countries via subsidiaries, affiliates and joint ventures (these terms will be defined later). The financial implications become more significant. The foreign 'arms' of a multinational firm normally have a different base or functional currency, which is the currency of the country where they are located. This setup results in

a greater currency and financial risk in general. As cross-border activity expands even further, the distinction between 'home' and 'abroad' becomes blurred, and difficulties arise as to the identification of the 'home country'. What is created in this case is a 'transnational firm'. It remains the case that the relationship between multinationals and FDI is very simple: firms become multinational (or transnational) when they undertake FDI. Thus, FDI represents an internal organizational expansion by multinationals. In this book, we shall use the term 'multinational corporation' (MNC) generally to imply the firms that indulge in FDI.

The link between FDI and MNCs is so close that the motivation for FDI may be used to distinguish between MNCs and other firms. Lall and Streeten (1977) distinguish among economic, organizational and motivational definitions of FDI. The economic definition places emphasis on size, geographical spread and the extent of foreign involvement of the firm. This definition allows us to distinguish between an MNC and (i) a large domestic firm that has little investment abroad; (ii) a small domestic firm that invests abroad; (iii) a large firm that invests in one or two foreign countries only; and (iv) a large portfolio investor that does not seek control over the investment. Parker (1974) classified 613 of the largest manufacturing firms in the world into 'MPE2', 'MPE1' and 'not MPE' (MPE standing for 'multinational producing enterprise'). According to this classification, MPE2 represents firms with more than five foreign subsidiaries, or more than 15 per cent of total sales produced abroad; MPE1 represents firms that are less globally orientated and have 2–5 subsidiaries or 5–15 per cent of sales produced abroad; and not MPE represents the rest of the firms. The organizational definition takes the size and spread for granted and emphasizes factors that make some firms more multinational than others. These factors pertain to the organization of these firms, centralization of decision-making, global strategy and the ability to act as one cohesive unit under changing circumstances. Finally, the motivational definition places emphasis on corporate philosophy and motivations. For example, an MNC is characterized by a lack of nationalism, and by being concerned with the organization as a whole rather than with any constituent unit, country or operation.

The 1999 *World Investment Report* (UNCTAD, 1999) defines multinational corporations (which it calls transnational corporations) as 'incorporated or unincorporated enterprises comprising parent enterprises and their foreign affiliates'. A parent enterprise or firm is defined as 'an enterprise that controls assets of other entities in

countries other than its home country, usually by owning a certain equity capital stake'. A foreign affiliate is defined as 'an incorporated or unincorporated enterprise in which an investor, who is resident in another economy, owns a stake that permits a lasting interest in the management of that enterprise'. Foreign affiliates may be subsidiaries, associates or branches.[6] UNCTAD (1999) distinguishes between them as follows:

- A subsidiary is an incorporated enterprise in the host country in which another entity directly owns more than a half of the shareholders' voting power and has the right to appoint or remove a majority of the members of the administrative, management or supervisory body.
- An associate is an incorporated enterprise in the host country in which an investor owns a total of at least 10 per cent, but not more than a half, of the shareholders' voting power.
- A branch is a wholly or jointly-owned unincorporated enterprise in the host country, which may take the form of a permanent office of the foreign investor or an unincorporated partnership or a joint venture. A branch may also refer to land, structures, immovable equipment and mobile equipment (such as oil drilling rigs and ships) operating in a country other than the investor's country.

Moreover, the UNCTAD (1999) lists the following facts and figures about multinationals:

1. Multinationals comprise over 500 000 foreign affiliates established by some 60 000 parent firms.[7]
2. The MNC universe comprises large firms mainly from developed countries, but also from developing countries and more recently from the countries in transition.
3. In 1997, the 100 largest non-financial MNCs held US$1.8 trillion in foreign assets, sold products worth US$2.1 trillion abroad and employed six million people in their foreign affiliates.
4. In 1997, the top fifty non-financial MNCs based in developing countries held US$105 billion in foreign assets. Most of these companies belong to Korea, Venezuela, China, Mexico and Brazil.
5. The twenty-five largest MNCs in Central Europe (excluding the Russian Federation) held US$2.3 billion in foreign assets and had foreign stakes worth US$3.7 billion.

6. The value of output under the common governance of MNCs amounts to about 25 per cent of global output, one third of which is produced in host countries. In 1998, foreign affiliate sales were about US$11 trillion.[8]

The question as to what MNCs are has been dealt with in the academic literature. Lall and Streeten (1977) identify the following 'salient features' of MNCs:

1. MNCs are predominant in certain monopolistic or oligopolistic industries characterized by the importance of marketing and technology.
2. The products of MNCs are new, advanced and cater for consumers who have relatively high incomes and sophisticated tastes, and who are responsive to modern marketing techniques.
3. The techniques of production MNCs use are the most advanced in their respective fields.
4. The expansion of an MNC tends to reproduce the oligopolistic conditions of the MNC's domestic market.
5. The maturing of MNCs may bring with it various commercial practices to bolster market dominance.
6. MNCs are attracted by large and growing economies with reasonably stable political conditions.
7. The organizational evolution of MNCs leads to a centralization of functions such as finance, marketing and research.
8. MNCs prefer complete or majority ownership of subsidiaries.
9. The increasing international role of MNCs has important implications for the structure of socio-political power in developed and developing countries.

Some attempts have been made to measures the extent of being 'multinational' according to a set of indicators. Dorrenbacher (2000) proposes a measure based on the following indicators: (i) structural indicators; (ii) performance indicators; and (iii) attitudinal indicators. Structural indicators include the number of countries where the firm is active, the number of foreign subsidiaries, the number of foreign employees, and the number of stock markets on which the firm's shares are listed. Performance indicators include foreign sales and operating income of foreign subsidiaries. The attitudinal indicators include management style and international experience of top management.

Indices (or composite indicators), which are calculated by combining individual indicators, can also be used as measures of multinationalization. These include the following measures:

1. The transnationality index of the UNCTAD. This indicator, which first appeared in UNCTAD's 1995 *World Investment Report*, aims to capture fully the extent of involvement in the world economy. It is based on three different ratios: (i) foreign sales to total sales; (ii) foreign assets to total assets; and (iii) foreign employment to total employment.
2. The transnational spread index of Ietto-Gillies (1998). This index is calculated by multiplying the average of the ratios used to calculate the transnationality index by the number of foreign countries in which a firm is active, as a proportion of the total number of countries where FDI has occurred minus one (the home country).
3. The degree of internationalization scale, which was suggested by Sullivan (1994). This indicator is based on (i) the ratio of foreign sales to total sales; (ii) foreign assets to total assets; (iii) the number of foreign subsidiaries to total subsidiaries; (iv) the international experience of top managers; and (v) the dispersion of international operations.

Empirical studies of the behaviour and characteristics of MNCs attempt to detect the characteristics that distinguish an MNC from purely domestic firms. The variables that have been found to be significant in the earlier literature are R&D expenditure, size of the firm, and foreign trade intensity, although other variables also appeared to be important. Vaupel (1971) obtained evidence showing that US MNCs (as compared with domestic firms): (i) incurred higher R&D as well as advertising expenditure; (ii) showed more net profit; (iii) had higher average sales; (iv) were more diversified; (v) paid higher wages in the USA; and (vi) recorded a higher export/sales ratio. Vernon (1971) reached a similar conclusion using the same data set. Lall (1980), however, found that R&D, economies of scale and the possession of skill advantages favour exports more than foreign production (FDI) by US MNCs, whereas product differentiation promotes more foreign production than exports. Horst (1972a), on the other hand, came to the conclusion that all of these variables can be accounted for by inter-industry differences, so that size remains the only significant distinguishing factor. A similar conclusion

was reached by Bergsten *et al*. (1978). Caves (1971) found strong rank correlation between the extent of product differentiation and the proportion of firms in an industry having foreign subsidiaries.

By using an econometric model of the probability that a firm becomes an MNC, Grubaugh (1987) obtained results supporting the importance of R&D expenditure, product diversity and size as characteristics of MNCs. Grubaugh (1987) tested three hypotheses to explain why firms would choose to become MNCs, based on three views of MNCs. The first view is that an MNC is essentially a firm that engages in capital arbitrage (MacDougal, 1960). The second view is that MNCs are oligopolists that compete by producing in various countries (Hymer, 1976). The third view emphasizes the intangible assets that firms acquire. These views of MNCs imply a certain relationship between whether or not a firm is an MNC and the characteristics of the firm (Dunning, 1977; Rugman, 1981). The capital arbitrage view implies that there is no significant difference between MNCs and domestic firms except the cost of capital and capital intensity. The second view implies the importance of the size of the firm and the diversity of its products. The third view implies the importance of knowledge (hence, R&D expenditure) and goodwill (hence, advertising expenditure). The importance of R&D is emphasized by Petit and Sanna-Randaccio (2000), who show that a firm that invests more in R&D is the one that is an MNC, whereas the rival is an exporter. Hence, they conclude that there is a positive relationship between international expansion and R&D expenditure, and that the latter leads to an increase in the likelihood of international expansion.

What does all of this tell us about the importance of MNCs? Lall and Streeten (1977) answer this question by suggesting that the significance of MNCs lies in the simple fact that they dominate overwhelmingly not only international investment but also international production, trade, finance and technology. They conclude that this domination makes any analysis of the structure of international economic relationships that does not take them into account unrealistic and irrelevant.

APPROACHES TO INTERNATIONAL BUSINESS

FDI is one of several approaches that business enterprises can use to enter foreign markets. The following is a common sequence that firms use to develop foreign markets for their products:

1. Export of the goods produced in the source country.
2. Licensing a foreign company to use process or product technology.
3. Foreign distribution of products through an affiliate entity.
4. Foreign (international) production, which is the production of goods and services in a country that is controlled and managed by firms headquartered in other countries.

Steps 3 and 4 involve FDI. Moving from step 1 to step 4 requires a larger commitment of resources, and in some respects greater exposure to risk. While this sequence may be a chronological path for developing foreign sales, it is not necessary that all four steps are taken sequentially, as some firms jump immediately to step 3 or step 4. UNCTAD (1999) identifies the following characteristics of international production:

1. International production arises when a firm exercises control over an enterprise located abroad, whether through capital investment or through contractual arrangements.
2. Technology flows play an important role in international production.
3. Innovation and research and development are at the heart of the ownership advantages that propel firms to engage in international production.
4. International trade is stimulated by international production because of the trading activities of MNCs.
5. International production generates employment opportunities that are particularly welcome in host countries with high rates of unemployment.
6. Financial flows associated with international production consists of funds for financing the establishment, acquisition or expansion of the foreign affiliates.
7. The capital base of international production, regardless of how it is financed, is reflected in the value of assets of foreign affiliates.

The choice between exporting and FDI depends on the following factors: profitability, opportunities for market growth, production cost levels, and economies of scale. For example, MNCs traditionally have invested in Singapore and Hong Kong because of the low production costs in these countries. For the same reasons, traditionally these countries have exported goods to other countries. Initially, exports precede FDI, but after having become familiar with factor and output

markets in the foreign country, a firm will establish a production facility there. Several motives exist for this change. FDI allows a firm to circumvent actual or anticipated barriers to trade. Another motive is the real appreciation of the domestic currency, which reduces the competitiveness of exports.

Step 2 is licensing, which may be defined as the supply of technology and know-how, or it may involve the use of a trademark or a patent for a fee. It offers one way to circumvent entry barriers to FDI. Under these circumstances licensing offers an opportunity to generate revenue from foreign markets that are otherwise inaccessible. Furthermore, the licence owner may often not have the capital, experience or risk tolerance associated with FDI. Firms prefer FDI to licensing in the case of complex technology, or when the risk of leakage of technological advantage to competitors exists.

Franchising is another form of entering a foreign market under contractual agreements. Companies with brand name products (Kentucky Fried Chicken and Burger King, for example) move offshore by granting foreigners the exclusive right to sell the product in a designated area. The parent company provides the technical expertise pertaining to the production process as well as marketing assistance for an initial fee and subsequent royalties related to turnover. UNCTAD (1999) defines royalties and licensing fees as 'receipts and payments of residents and non-residents for (i) the authorised use of intangible, non-produced, non-financial assets and proprietary rights such as trade marks, patents, processes, techniques, designs, manufacturing rights, franchises, etc.; and (ii) the use, through licensing agreements of produced originals or prototypes, such as manuscripts, films, etc.'

FDI may take one of three forms: greenfield investment, cross-border mergers and acquisitions (M&As), and joint ventures. Greenfield investment occurs when the investing firm establishes new production, distribution or other facilities in the host country. This is normally welcomed by the host country because of the job-creating potential and value-added output. Sometimes, the term 'brownfield investment' is used to describe a situation where investments that are formally an acquisition resemble greenfield investment. This happens when the foreign investor acquires a firm but replaces almost completely the plant and equipment, labour and the product line. This concept has been used most to describe acquisitions in transition economies (Meyer and Estrin, 1998).

FDI may occur via an acquisition of, or a merger with, an established firm in the host country (the vast majority of M&As are indeed acquisitions rather than mergers). This mode of FDI has two advantages over greenfield investment: (i) it is cheaper, particularly if the acquired project is a loss-making operation that can be bought cheaply; and (ii) it allows the investor to gain a quick access to the market. Firms may be motivated to engage in cross-border acquisitions to bolster their competitive positions in the world market by acquiring special assets from other firms or by using their own assets on a larger scale. A large number of M&As fail in the sense that the firms engaging in this activity do not produce better results in terms of share prices and profitability than those firms that do not indulge in this activity. However, the extent of failure depends crucially on the success criteria, which means that the failure rate may be high or low, depending on these criteria (Hopkins, 1999).

Whether a firm would choose M&As or greenfield investment depends on a number of firm-specific, host country-specific and industry-specific factors, including the following (UNCTAD, 2000):

1. Firms with lower R&D intensity are more likely to indulge in M&As than those with strong technological advantages.
2. More diversified firms are likely to choose M&As.
3. Large MNCs have a greater tendency to indulge in M&As.
4. There is weak support for the proposition that advertising intensity leads to more acquisitions.
5. Cultural and economic differences between the home country and the host country reduce the tendency for M&As.
6. Acquisitions are encouraged by capital market imperfections and financial crises.
7. MNCs with subsidiaries in the host country prefer acquisitions.
8. The tendency towards M&As depends on the supply of target firms.
9. Slow growth in an industry favours M&As.

McCann (2001) presented a model in which he explained cross-border acquisitions involving UK firms during the period 1987–95 using panel data analysis. He found that models which explain cross-border acquisitions through capital market imperfections are inadequate, but he also found the exchange rate, stock prices and corporate tax differentials to be important determinants. The data on M&As show that acquisitions dominate the scene, as less than 3 per cent of

cross-border M&As by numbers are in fact mergers. In reality, even when mergers are supposedly between two equal partners, most are, in reality, acquisitions. For practical purposes, M&As are actually mergers.

Cross-border acquisition of businesses is a politically sensitive issue, as most countries prefer to retain local control of domestic firms. It follows that, while countries may welcome greenfield investments, foreign firms' bids to acquire domestic firms are often resisted, and sometimes even resented. The underlying argument here is that M&As are less beneficial than greenfield FDI, and may even be harmful, because they do not add up to productive capacity but rather represent a transfer of ownership that may be accompanied by layoffs or the termination of some beneficial activities. If mergers and acquisitions take place in some sensitive areas, such as the media, then it may seem (perhaps justifiably) like a threat to the national culture or identity.

Whether or not cross-border acquisitions produce synergetic gains, and how such gains are divided between acquiring and target firms, are important issues from the perspective of shareholders' welfare and public policy. Synergetic gains are obtained when the value of the combined firm is greater than the stand-alone valuations of the individual (acquiring and target) firms. If cross-border acquisitions generate synergetic gains, both the acquiring and the target firms' shareholders gain wealth at the same time. In this case, one can argue, both from a national and a global perspective, that cross-border acquisitions are mutually beneficial and thus should not be thwarted. Moreover, it is sometimes argued that the perceived negative effects of M&As may materialize in the short run only, while several benefits emerge in the long run. The latter include new sequential investments, transfer of new technology, and the generation of employment.[9]

Synergetic gains may or may not arise from cross-border acquisitions, depending on the motive of acquiring firms. In general, gains will result when the acquiring firm is motivated to take advantage of market imperfections such as mispriced factors of production, or to cope with trade barriers. Several studies have investigated the impact of cross-border acquisitions. For example, Doukas and Travlos (1988) investigated the impact of international acquisitions on the stock prices of US bidding firms. The results show that shareholders of the bidding firms experience significant positive abnormal returns when firms expand into new industries and markets. Harris and

Ravenscraft (1991) studied shareholder wealth gains for US firms acquired by foreign firms. They concluded that US target firms experience higher wealth gains than when they are acquired by US firms.

FDI can also take the form of joint ventures, either with a host country firm or a government institution, as well as with another company that is foreign to the host country. One side normally provides the technical expertise and its ability to raise finance, while the other side provides valuable input through its local knowledge of the bureaucracy as well as of local laws and regulations. Buckley and Casson (2000b) present a model that explains the formation of joint ventures in terms of nine distinct factors: (i) market size; (ii) pace of technological change; (iii) interest rates; (iv) cultural distance; (v) protection of independence; (vi) missing patent rights; (vii) economies of scope; (viii) technological uncertainty; and (ix) economies of scale. This model allows them to arrive at detailed predictions about how the formation of joint ventures varies with industries, between industries, across countries and over time.

HISTORY OF FDI

In the nineteenth century, foreign investment was prominent, but it mainly took the form of lending by Britain to finance economic development in other countries as well as the ownership of financial assets. However, a recent article by Godley (1999) analyses some cases of FDI in British manufacturing industry prior to 1890, and shows that from 1890 onwards the bulk of FDI was in the industrial goods sector. Godley also shows that investors in Britain prior to 1890 were primarily in the consumer goods sector, and that they mostly failed because they were narrowly focused and driven entirely by concern about enhancing access to the British market. One exception was the Singer Manufacturing Company. As a result of its enthusiastic commitment to FDI, the company emerged as the world's first modern MNC and was one of the largest firms in the world by 1900.

In the interwar period of the twentieth century, foreign investment declined, but direct investment rose to about a quarter of the total. Another important development that took place in the interwar period was that Britain lost its status as the major world creditor, and the USA emerged as the major economic and financial power. In the post-Second World War period, FDI started to grow, for two reasons. The first was technological – the improvement in transport and com-

munications which made it possible to exercise control from a distance. The second reason was the need of European countries and Japan for US capital to finance reconstruction following the damage inflicted by the war. Moreover, there were some US tax laws that favoured FDI. By the 1960s, all these factors were weakening to the extent that they gave rise to a reversal of the trend towards growth in FDI. First, various host countries started to show resistance to the US ownership and control of local industry, which led to a slowdown of outflows from the USA. Second, host countries started to recover, initiating FDI in the USA, and leading to a decline in the net outflow from the USA. The 1970s witnessed lower FDI flows, but Britain emerged as a major player in this game as a result of North Sea oil surpluses and the abolition of foreign exchange controls in 1979.

The 1980s witnessed two major changes and saw a surge in FDI. The first change was that the USA became a net debtor country and a major recipient of FDI with a negative net international investment position. One of the reasons for this development was the low saving rate in the US economy, making it impossible to finance the widening budget deficit by resorting to the domestic capital market, and giving rise to the need for foreign capital, which came primarily from Japan and Germany. Another reason was the restrictive trade policy adopted by the USA. The other major change in the 1980s was the emergence of Japan as a major supplier of FDI to the USA and Europe. Motivated by the desire to reduce labour costs, Japanese direct investment also expanded in South East Asia.

The surge in FDI in the 1980s is attributed to the globalization of business. It is also attributed by Aizenman (1992) to the growing concern over the emergence of managed trade. Moreover, it is argued that FDI benefits both MNCs and host country, and this is why there has been tolerance towards FDI. Another reason for the surge in FDI is the increase in FDI inflows to the USA as a result of the depreciation of the US dollar in the second half of the 1980s. The total flows of FDI from industrial countries more than quadrupled between 1984 and 1990.

In the period 1990–2, FDI flows fell as growth in industrial countries slowed, but a strong rebound subsequently took place. This rebound is attributed to three reasons: (i) FDI was no longer confined to large firms, as an increasing number of smaller firms became multinational; (ii) the sectoral diversity of FDI broadened, with the share of the service sector rising sharply; and (iii) the number of countries that were outward investors or hosts of FDI rose considerably.

Moreover, the 1990s brought considerable improvements in the investment climate, triggered in part by the recognition of the benefits of FDI. The change in attitude, in turn, led to a removal of direct obstacles to FDI and to an increase in the use of FDI incentives. Continued removal of domestic impediments through deregulation and privatization was also widespread.

Another important feature of the 1990s was the decline in the importance of Japan as a source of FDI, caused by the burst of the Japanese bubble economy. The late 1990s were characterized by cross-border M&As (which were motivated by deregulation and enhanced competition policy) as the driving force behind FDI. Moreover, the trend towards the liberalization of regulatory regimes for FDI continued. By the end of 1998, the number of treaties for the avoidance of double taxation had reached a total of 1871. In 1998 and 1999 some changes were introduced to (host) government policies on FDI, strengthening the trend towards the liberalization, protection and promotion of FDI (UNCTAD, 2000).[10] It seems that this trend will continue for a long time to come, which means that the growth of FDI will be robust in the foreseeable future.

RECENT TRENDS

In this section we examine briefly the recent trends in FDI. A more detailed account of the global and regional trends up to 1999 can be found in the 2000 *World Investment Report* (UNCTAD, 2000). Before we examine the figures, it may be worthwhile to try to anticipate what the pattern has been like on the basis of some theoretical considerations. Lipsey (2000) suggests that if FDI flows represented mainly responses to differences among countries in the scarcity and price of capital, countries would tend mainly to be sources or recipients of FDI (capital-surplus and capital-deficit countries respectively). Given the size of the economy, the levels of outflows and inflows should therefore be negatively related. This relationship is also obtained by viewing FDI flows as depending on economic conditions. If the economy is in a boom, FDI inflows will increase and FDI outflows will decrease. And if the economy is in a slump, then FDI inflows will decrease and outflows will increase. Hence, FDI outflows and inflows should be correlated negatively. Lipsey (1999, 2000) shows that this is not the case. The positive relationship is attributed to the possibility that economic factors that encourage inward flows also encourage

outward flows. Lipsey also suggests that the coexistence of outward and inward stocks of FDI arises from an alteration between inflows and outflows.

FDI flows comprise the capital provided (either directly or through related enterprises) by a foreign direct investor to an FDI enterprise, or capital received from an FDI enterprise by a foreign direct investor. From the perspective of a particular country, FDI flows may be inward (when a foreign country invests in the country in question) or outward (when the home country invests abroad). FDI flows consist of the following items:

- Equity capital, which is the foreign investor's purchases of shares in an enterprise in a foreign country.
- Reinvested earnings, which comprise the investor's share of earnings not distributed as dividends by affiliates or remitted to the home country, but rather reinvested in the host country.
- Intra-company loans, which refer to short-term or long-term borrowing and lending of funds between the parent company and its affiliates.

FDI inflows and outflows during the period 1994–9 are shown in Table 1.1. These figures are calculated on a net basis; that is, as capital transactions' credits less debits between investors and their affiliates.

Table 1.1 FDI inflows and outflows (US$bn)

Region/country	*1988–93**	*1994*	*1995*	*1996*	*1997*	*1998*	*1999*
Inflows							
European Union	78.5	76.9	114.4	108.6	128.6	248.7	305.1
USA	44.8	45.1	58.8	84.5	105.5	186.3	275.5
Japan	0.7	0.9	0.4	0.2	3.2	3.2	12.7
Developing countries	46.9	104.9	111.9	145.0	178.8	179.5	207.6
World	190.6	256.0	331.9	377.5	473.1	680.1	865.5
Outflows							
European Union	107.2	120.7	159.0	182.3	223.7	425.5	509.8
USA	39.3	73.3	92.1	84.4	99.5	146.1	150.9
Japan	32.5	18.1	22.5	23.4	26.1	24.2	22.7
Developing countries	23.5	42.1	50.3	57.8	64.3	33.1	65.6
World	221.4	282.9	357.5	390.8	471.9	687.1	799.9

Note: *Annual average.
Source: UNCTAD (2000).

Net decreases in assets or net increases in liabilities are credits (recorded with a positive sign on the balance of payments), while net increases in assets or net decreases in liabilities are debits (recorded with negative signs on the balance of payments). In this table, the negative signs are deleted for convenience. It is obvious that FDI outflows and inflows are positively correlated, more so in the case of the European Union (EU). During the 1994–9 period, FDI inflows to the EU grew at an average annual rate of 31.7 per cent, whereas outflows grew at a rate of 33.4 per cent. Inflows grew much faster than outflows in the case of Japan and the USA, reflecting the attractiveness of the USA as a destination for FDI. In the case of Japan, however, the disparity between the growth rates of inflows and outflows is a reflection of the fact that Japan traditionally has been a source rather than a recipient country of FDI. Thus, the growth in inflows is measured relative to a very low initial value (US$0.7 billion in 1994). It seems, however, that Japan has more recently become a major recipient country, with inflows totalling US$12.7 billion in 1999. In that same year, the EU accounted for 35.3 per cent of FDI inflows and 63.7 per cent of total outflows, which means that the EU is the largest source region. Japan's economic problems in the 1990s may explain the declining role of Japan as a source of FDI, which in 1999 accounted for 2.8 per cent of total outflows (less than the contribution of developing countries). In 1999, the USA, like the developing countries, was a net recipient of FDI (capital importer), whereas the EU and Japan were capital exporters. While in theory total outflows should be equal to total inflows, this is not the case in practice because of measurement errors. Remember that these statistics were obtained initially from national sources.

Now we turn to FDI stocks, which represent the value of the share of their capital and reserves (including retained profits) attributable to the parent enterprise, plus the net indebtedness of affiliates of the parent enterprise. Like FDI flows, stocks can also be inward or outward. Table 1.2 reports FDI stocks for selected years. FDI stocks are estimated either by cumulating FDI flows over a period of time, or by adding flows to an FDI stock that has been obtained for a particular year from national office sources or the IMF data series on assets and liabilities.

During the period 1980–99, the growth of FDI stocks echoed that of FDI flows. First, there is high correlation between inward and outward stocks. The growth rates of inward and outward stocks during this period were very close. In the case of the EU, the inward stock grew at an annual rate of 12.1 per cent, while the outward stock grew

Table 1.2 FDI inward and outward stocks (US$bn)

Region/country	*1980*	*1985*	*1990*	*1995*	*1998*	*1999*
Inward stocks						
European Union	185.7	236.4	723.5	1050.3	1451.2	1652.3
USA	83.1	184.6	394.9	535.6	811.8	1087.3
Japan	3.3	4.7	9.9	33.5	26.1	38.8
Developing countries	121.2	218.1	377.4	739.5	1241.0	1438.5
World	495.2	763.4	1761.2	2743.4	4015.3	4772.0
Outward stocks						
European Union	212.6	292.7	789.4	1303.2	1920.4	2336.6
USA	220.2	251.0	430.5	699.0	980.6	1131.5
Japan	19.6	44.0	201.4	238.5	270.0	292.8
Developing countries	16.3	32.4	81.9	258.3	403.9	468.7
World	523.2	707.1	1716.4	2870.6	4065.8	4759.3

Source: UNCTAD (2000).

at a rate of 13.4 per cent. Even in the case of Japan, the growth rates were close (13.9 and 15.3 per cent, respectively). In 1999, the EU accounted for 34.6 per cent of the world inward stocks and 49.1 per cent of the world outward stocks. In terms of the net asset value (outward stocks minus inward stocks), the USA and developing countries were in deficit, whereas Japan and the EU were in surplus. Again, the world inward and outward stocks are not equal because of measurement errors.

Table 1.3 reports some statistics on cross-border mergers and acquisitions. These figures are published by UNCTAD based on data provided by KPMG Corporate Finance. By comparing the figures in Table 1.1 with those in Table 1.3, it is obvious that M&As have become the dominant form of FDI. In 1999, M&As accounted for over 80 per cent of total FDI inflows, and over 90 per cent of total FDI outflows. In developing countries, the two ratios were 8.9 per cent and 63 per cent, respectively. For the EU, however, M&As were dominant. This shows that whether FDI takes the form of greenfield investment or M&As depends in part on the level of development in the host country, since this factor determines the supply of target firms.

The statistics that we have considered on FDI flows, stocks and M&As serve to give a general indication as to what is happening, but one should take these statistics with a large pinch of salt. Because of measurement errors and accounting valuation problems, which apply particularly to stocks, individual figures tend to be inaccurate. As we

Table 1.3 Cross-border mergers and acquisitions (US$bn)

Region/country	*1993*	*1994*	*1995*	*1996*	*1997*	*1998*	*1999*
Sellers							
European Union	38.5	55.3	75.1	81.9	114.6	187.9	344.5
USA	19.9	44.7	53.2	68.1	81.7	209.6	233.0
Japan	0.09	0.8	0.5	1.7	3.1	4.0	15.9
Developing countries	12.8	14.9	16.0	34.7	64.6	80.8	64.6
World	83.1	127.1	186.6	227.0	304.9	531.7	720.1
Purchasers							
European Union	40.5	63.9	81.4	96.7	142.1	284.4	497.7
USA	21.4	28.5	57.3	60.7	80.9	137.4	112.4
Japan	1.1	1.1	3.9	5.7	2.7	1.3	9.8
Developing countries	10.4	10.2	12.8	28.1	32.5	19.2	41.3
World	83.1	127.1	186.6	227.0	304.9	531.7	720.1

Source: UNCTAD (2000), based on data provided by KPMG.

have seen, total inward flows are not equal to total outward flows, and the same is true for stocks. But since planet Earth does not yet have financial relationships with other planets from our solar system or from distant galaxies, inward flows (stocks) should be equal to outward flows (stocks). Moreover, by comparing the figures in Table 1.1 with those in Table 3.1, we find that the value of M&As in EU countries in 1999 was greater than the FDI inflows into the region, which does not make sense. Several other discrepancies can be observed. There are also problems with the measurement of M&As. For example, M&A statistics are compiled either on the basis of announcement or on the basis of completion. The treatment of additional acquisitions may also differ from one collecting agency to another. Moreover, the available data on M&As include portfolio investment, in which case it is necessary to extract transactions corresponding to FDI (in terms of control).[11]

It seems that, as economists, we have to live with the measurement errors problem, but the consolation is that these statistics give a good picture of the general trends in FDI. The next step is to study what determines FDI, and this is the subject matter of Chapter 2.

3 The Effects of Foreign Direct Investment

FDI involves the transfer of financial capital, technology and other skills (managerial, marketing, accounting, and so on) as we have seen so far. This process gives rise to costs and benefits for the countries involved: the investing country (the source of the investment) and the host country (the recipient or the destination of the investment). It is not clear, however, what costs are borne and what benefits are enjoyed by the two countries, at least not quantitatively. There is even a fundamental disagreement on what constitutes the costs and benefits of FDI from the perspectives of the two countries. This disagreement is indicated by the big gap between those holding pro-globalization, free-market views, and those with anti-globalization, anti-market views. Moreover, the division of welfare gains between the host country and the investing country does not only depend on given market prices, but also on the relative strength of the two countries in bargaining over the terms of the agreement governing a particular FDI project. Nevertheless, one country's losses are not necessarily the other country's gains. Kindleberger (1969), for example, argues that the relationship arising from the FDI process is not a zero-sum game. *Ex ante*, both countries must believe (justifiably or otherwise) that the expected benefits to them must be greater than the costs to be borne by them, because an agreement would not otherwise be reached, and the underlying project would not be initiated. However, believing in something *ex ante* is no guarantee that it will materialise *ex post*.

One way to explain the effect of FDI is to use the conventional multiplier process, but an attempt like this will be made more difficult by the qualitative differences between domestic investment and FDI, which are bound to have different effects from each other. One reason for the differences in the effects is that FDI is controlled by parties over which there is limited local jurisdiction. MNCs are less dependent on their host countries or countries of origin than local firms, and this makes them difficult to control. The fact that the investor undertaking an FDI project is foreign to the host country creates

economic, political and social effects that impinge upon the costs and benefits of FDI.

The effects of FDI on the host country can be classified into economic, political and social effects. The basic presumption that is found in the literature, based on the principles of neoclassical economics, is that FDI raises income and social welfare in the host country unless the optimum conditions are distorted significantly by protection, monopoly and externalities (Lall and Streeten, 1977). We have to bear in mind that MNCs operate in such a way as to maximize profits worldwide, and in the process they shift resources to areas where returns are high, and buy inputs where their prices are low (after all, they are profit maximizers). On the surface, it would seem that this is some sort of efficiency that should lead to an increase in world welfare. However, the problem is that MNCs exist and operate primarily because of market imperfections, which casts doubt on the validity of the conclusion that FDI leads to an increase in welfare. Unless all markets are perfect, growth in one sector may not be beneficial: improved resource allocation has to be judged against increases in market imperfections. If we assume, for the sake of argument, that markets are perfect and that there are constant returns to scale, then if capital is allowed to move freely, it would flow from a low-return country to a high-return country. This causes the rate of return on capital to fall in the high-return country, and rise in the low-return country, and this would involve a gain in world output (for a diagrammatic illustration, see Winters, 1991, pp. 227–8). But even in this case, some distributional changes occur between labour and capital. In any case, it seems that the conventional wisdom that FDI is always welfare-improving is no longer a conventional wisdom. Leahy and Montagna (2000b) challenge this conventional wisdom, and show that direct product market competition makes welfare losses more likely, because MNCs capture market shares from the indigenous firms.

The economic effects of FDI include the implications for (macro and micro) economic variables such as output, the balance of payments, and market structure. The political effects include the question of national sovereignty, as the sheer size of the investing MNCs may jeopardize national independence (who runs the host country, the prime minister or the CEO of the investing MNC?).[1] The social issues are concerned mainly with the creation of enclaves and a foreign elite in the host country, as well as the cultural effects on the local population (for example, customs and tastes).[2] Naturally, social issues are

more likely to arise when there are significant economic, social, and cultural differences between the investing and host countries. For example, it is plausible to think that the social and cultural impact of Australian FDI in New Zealand, if any, will be less than that of Australian FDI in Malaysia, or American FDI in Saudi Arabia.[3] In this chapter (and this book) we are concerned mainly with the economic effects of FDI.

The economic effects of FDI can be classified into macro effects and micro effects. The usual convention in analysing the macro effects of FDI is to treat it as a rise in foreign borrowing. If there is unemployment and capital shortage (as is typically the case in developing countries) such borrowing leads to a rise in output and income in the host country. FDI will, under these conditions, have a beneficial effect on the balance of payments, but an indeterminate effect on the terms of trade (depending on whether the impact of increased output falls on import substitutes or exports). Formal analysis of the macro effects of FDI in a neoclassical framework can be traced back to the analysis by MacDougal (1960), who used partial-equilibrium, comparative-static analysis to show that the host country would gain mainly through taxes on foreign profits. Real wages, the argument goes, would also rise at the expense of profits because of the declining marginal productivity of capital. This kind of analysis has been extended in several ways, including: (i) the analysis of optimum taxes; (ii) dynamic growth models with foreign investment and trade using pure trade theory; (iii) comparative statics general equilibrium models; and (iv) comparative statics trade theory incorporating capital movements.

The micro effects of FDI pertain to structural changes in the economic and industrial organization. For example, an important issue is whether FDI is conducive to the creation of a more competitive environment, or conversely to a worsening of the monopolistic and/or oligopolistic elements in the host economy. In general, the micro effects pertain to individual firms and individual industries, particularly those that are closely exposed to, and associated with, FDI. Markusen and Venables (1997) put forward the idea that the (micro) effects of FDI on the host country may operate through many different channels. They present a simple analysis of two of these channels: product market competition, through which MNCs may substitute for domestic firms, and linkage effects, through which MNCs must be complementary. Hence, it is possible for FDI to act as a catalyst, leading to the development of local industry, which may in turn become so strong as to reduce both the relative and absolute

position of MNCs in the industry. This analysis, they argue, fits well with some of the case study literature on South East Asian economies.

While we may consider the costs and benefits of FDI to be borne by, or accruing to, the two countries (the investing country and the host country), we shall refer invariably to the MNC as the 'investing party' and the host country as the 'recipient party'. Furthermore, most of our discussion in this chapter will deal with the effects of FDI on host developing countries, although some reference will be made to the effects on the economy of the investing country. The emphasis on the effects on developing countries results from the ever-increasing importance of FDI in the theory and practice of economic development. Recall that FDI is sometimes hailed as the only salvation for developing countries, and the only way out of the vicious circle of poverty. Our conclusion will not be that sanguine: while FDI has its benefits, it also has its costs, and hence the effects on developing countries may not be favourable in all cases all the time.

We shall now discuss the economic effects of FDI in turn. Just like the theories of FDI, there is a significant overlap in the discussion of these effects. For example, the provision of capital, as performed by FDI, overlaps significantly with the effect on the balance of payments and the effect on output. Moreover, technology is believed to be the main conduit of the effect of FDI on growth and productivity, yet we shall examine the effects of FDI on growth, productivity and technology separately. And the effect on the welfare of the host country encompasses some of the individual effects. While the main body of this chapter contains a brief review of related empirical evidence, the findings of recent empirical studies are tabulated in the appendix (Table A3.1).

THE PROVISION OF CAPITAL

The two-gap model, which is often used in development economics, shows that developing countries typically encounter the problem of increasing their saving to match their investment needs, and that of financing imports through export earnings (see, for example, McKinnon, 1963). The first problem arises from the saving gap (the difference between investment and saving), whereas the second problem arises from the foreign exchange gap (the difference between imports and exports).[4] It is often argued that FDI contributes to filling these two gaps, not only because MNCs have better access to financial

markets, but also because: (i) FDI by a particular MNC in a particular project may encourage other MNCs to participate in the same project; (ii) such an action may encourage the flow of official development aid from the investor's home country;[5] and (iii) by offering locals attractive investment opportunities, FDI may mobilize domestic saving. By filling (or contributing to the filling of) the foreign exchange gap, FDI obviously has a positive effect on the balance of payments, which we shall examine shortly. It is arguable, therefore, that the net impact of FDI on the quantum of capital flows to developing countries is usually positive, because it leads to an increase in the inflow of financial resources available for investment. In this respect, FDI offers certain advantages over other sources of foreign finance. First, it is more stable than other financial flows. Second, FDI inflows represent a long-term commitment to the host country.[6] Third, FDI is easier to service than commercial loans, since profits tend to be linked to the performance of the host economy (that is, they are procyclical).

Razin *et al*. (1999a) examined the role of FDI in the financial markets of the host country. They argue that, in the absence of a well-developed domestic credit market (in which case, domestic savings cannot be efficiently channelled into domestic investment) FDI can play a double role. It provides a vehicle for reviving the domestic market through which domestic savings can be channelled to finance domestic investment, and it supplies foreign savings on top of domestic savings to finance domestic investment. The second role provides the traditional gains from trade to the home country. The first role is by no means costless: as the equity market is characterized by asymmetric information, it does not always generate the correct signals about the social rates of return on domestic capital. As a result, Razin *et al*. argue, there are some welfare losses that offset some or all of the gains stemming from the mere channelling of domestic savings into domestic investment. When a well-developed domestic credit market exists (through which domestic savings can be channelled into domestic investment even in the absence of an equity market) then the first role played by FDI does not generate any gain. They conclude that, when FDI can be levered domestically, the traditional gains from trade associated with the second role of FDI is curtailed severely. As a result, the total net effect of FDI on the welfare of the domestic economy could well be negative.

Lall and Streeten (1977) cast doubt on the ability of FDI to perform the function of providing capital, for at least three reasons. First, direct investment is a relatively expensive source of foreign capital.

Second, the actual capital inflow provided by the MNC may not be very large (FDI may be financed by borrowing in the host country). Indeed, MNCs can, through their market power, raise cheap funds and crowd out other socially desirable activities in the host country. Third, the capital contribution of the MNC may take the form of machinery or capitalized intangibles, such as know-how and goodwill. For these reasons, FDI provides little, and expensive, capital.

THE EFFECT OF FDI ON OUTPUT AND GROWTH

One of the most important aspects of FDI is its effect on output and therefore growth in the host country. This effect naturally is more important for developing countries, where inward investment is viewed as a means of boosting economic development. For the effect on output to materialize, a necessary condition is an increase in the capital stock of the host country as a result of the investment or, in the case of a take-over, a more efficient utilization of existing resources. The output effect will be less pronounced if FDI takes the form of a take-over (M&As).

Theories of economic growth and development focus on the increase in real per capita income and relate this increase to certain major factors such as capital accumulation, population growth, technological progress and the discovery of new natural resources. The various ways in which these factors are interrelated have given rise to different theories of economic development. In these theories, however, capital accumulation is seen as the driving force behind faster growth.[7] It is then obvious that FDI, by affecting capital accumulation, ought to be capable of influencing economic development.

Of course, technology (or technological progress) also plays a big role. In contrast to the traditional Solow growth model, where technological change is assumed to be exogenous, the recent growth literature highlights the dependence of economic growth on the state of domestic technology relative to that of the rest of the world. Thus, growth rates in developing countries are in part explained by a catch-up process in the level of technology. In a typical model of technology diffusion, the growth rate of a developing country depends on the extent of the adoption and implementation of new technologies that are already in use in leading countries. According to this view, what matters is the adoption and adaptation of foreign technology, a proposition that is supported strongly by the Japanese experience.

The effect of FDI on output can be explained in terms of the multiplier model, but it is rather difficult to quantify the multiplier associated with FDI. This is because, apart from the leakages associated with domestic investment (such as taxes and imports), FDI has leakages of its own, such as the import content (which represents foreign claims on domestic output) and remittances (in the form of dividends, interest payments, fees and royalties), which again represent claims on local output. On the other hand, FDI may be associated with import substitution, a factor that complicates the matter further.

The effect of FDI on the level, composition and growth of the output of the host country also depends to a large extent on the macroeconomic policy in operation in that country. In general, it seems that FDI can exert an impact on the output of the host country if it is possible to absorb surplus resources and/or improve efficiency through alternative allocations. Specifically, however, the following outcomes are possible:

1. If the host government pursues a macroeconomic policy that always achieves full employment, then inward FDI would not affect the size of output, provided it is as efficient as any domestic means of resource utilization.
2. If FDI absorbs resources that would otherwise remain unemployed, then the output generated by FDI net of remittances represents a net gain to the output of the host country.
3. If FDI is capable of improving the efficiency of domestic resources by shifting them from less efficient to more productive sectors of the economy, then domestic output would rise.

Lall and Streeten (1977) argue that the domination of a developing economy by an MNC may be detrimental economically to growth and development, for at least three reasons. First, the MNC's activity may lead to a lower rate of accumulation domestically, because a proportion of the profits generated by this activity is repatriated rather than invested in the host country. Second, the MNC's presence may lead to some adverse developments, such as a greater incidence of undesirable practices (for example, derogatory transfer pricing), or weaken the control over economic policy.[8] Third, the MNC may affect the market structure adversely, making it less competitive, as we shall see later.

What is the empirical evidence on this issue? We start with the evidence presented by Borensztein *et al*. (1995). This piece of evidence

is based on a model of endogenous growth in which the rate of technological progress is the main determinant of long-term economic growth. It is assumed that technological progress takes place through the process of capital deepening that results from the introduction of new varieties of capital goods. MNCs typically are portrayed as possessing more advanced knowledge and to be capable of introducing new capital goods at a lower cost. However, it is argued that the application of more advanced technology requires the presence of a sufficient level of human capital in the host country, which means that if this condition is not satisfied, then the absorptive capacity of the developing host country will be very limited. Hence, the model highlights the role of both the introduction of more advanced technology and the requirement of absorptive capacity in the host country as determinants of economic growth. This line of reasoning brings to the surface the issue of complementarity between FDI and human capital in the process of economic growth.

Borensztein *et al*. (1995) test the effect of FDI on economic growth in a cross-country regression framework, utilizing data on FDI flows from industrial countries to sixty-nine developing countries over two decades. Their results suggest the following conclusions:

1. FDI is an important vehicle for the transfer of technology, contributing relatively more to growth than does domestic investment.
2. For FDI to produce higher productivity than domestic investment, the host country must have a minimum threshold stock of human capital.
3. FDI has the effect of increasing total investment in the economy more than proportionately, which suggests the predominance of complementarity effects with domestic firms.

Feldstein (1994) examines the effect of outward FDI on the national incomes of the home and host countries in the presences of taxes and tax credits (see Chapter 6). He argues that the national income of the home (investing) country depends on the relative importance of two factors acting in different directions: the loss of tax revenue to the foreign (host) government, and the increased use of foreign debt. He develops an analysis of these two factors in the presence of a segmented international capital market, in which most national saving remains in the country where saving arises.

The idea underlying Feldstein's thinking in this respect is very simple. Firms that invest abroad pay taxes on the profit of their

foreign subsidiaries to the governments of the host countries, which means that one consequence of outward FDI is the loss of revenue by the home government to the host government. If investing firms receive tax credit for the taxes that they pay to the foreign government, they (the firms) will invariably be indifferent between taxes paid to the host government and those paid to the home government. The tax credit causes the firms to invest abroad until the after-tax rate of return on the foreign investment is equal to the after-tax rate of return on domestic investment. Since the home country receives the full before-tax return on domestic investment (in the form of taxes on the income realized by the investing firms) but only the after-tax return on the investment of foreign subsidiaries, critics of the foreign tax credit argue that it causes an excessive amount of FDI and a reduction in the home country's national income.

Feldstein, however, questions this argument, on the grounds that it fails to take into account the fact that firms that invest abroad increase their use of foreign debt as they increase the extent of their FDI. Although each firm's overall leverage may be unaffected by the extent of its FDI, the home country as a whole benefits from the use of the additional low-cost credit supplied by foreign creditors. While the debt capital may in theory be available to the investing firms (and therefore to the benefit of the home country) through international portfolio investment, the evidence on segmentation of global capital markets implies that this would not occur in practice. For example, Feldstein and Horioka (1980) have shown, in a frequently-cited paper, that the saving generated in a country tends to stay in that country.[9] Hence, the advantage of FDI, according to Feldstein's reasoning, is that it allows the investing firms and the source country of FDI to utilize foreign debt capital without requiring that capital to cross national borders.

But why would firms increase their foreign borrowing as they expand their FDI? Feldstein answers this question by listing a number of reasons. First, such borrowing is one way to hedge the value of foreign profits caused by fluctuations in exchange rates. Second, there may be some restriction on the amount of domestic interest that firms can deduct in calculating domestic taxable income when they have overseas operations. Third, firms may be able to borrow at a lower cost in countries where collateral is available. Fourth, local debt can be used as an anti-expropriation device (see Chapter 5). Irrespective of the relative importance of these factors, foreign borrowing is useful as long as the real after-tax cost of foreign borrowing is less than the

after-tax return on the foreign assets acquired with these funds. The net effect of an additional dollar of outward FDI on the national income of the home country thus depends on the relative importance of the tax paid to the foreign government, and the advantage obtained by using foreign debt. The evidence presented by Feldstein shows that the favourable leverage effect is likely to outweigh the loss of revenue to the foreign country. When a firm equates the after-tax rate of return on the domestic investment and investment in a foreign subsidiary, the national rate of return to the home country is higher on foreign investment than on domestic investment. Feldstein's calculations imply that a dollar of cross-border outward FDI raises the present value of US national income by US$1.72.

THE EFFECT OF FDI ON EMPLOYMENT AND WAGES

There is a relationship between investment and employment. In his *General Theory*, Keynes (1936) suggested the existence of a direct relationship between investment and employment. However, there is still considerable divergence in views among economists about the employment effects of FDI (see, for example, Pugel, 1985). Baldwin (1995) argues that this debate encompasses three key issues: (i) the extent to which FDI substitutes for domestic investment; (ii) the extent to which FDI stimulates increases of exports of intermediate goods and capital goods; and (iii) whether FDI involves the construction of new plants or simply the acquisition of existing facilities. In general, the employment effects of FDI may be summarized as follows:

1. FDI is capable of increasing employment directly, by setting up new facilities, or indirectly by stimulating employment in distribution.
2. FDI can preserve employment by acquiring and restructuring ailing firms.
3. FDI can reduce employment through divestment and the closure of production facilities.

The available evidence suggests that the effect of FDI on employment is low. Vaitsos (1976) analysed the employment effects of MNCs with reference to four characteristics: scale, concentration, foreignness, and transnationality. He produced evidence indicating that the overall employment effects of the activities of MNCs on the host

countries has been relatively small. Tambunlertchai (1976) evaluated the contribution of foreign firms to the host country by reference to four criteria: (i) contribution to national income; (ii) creation of employment; (iii) utilization of domestic resources; and (iv) earnings and savings of foreign exchange. His empirical evidence suggested that FDI was unable to render a significant contribution to the host country in terms of these criteria because of high capital intensity and import dependency. Feldstein (1994) argues that there is ample evidence that total employment in an economy with a well-functioning labour market will not be affected by the volume or character of FDI. For example, Graham and Krugman (1991) concluded that the net impact of FDI on US employment is approximately zero.

The desirability of FDI with respect to its effect on the demand for labour at home is a controversial issue. The critics of outward FDI argue that such investment destroys jobs at home because the output of foreign subsidiaries becomes a substitute for exports from the home country. Proponents of outward FDI argue that it creates jobs, because domestic firms export more when they have foreign subsidiaries. It is arguable that the technological bias and capital intensive nature of most investment characterizing oligopolistic industries (made possible by their R&D expenditure, and necessitated and supported by their large size) means that investment is unlikely to promote considerable labour usage.

Earlier work on home country employment effects encompasses the important aspect of whether production by foreign subsidiaries of a home country's firms is a substitute or a complement to domestic production by the parent firms or by other home country firms (for a survey, see Blomstrom and Kokko, 1994). Blomstrom *et al*. (1997) argue that the difficulty associated with these studies is the lack of convincing counter-factual situations. They therefore raise the following questions. What would have happened in the absence of foreign production? Would the parent firm have supplied, by exporting, the markets now served by subsidiaries? Would the markets now served by subsidiaries' production, or by some combination of subsidiary and home country production, have been lost to the parent firm?

Evidence from US studies suggest either a positive relationship or no relationship between US-owned production in a market and exports to that market by the parent company and by US companies in general (see, for example, Blomstrom *et al*., 1988). It has also been found that there is a negative relationship between US-owned

production and exports to the host country from other sources. A positive relationship was found across firms between production abroad and firm exports to the world, suggesting that such production had not materialized at the expense of exports to third countries (Lipsey and Weiss, 1981, 1984). Studies of Swedish firms have reported some mixed results, with a long period of findings of positive relationships (Swedenbourg, 1979) and some more recent reports of negative ones, particularly in third-country markets (Svensson, 1996). The main reason for finding a positive relationship is the role of FDI in the rivalry markets. The reason for the ambiguity of the results of most of these studies is that they do not take account of a firm's most important motivation for producing in a foreign country: the opportunity to increase its market share or even the size of the market itself, or to defend its existing market share. This kind of investigation would require data on the size of particular product markets in host countries, which is difficult to obtain.

Other aspects of home country effects that have been studied include competition between home and foreign markets for an MNC's capital resources, the extent to which expansion of offshore production reduces the demand for labour at home, and the relationship between foreign production and home country wage levels. A study of US firms indicated that home and foreign investments were not independent, and that an increase in plant and equipment investment in foreign operations caused a decrease at home because it raised the firm's cost of capital (Stevens and Lipsey, 1992). Brainard and Riker (1997) have concluded that foreign subsidiaries' employment by US MNCs substituted only modestly for US parent employment at the margin. They find much stronger substitution among workers in US foreign subsidiaries located in different low-wage host countries. Furthermore, US wage studies have suggested a positive relationship across firms between foreign activity and home country wage levels (Kravis and Lipsey, 1988). This relationship may reflect an allocation of low-risk activities to foreign operations. Finally, Blomstrom *et al*. (1997) obtained results on the effect of FDI on home country employment using data on US and Swedish manufacturing MNCs. The Swedish experience seems to differ from that of the USA. The results for Sweden show that, given the level of sales by the parent, MNCs with more sales abroad will also have higher employment in the parent company. In sharp contrast to US MNCs, production by Swedish MNCs in both developed and developing counties seems to have positive effects on parent employment.

A related issue is the effect of FDI on relative wages. Feenstra and Hanson (1995) examined the increase in the relative wages of skilled workers in Mexico during the 1980s. They linked rising wage inequality in Mexico to capital inflows, whose effect was to shift production in Mexico towards relatively skill-intensive goods, thereby increasing the relative demand for skilled labour. They also examined the impact of FDI on the share of skilled labour in total wages in Mexico during the period 1975–88. The results they obtained indicated that growth in FDI is correlated positively with the relative demand for skilled labour. Driffield and Taylor (2000) examined the effect of inward FDI on the British labour market, providing evidence indicating that FDI leads to an increase in wage inequality and the use of relatively more skilled labour by local firms.

While there is a general agreement on the proposition that relative wage changes are caused by an increase in relative demand for skilled labour, economists are not in agreement over the reasons for the shift in demand. Two explanations are normally given. One is that the advent of information technology has caused firms to switch towards production techniques that are biased in favour of skilled labour (Davis and Haltiwanger, 1991; Lawrence and Slaughter, 1993; Berman *et al*., 1994). The other explanation is that an increase in import competition from low-wage countries has shifted resources towards industries that use skilled labour relatively and intensively (Borjas and Ramey, 1993; Leamer, 1993). However, Feenstra and Hanson (1995) argue that the rise in wage inequality across dissimilar countries is consistent with a third explanation: capital flows from North to South, and a corresponding rise in outsourcing by Northern MNCs, have contributed to a worldwide increase in the relative demand for skilled labour. A flow of capital from North to South, which is identified as outsourcing by Northern firms, shifts an increasing portion of input production to the South. These activities are, from the North's perspective, ones that use relatively large amounts of unskilled labour, but from the South's perspective are ones for which the reverse is true. The result is an increase in the relative demand for skilled labour in both countries, which in turn causes the relative wage of skilled labour to rise in both regions. This is called the 'capital-accumulation-outsourcing hypothesis'. Feenstra and Hanson (1995) examined the effect of FDI on the relative demand for skilled labour in Mexico. They found the growth in FDI to be correlated positively with the relative demand for skilled labour. In the regions where FDI was most concentrated, growth in FDI could account for over 50 per cent of the

increase in the share of skilled labour in total wages that occurred during the late 1980s. This is consistent with the hypothesis that outsourcing by MNCs has contributed significantly to the increase in the relative demand for skilled labour in Mexico.

The final issue to be discussed in this section is the effect of FDI on factor demand elasticity. Hatzius (2000) argues that the liberalization of FDI has made labour costs more important to domestic investment and long-run labour demand. If higher labour costs induce firms to relocate production abroad, domestic employment will fall. In recent years, this has led some economists to argue that falling barriers to FDI have made wage moderation more important for preserving employment. Unless wages are kept under control, the argument goes, capital will migrate to countries with lower labour costs, thus leading to a rise in unemployment.

In an open economy, rising labour costs tend to reduce labour demand, for three reasons. First, the typical firm produces less total output in response to higher unit costs (the output or scale effect). Second, capital is substituted for labour (the substitution effect). Third, some firms may move abroad (the location effect). With high FDI barriers, a given small change in production costs will push only a small number of firms to relocate. As FDI barriers fall, the same production cost change will induce more firms to move. Hence, the elasticity of the capital stock and labour demand with respect to production costs rises as relocation costs fall.

Hatzius (2000) presents a simple equation relating FDI to unit labour costs and other control variables, as suggested by Cushman (1987), Culem (1988), Lucas (1993), Moore (1993), Pain (1993), Klein and Rosengren (1994), Bajo-Rubio and Sosvilla-Rivero (1994), Wang and Swain (1995), and Barrell and Pain (1996). Control variables include the deviation of output from trend in the source and host countries, which is used to control for short-term business cycle fluctuations, the relative number of days lost to strikes and lockouts to control for the industrial relations climate, and the relative real long-term interest rate to control for differences in the return on financial assets across countries. Dummy variables are also included to control for unobserved factors that drive the international propensity to invest abroad. Hatzius obtains evidence from British and German data that suggests the following conclusions. First, high unit labour costs encourage FDI outflows and discourage FDI inflows. Second, the effect of unit labour cost on domestic manufacturing investment was more negative in the high-FDI 1980s

than in the low-FDI 1970s, and this change was concentrated in high-FDI industries. The estimates suggest that the long-run labour demand elasticity may have risen substantially.

What can we conclude from the discussion of the employment effect so far? While FDI may have positive employment effects under certain conditions, it can certainly have adverse effects on employment and industrial relations in the host country. This is probably what prompted the OECD to issue its guidelines on how MNCs should deal with the employment and industrial relations issue. These guidelines boil down to urging MNCs, *inter alia*, to (i) respect the rights of their employees represented by trade unions; (ii) contribute to the abolition of child labour; (iii) avoid discrimination against employees; (iv) encourage the development of collective agreements; (v) provide a true picture of the performance of the firm; and (vi) take into account the livelihood of their employees when they consider changes in operations.

THE BALANCE OF PAYMENTS EFFECT

The balance of payments effect is more important for developing countries than for developed countries. This is because foreign exchange is regarded as a scarce resource affecting growth through the foreign exchange gap. Hence, any effect of FDI may mitigate or worsen the constraints imposed by the balance of payments on the attainment of macroeconomic objectives pertaining to growth and employment. In general, FDI is often blamed for its balance of payments effect: the investing country faces a sudden deficit when the FDI occurs, while the host country faces a small perpetual deficit as a result of profit repatriation. After all, a profitable FDI project with profits repatriated in foreign currency must necessarily result in greater balance of payments outflows than a similar project financed locally.

The balance of payments effect has certain features. First, there is a distinction between direct and indirect balance of payments effects. Hence, the effect of FDI must be examined in terms of (i) the absorption of the host country's factor input in the production process; (ii) the proportions of output sold in the host country and abroad; and (iii) the distribution of the value of output between the host country's factor inputs, the host government (in the form of tax revenue), and the retained share. The direct effect, which is reflected

immediately in the foreign exchange gap, results from the flows associated with the investment. Inflows include (i) exports and inflows of equity capital; and (ii) loans from abroad net of capital and loans repatriated. Outflows include: (i) value of capital goods imported; (ii) the value of raw materials and intermediate goods imported; (iii) royalties and technical fees paid abroad after tax; and (iv) net after-tax profits and interest accruing abroad. The direct effect does not tell the whole story, and it is deficient in two important respects. First, it does not show what would have happened had the foreign investment not occurred. Second, it does not give an idea of the effect of FDI on the balance of payments via domestic sales and the use of local resources. The latter is the indirect effect.

The second feature of the balance of payments effect is that it takes two forms. The initial, one-off, effect leads to an improvement in the capital account of the host country by the amount of the investment less the value of any imported machinery. The second form is the continuing effect, which is by far the most important.

The third and final feature is that the balance of payments effect is the result of (i) the export effect; (ii) the import substitution effect; (iii) the import effect; and (iv) the remittances effect. The first two effects lead to an improvement in the balance of payments, whereas the remaining two lead to its deterioration.

The empirical evidence on the balance of payments effect of FDI indicates that there is a difference between developed and developing countries, particularly with respect to investment in manufacturing industries. In assessing the impact of US FDI in Britain, Dunning (1961, 1969) estimated a positive effect of around 15 per cent of the total capital invested. However, Dunning only dealt with the direct effect of FDI, which results in observable flows in the balance of payments. He did not consider the indirect effect arising from changes in the income of residents, or changes in consumption patterns.

There is evidence indicating that the balance of payments of developing countries benefit from FDI in extraction, but not in manufacturing (for example, Vaitsos, 1976). Investment in manufacturing seems to have detrimental effects on the balance of payments of developing countries because of the high import content of the investment as well as the mechanism of transfer pricing of MNCs. The high import content of the output of MNCs results from the unavailability of local products and materials, the uncompetitiveness of local prices, and inferior quality.

THE EFFECT OF FDI ON TRADE FLOWS

Romer (1975) makes a very interesting observation about the relationship between the involvement of countries in FDI and their involvement in trade, based on the economic history of four industrial countries: the UK, Germany, Japan, and the USA. He puts forward the proposition that an industrial country passes through the following four stages, characterized by the country's shares of trade and FDI: (i) the share in world exports of manufactured goods rises; (ii) the trade share stabilizes, while the share in FDI starts rising; (iii) the trade share starts to fall; and (iv) the FDI share starts to fall. More recently, attempts have been made to integrate the theory of FDI with the theory of international trade (see, for example, Markusen, 2000).

Concern has been expressed over the effect of FDI on trade flows. Although it may appear that MNCs are trade-intensive firms, this is more a reflection of the activities in which they indulge rather than their own behaviour. Solomon and Ingham (1977) suggest that MNCs export fewer engineering products than do domestic firms, while Panic and Joyce (1980) suggested that MNCs' exports were stagnant over the 1970s. More recently, Goldberg and Klein (1997) have shown that FDI directed into developing countries affects their trade flows with industrial countries even after controlling for the effect of the exchange rate. There is also some evidence indicating that subsidiaries tend to import parts and capital equipment from the parent MNC, located in the home country. Hence, it seems that by indulging in FDI, MNCs affect the size and direction of trade flows. In a study of US FDI in developing countries, Rock (1973) found a significant correlation between the trade of these countries in the USA and FDI in these countries.

The most critical issue about the relationship between FDI and trade is whether they are complements or substitutes. In other words, to what extent do production and sales by subsidiaries in a foreign market replace or help to increase exports to the same market? One reason we should believe that FDI and trade are substitutes is that they are two alternative modes of entry, as we have seen. However, there are reasons to believe that FDI does not replace exports, but rather stimulates them. One reason for this is that FDI enables firms to establish a larger distribution base, thus enlarging the line of products sold in the foreign market over and above what could be achieved via exports. Moreover, production in the foreign country invariably requires the import of intermediate products from the

home country (and hence exports). This argument also applies to imports by the home country. If a foreign subsidiary can produce goods more cheaply abroad and export them to the home country, this obviously means that FDI leads to increasing imports by the home country.

There is now a consensus that whether trade and FDI are complements or substitutes depends on whether FDI is horizontal – as in Markusen (1984), or vertical – as in Helpman (1984). Whether FDI is horizontal or vertical depends on various country characteristics. For example, if countries have significantly different factor endowments, then vertical FDI dominates. On the other hand, horizontal FDI dominates if countries are similar in size and relative endowments, and if trade costs are moderate to high. Markusen *et al*. (1996) tried to predict the relationship between FDI and country characteristics. Naturally, it seems that the same criteria would determine whether FDI and trade are complements or substitutes.

But why would this relationship exist? In horizontal FDI, firms serve foreign markets by setting up plants there to provide identical goods (Markusen, 1984). Hence, exports from the source country to the host country will decline, implying that they are substitutes. In vertical FDI, MNCs separate different production stages geographically across countries, to take advantage of lower factor prices (Helpman, 1984). Specifically, the unskilled-labour-intensive stages of production are located in a low-wage country. In this case, there will be an increase in the exports of final products from the host country (the cheap labour country), while there is also an increase in the exports of intermediate products by the MNC (from the source country) to the host country where the subsidiary is located. Hence, FDI and trade are complements in this case. The question really boils down to the relationship between the sales of MNCs and the volume of trade.

The available empirical evidence is mixed. Most empirical studies based on cross-sectional industry and firm level data indicate a positive relationship. For example, Lipsey and Weiss (1981, 1984) and Blomstrom *et al*. (1988) found a predominantly positive relationship. Pain and Wakelin (1998) considered a time series relationship between manufacturing exports and FDI for eleven OECD countries, with mixed results. Blonigen (1999) argues that most of the empirical studies of the relationship between trade and FDI indicate complementarity (for example, Belderbos and Sleuwaegen, 1998). He attributes this finding to aggregation bias and shows that the substitution

effect is easy to identify in product level data. Finally, Amiti *et al*. (2000) explain why the evidence is mixed in terms of country characteristics, hence attributing the difference to whether FDI is horizontal or vertical.

THE EFFECT OF FDI ON PRODUCTIVITY

Productivity is likely to rise and unit cost likely to decline if: (i) FDI is export-promoting and the products of the subsidiary are destined for the large world markets; and (ii) the underlying conditions and policies allow the installation of plants designed to achieve full economies of scale. On the other hand, if FDI is import-substituting and the size of the market is too small to allow the installation of the optimum plant size, then productive efficiency may not be achieved. There are, however, some reasons for believing that productive inefficiency may not be important. First, the empirical evidence indicates that unit costs of operating a plant smaller than the optimum size are not significantly higher than those of the most efficient scale. Second, even if investment was mainly import-substituting, any scope for some exports leads to an increase in the size of the market and allows for the utilization of a higher capital intensity technology.

Petrochilos (1989) argues that productivity is likely to be affected by a host of other factors that are not necessarily exclusively applicable to FDI. The general factors include: (i) the full utilization of the firms' resources; (ii) the quality of existing manpower; (iii) the climate of industrial relations; and (iv) the existence or otherwise of restrictive practices. The effect of FDI on productivity is channelled through technology diffusion, an issue we shall deal with next.

FDI AND TECHNOLOGY

Technology diffusion plays a central role in the process of economic development (see, for example, Nelson and Phelps, 1966; Jovanovic and Rob, 1989; Segerstrom, 1991). The interaction between FDI and technology is considered to be of great and critical importance in the discussion of FDI. Indeed, the transfer of technology has perhaps become the predominant issue around which discussions of MNCs and their dealings with developing countries evolve. This is because technology is believed to be a vital source of economic growth, capital

accumulation, trade, and even changes in the organization of social relations and the relations of production. Particular problems in this respect take the form of how foreign technology is transferred to, and absorbed by, the host country and how it affects that country's economy. It is because of problems like these that the anticipated positive effects of technology on developing countries may not materialize. The importance of technology, particularly for developing countries, has prompted the OECD to issue guidelines urging MNCs to: (i) ensure that their activities are compatible with the technology plans of the host countries; (ii) adopt practices that allow the transfer and rapid diffusion of technology; (iii) address local market needs in an exercise pertaining to technology; (iv) license technology on reasonable terms and conditions; and (v) develop ties with local universities and research institutes. As we shall discover from the following discussion, it is doubtful whether MNCs' practices are consistent with these guidelines.

Technology is the product of R&D, aiming at the invention of new products or techniques of production, or both (Petrochilos, 1989, p. 36). Johnson (1970) considered the transfer of technology to be the crucial element of the FDI process. Any new technology confers an advantage on its owner, the original investor in R&D. In the case of new products, the advantage comes in the form of monopoly power. The owners of a particular type of technology have the options of selling the technology, licensing it, or exploiting it directly in production. The first option is unlikely to be considered by large firms engaged in R&D, but it may appeal to a sole inventor. On the other hand, licensing is considered as a suitable form of transfer to firms in the host country, but it may be limited because of the need of the owners to maintain control over business secrets, patents and trademark rights. Reasons for licensing may involve external factors, such as the prohibition of FDI by the host country, and internal factors such as the desire to receive a return on a sunk cost asset, technology or a product approaching obsolescence. If FDI cannot take place, for whatever reason, the transfer of technology can be accomplished through a technical assistance agreement.

Technology diffusion can take place through a variety of channels, including imports of high-technology products, adoption of foreign technology, and acquisition of human capital through international study. FDI by MNCs is considered to be a major channel for the access to advanced technologies by developing countries. In addition to its effects on technological progress, FDI contributes to growth by increasing capital accumulation in the host country. Knowledge

transferred from the MNC to its subsidiaries may leak out to the host country, giving rise to an externality known as the spillover effect from FDI. Various channels for spillover have been suggested: labour turnover from MNCs to local firms; technical assistance and support to suppliers and customers; and the demonstration effects on local firms in issues such as choice of technology, export behaviour, and managerial practices.

It appears that the benefits of foreign technology accruing to the investing firm and the source country are substantial (for example, increased monopoly power). However, the corresponding benefits of foreign technology accruing to the host country may not be so obvious, and may turn out to be negligible or even negative. This may be caused partly by the inability of the host country to absorb the foreign technology properly. The main reason seems to be the fact the technology is created to suit a particular environment of factor endowments but it is then used indiscriminately in different environments, in which case it may work against the interests of the host country. Technology may also have a detrimental effect on employment: new foreign technology can destroy jobs in existing industries, render obsolete human skills, and necessitate retraining of the redundant labour force, thus accentuating further the difficulties on the employment front.

Lall and Streeten (1977) cast doubt on the proposition that FDI plays an important role in technology diffusion. First, they consider the appropriateness of technology with respect to the products that are made with the technology transferred, and to the factor endowments of host countries. It is not only that the products are developed in high-income countries, but Lall and Streeten also argue that it is in the nature of MNCs that their products are excessively sophisticated in relation to the needs of developing countries (will Rolex allow the production of cheap versions of its watches to suit low-income consumers, for example?). Winters (1991, p. 229) agrees with this argument, stating that MNCs frequently pass on old technologies, which can be too capital-intensive for the local economy. Such a technology transfer, according to Winters, would create a 'dualistic structure' in the host country containing a small advanced industrial sector linked to the outside world, surrounded by a large, capital-starved sector. If the technology is capital-intensive in relation to factor endowment in the host country, then technology transfer would result in: (i) worsening employment; (ii) worsening income inequality; (iii) distorting influences on the technology used by other firms; and (iv) bias in production towards sophisticated and differentiated products. It is by

no means controversial that much of modern technology cannot be adapted to suit developing countries. Reuber *et al*. (1973) found, in a survey of the issue of adapting foreign technology, that most foreign firms generally introduce their production technologies in the host country intact. If adaptation takes place then it must be in response to (i) the need to scale down the volume of production in keeping with the size of local markets; (ii) the difficulty of achieving and maintaining acceptable standards of quality control; and (iii) the local customs and legal regulations of the host country.

Lall and Streeten (1977) further argue that, even if we disregard the appropriateness of the technology, the role played by MNCs and FDI in the transfer of technology may be limited, for the following reasons. First, there are several sources of technology besides MNCs (such as small consultants who are not interested in FDI). Second, the relative importance of MNCs in transferring technology depends on several factors, including the commercialization of the technology. Third, FDI may not be the only way of acquiring technology from an MNC. Fourth, and perhaps more importantly, the price set for the transfer of technology depends on the form of transfer and the bargaining skills of the parties (the MNC and the host country). There is ample evidence that MNCs have greater bargaining power than the developing host countries ('money talks'). Narula and Dunning (1999) argue that the balance in bargaining power has shifted in favour of MNCs. Furthermore, the conditions attached to the transfer of technology may restrict the freedom of the party acquiring the technology to buy and sell commodities related to the technology. Finally, the patent system as embodied in the Paris Convention and national laws imposes further constraints in the form of restrictive clauses, and the sale of intermediate and capital goods on which transfer pricing can be used.[10] Winters (1991) argues that MNCs are very skilled and powerful negotiators, which enables them to strike a very favourable bargain in a bilateral negotiation with the government of a poor country. This is particularly the case because the negotiating poor country has little information about the underlying technology. As a result, the MNC will be in a position to demand high royalties and impose severe licensing restrictions. In summary, while FDI may in theory be the fastest and most efficient way of gaining access to the latest technology, some costs arise because the monopolistic power of MNCs allows them to extract rents.

Work on economic growth has highlighted the contribution of FDI to the technological progress, and hence growth, of developing countries.

However, the evidence is rather mixed. Findlay (1978) postulates that FDI boosts the rate of technical progress in the host country through a contagion effect from the more advanced technology and management practices used by foreign firms. Wang (1990) incorporates this idea into a model more in line with the neoclassical framework, by assuming that the increase in 'knowledge' applied to production is determined as a function of FDI. There are a number of studies examining spillovers from FDI. Positive spillovers were found in Australia (Caves, 1974b), Canada (Globerman, 1979), and Mexico (Blomstrom and Persson, 1983). No positive spillovers were found in Morocco (Haddad and Harrison, 1993) or Venezuela (Aitken and Harrison, 1999).

The difference in the results suggests that the ability to benefit from foreign technology is not automatic as it is affected by various economic and technological factors. Findlay (1978) presents a dynamic model of technology transfer through FDI from developed to developing countries. He concludes that spillovers are greater, the wider the technology gap. However, it is also arguable that wide gaps may constitute an obstacle to spillovers. Wang and Blomstrom (1992) take competition into account, arguing that the greater the extent of competition, the greater will the transfer of technology also be, from the MNC to subsidiaries, and this is likely to leak out to local firms. A high technology gap combined with a low degree of competition tends to prevent spillovers. Sjoholm (1999) examines spillovers from FDI in the Indonesian manufacturing sector. The results show that high competition tends to increase the degree of spillovers from FDI, because the degree of competition affects the choice of technology transferred from the MNC to subsidiaries. It also seems that domestic competition rather than competition from imports affects spillovers from FDI. The fruits of technology transfer are therefore by no means certain, or they may be too sour to be enjoyed.

FDI AND TRAINING

Foreign investors, much as they dislike to spend on the training of locals, realize that such expenditure may be crucial for the success of their investment. Therefore, expenditure on training becomes part of the initial investment and another sunk cost. While it is true that foreign subsidiaries can rely on expatriate personnel, at least at the beginning of operations, they have a strong incentive to limit the number of such personnel working

in the host countries, and so they start to use more local people as soon as they can. This is partly because of cost considerations, since the remuneration of an expatriate tends to be higher than that of a local employee. Using local people may also result from pressure from the host government.

The effects of FDI on the training of local employees are difficult to quantify. For example, the extent to which subsidiaries use an appropriate combination of local and foreign personnel is difficult to know. The capital-intensive nature of most FDI implies that the number of local workers who are likely to be involved in training is not large. Reuber *et al*. (1973) reached the conclusion that, even allowing for the fact that training costs could not be identified properly, costs of training local people are not large enough to make a significant contribution to the improvement of the skills of these people.

Training may sometimes be considered under the general heading of 'organization and management', in the sense that the host country will benefit from the 'managerial superiority' of MNCs, and these benefits would show up as lower costs and prices. Lall and Streeten (1977) consider three kinds of managerial benefits: (i) managerial efficiency in operations arising from better training and higher standards; (ii) entrepreneurial ability in seeking out investment opportunities; and (iii) externalities arising from training received by employees (technical, executive, accounting, and so on). However, Lall and Streeten (1977) cast doubt on the proposition that the practices of MNCs are necessarily more efficient than those of smaller firms. They also argue that the practices of MNCs may be irrelevant to the host country, in which case the training will be useless and may even be harmful. Furthermore, they argue that the hierarchical structure of MNCs may entail costs to the host countries. These costs are: (i) dependence and subordination; (ii) transfer pricing; and (iii) the suppression of local entrepreneurship.

FDI AND INTER-INDUSTRY LINKAGES

FDI can influence the economy of the host country via inter-industry linkages. To the extent that foreign subsidiaries establish links with local suppliers for locally-produced materials and parts, FDI can help to provide local firms with increased opportunities that in turn affect their

employment and income positions. These are called backward linkages. Forward linkages can also be established for distribution purposes.

Assuming that they have the choice, there is no reason to assume that the subsidiaries have a preference for locally-produced materials and parts compared to importing them. If these materials are produced by the parent firm or other subsidiaries, there is less scope for inter-industry linkages, particularly if the locally-produced goods are of inferior quality. Baranson (1966) cites evidence showing that supplies from the car industry in Mexico and the steel industry in India have a greater cost and are of a lower quality than the components produced in the US.

Petrochilos (1989, p. 44) warns of the hazard of exaggerating the importance of inter-industry linkages for the developing economies. Most MNCs operating abroad in the manufacturing sector are sufficiently vertically integrated or have incentives to engage in inter-subsidiary transactions that limit the scope for developing strong and extensive ties with local suppliers. Risk considerations may indicate that extensive ties are imprudent, while the state of industrial relations in certain host countries may be particularly significant in this respect. Also, for the subsidiary to minimize risk there is the option of a take-over of the local supplier.

THE EFFECT OF FDI ON MARKET STRUCTURE

FDI is likely to affect the structure of the industries it is directed towards. It may be responsible for improving the competitive forces or for worsening the monopolistic or oligopolistic elements in the host economy. Caves (1971, 1974b) argues that the entry of a foreign subsidiary into local markets can force more active rivalry and an improvement in performance than would a domestic entry at the same scale. This is because FDI is thought of as a vehicle for disseminating the transfer of technology, including a higher level of technical efficiency. On whether or not this will materialize depends the actual practice of MNCs, and this is what has prompted the OECD to issue some relevant guidelines for MNCs, with the aim of encouraging behaviour that is conducive to boosting competition. According to these guidelines, MNCs should (i) refrain from entering into or carrying out anti-competitive agreements such as price fixing; (ii) conduct their activities in a manner that is consistent with local competition laws; and (iii) co-operate with the competition authorities.

On the grounds of allocative efficiency, it is arguable that FDI can provide a significant increase in competition in the host country. Kindleberger (1969) suggests that the main impact of FDI is widening the scope for competition. This is because it is typical that foreign subsidiaries, backed up by strong parents, can compete effectively with local oligopolists and break the latter's grip on the local market. By reducing monopolistic/oligopolistic distortions, FDI can improve the allocation of resources in the host country.

The preceding discussion presupposes the existence of large local firms dominating local markets that can only be challenged by equally powerful rivals. If small local firms cannot provide this rivalry, MNCs will act as a catalyst for more competitive behaviour. However, if this condition is not satisfied, the danger is that foreign subsidiaries may dominate local markets. Reuber *et al*. (1973) argue that FDI may 'preempt the development of indigenous firms and managers capable of establishing and maintaining a strong countervailing influence'. In this case, there would be a worsening of market concentration and the possibility that monopolistic or oligopolistic practices might arise, leading to worsening industrial performance. Moreover, Reuber *et al*. argue that the entry of foreign subsidiaries might raise the level of concentration in the host country because their presence might exert pressure for mergers among local firms. Lall and Streeten (1977) argue that MNCs may induce a very high degree of oligopolistic concentration, imposing diminished price competition. Newfarmer and Mueller (1975) present evidence from Mexico and Brazil supporting the proposition that the entry of MNCs significantly speeds up the process of oligopolization in developing host countries.

FDI AND THE ENVIRONMENT

It is arguable that, because MNCs have significant financial, political and negotiating power, they can get away with causing a lot of damage to the environment, particularly in developing countries that are trying to attract FDI. Indeed, one of the reasons why MNCs choose to locate production facilities in developing countries is that these countries have less stringent environmental damage requirements. Indeed, the governments of these countries may even inflict damage on the environment in an attempt to attract FDI. Recall our earlier reference in Chapter 2 to the 1995 advertisement in *Fortune*, in which the government of the Philippines declared its willingness to 'fell [sic]

mountains', 'raze jungles', 'fill swamps', 'move rivers' and 'relocate towns' to please foreign investors.

Some work has been done on the environmental effects of FDI, again without reaching a consensus view on whether FDI is good or bad for the environment. A recent publication by the OECD (1999) deals with the effect of FDI on the environment, and explores the role of host countries in developing and implementing coherent policies to ensure that proposed projects are environmentally sound. It also considers the strengths and weaknesses of voluntary corporate environmental management. One of the arguments put forward by the OECD is that FDI can be either 'a boon or a bane for the environment', depending on the specific circumstances. It seems, however, that it is more likely that FDI is a boon for the environment in a developed country and a bane for it in a developing country. A major motivation for the anti-globalization movement is the environmental damage believed to be inflicted by FDI and the operations of MNCs in developing countries.

There is obviously some concern about the environmental effects of FDI that has prompted the OECD to issue some guidelines as to how MNCs should tackle environmental issues. The OECD urges MNCs to 'take due account of the need to protect the environment, public health and safety' and to 'conduct their activities in a manner contributing to the wider goal of sustainable development'. Specifically, the OECD's guidelines on the environment encourages MNCs, *inter alia*, to: (i) provide information on the potential environmental impact of their activities; (ii) consult with the communities affected directly by the environmental policies; and (iii) maintain contingency plans for preventing, mitigating and controlling serious environmental damage.

It is interesting to note that, in its 8 February 1992 issue, *The Economist* reported a World Bank internal memo in which, by using three arguments, a bank official allegedly advocated the idea of exporting more pollution to developing countries. The first argument was that a given amount of health-impairing pollution should be done in a country with the lowest cost; that is, the country with the lowest wages. The second was that the costs of pollution are likely to be non-linear, as the initial increments of pollution probably have very low costs. Finally, the third argument was based on the idea that demand for a clean environment for health reasons is likely to have very high income elasticity. No wonder, then, that anti-globalization demonstrations are directed against the World Bank and the IMF as much as they are directed against MNCs.

MODELLING THE EFFECTS OF FDI

Most of the empirical work that has been done on the effects of FDI is based on the single equation approach using time series or cross-section aggregated or disaggregated data. Typically, the underlying model would consist of an equation in which the dependent variable is the variable hypothesized to be affected by FDI, while FDI, whatever the measure may be, appears as an explanatory variable. Other explanatory variables are used to control for the effect of other variables on the dependent variable. For example, Borensztein *et al*. (1995) investigated the effect of FDI on economic growth by specifying a relationship of the form:

$$g = f(I^F, H, Y_0, X) \tag{3.1}$$

where g is the growth rate, I^F is foreign direct investment, H is the stock of human capital, Y_0 is the initial level of output, and X is a vector of variables that are frequently used as determinants of growth, such as government expenditure, and variables representing foreign exchange and trade restrictions. The implication of Equation (3.1) is that the growth rate is determined by FDI and other factors, and in this sense causality runs from FDI to growth. Now, compare this with models of FDI determination, such as the model used by Yang *et al*. (2000) to study FDI in Australia. This (time series) model is written as:

$$\begin{aligned} I_t^F = {} & \alpha_0 + \alpha_1 \Delta i_t + \alpha_2 \Delta E_t + \alpha_3 \Delta Y_t + \beta_4 \Delta W_t \\ & + \beta_5 O_t + \beta_6 D_t + \beta_7 \pi_t \end{aligned} \tag{3.2}$$

where i is interest rate, E is the effective exchange rate, Y is output, W is the wage rate, O is openness, D is a measure of industrial disputes, and π is the inflation rate. The implication of Equation (3.2) is that output (which is a proxy for market size) determines FDI, and hence causality runs from output to FDI.

It is more plausible, however, to postulate that both output and FDI are endogenous variables that affect each other within a macroeconomic system. Hence, a simultaneous equation model may be more appropriate as a representation of both the determination and the effects of FDI.[11] Unfortunately, most of the empirical work on FDI is based on the single equation approach, but one notable exception is

Petrochilos (1989), who specified and estimated a simultaneous equation econometric model designed to show the effect of capital formation on the growth of output in general, and the influence of FDI in particular. In this model, FDI is an endogenous variable which is determined within the system while affecting the other variables. A simultaneous equation model makes a lot of sense, since FDI affects and is affected by the other variables, and because postulating bi-directional causality is highly plausible. In a sense, this model can be used to test hypotheses concerning the determinants and the effects of FDI at the same time. The model consists of ten behavioural equations and three identities which can be written as follows (the notation is explained in Table 3.1).

Table 3.1 Notation for the simultaneous equations model

Endogenous variables		*Exogenous variables*	
Variable	*Definition*	*Variable*	*Definition*
C	Private consumption	F^H	Long-term housing finance
I^D	Gross private domestic investment in manufacturing industry	D	Tariffs
I^F	Gross private foreign investment in manufacturing	I^G	Gross public investment
I^H	Gross private residential investment	N	Employment in manufacturing
I^O	Other gross private investment	R	Discount rate
I	Total gross investment	S	Political dummy
M^C	Imports of capital goods	T	Payments for foreign technology
M^R	Imports of raw materials and intermediate goods	Z	Residual (government consumption + exports + subsidies + change in stock − indirect taxes)
M^O	Other imports	K	Difference between gross domestic product and personal disposable income
Y^M	Manufacturing output		
Y^O	Non-manufacturing output		
Y^D	Personal disposable income		
Y	Gross domestic product		

$$C_t = F_1(Y_t^D, C_{t-1}) \tag{3.3}$$

$$I_t = F_2(Y_t^D, F_{t-1}^H) \tag{3.4}$$

$$I_t^D = F_3(Y_{t-1}, I_{t-1}, D_t) \tag{3.5}$$

$$I_t^F = F_4(Y_{t-1}, D_t, R_{t-1}, S_t) \tag{3.6}$$

$$I_t^O = F_5(Y_{t-1}^D, I_{t-1}^O) \tag{3.7}$$

$$M_t^C = F_6(I_{t-1}^D, I_{t-1}^F, T_t) \tag{3.8}$$

$$M_t^R = F_7(I_{t-1}^D, I_{t-1}^F, Y_{t-1}^D) \tag{3.9}$$

$$M_t^O = F_8(Y_t^D, C_{t-1}) \tag{3.10}$$

$$Y_t^M = F_9(I_{t-1}^D, I_{t-1}^F, N_t, T_t) \tag{3.11}$$

$$Y_t^O = F_{10}(Y_{t-1}, I_{t-1}^O) \tag{3.12}$$

$$I_t = I_t^H + I_t^D + I_t^F + I_t^O + I_t^G \tag{3.13}$$

$$Y_t = Y_t^M + Y_t^O = I_t + C_t + Z_t - M_t^C - M_t^R - M_t^O \tag{3.14}$$

$$Y_t^D = Y_t^M + Y_t^O - K_t \tag{3.15}$$

The model was estimated using linear and log-linear specifications and Greek data by employing more than one estimation method. The main findings are as follows:

1. FDI in manufacturing is explained by the factors typically found in the literature, such as the size of the market (proxied by GDP), tariffs, political stability, and the discount rate.
2. Manufacturing output is determined primarily by domestic investment in manufacturing, imports of foreign technology and employment in large-scale industry and, to a lesser extent, by investment in manufacturing by foreign subsidiaries.
3. Capital formation played an important role in the development process of the Greek economy during the period 1953–78.
4. The direct employment effect arising from the operations of foreign subsidiaries is not significant.
5. The contribution of FDI to the balance of payments is difficult to evaluate because of data problems.

A FINAL REMARK

In this chapter we discussed the effects of FDI, a highly controversial and contentious issue. There is no doubt that FDI affects both home and host countries. In theory, the effects on the host country can be highly positive, but the benefits are not realized automatically. There are certain conditions that have to be satisfied for a positive effect to materialize, and these conditions are more likely to be satisfied by a developed rather than by a developing host country. The empirical evidence (more of which is presented in the appendix to this chapter) is so mixed that it cannot resolve the underlying issues.

So far we have dealt with the characteristics, determinants and effects of FDI. In the following six chapters we shall deal in more detail with some aspects of FDI and the behaviour of MNCs that we have encountered in the book so far. There is no doubt that MNCs are not charities and that they are profit-seeking firms. Therefore, FDI is triggered, first and foremost, by the profitability (or anticipated profitability) of the underlying projects, at least in the long run. In Chapter 4 we shall study international capital budgeting, which tells us how MNCs determine the financial feasibility of FDI projects.

Appendix

Table A3.1 Summary of selected recent studies of the effects of FDI

Study	*Issue under investigation*	*Findings*
Zukowska-Gagelmann (2000)	Examining the effect of FDI on productivity growth	FDI has a negative impact on the performance of the most productive local firms
Driffield and Taylor (2000)	The labour market impact of inward FDI in the UK	FDI leads to an increase in wage inequality and the use of skilled labour in domestic firms
Kearns and Ruane (2001)	Relationship between FDI and growth in Ireland	FDI has been beneficial to Ireland. R&D-active firms provide greater benefits
Fan and Dickie (2000)	Contribution of FDI to growth and stability in Asian countries	FDI accounts for 4–20 per cent of GDP growth
Xu and Wang (2000)	International trade and FDI as channels for technology diffusion	No evidence that FDI is a significant channel for technology diffusion

Nachum (1999)	Impact of FDI on international competitiveness	FDI weakens the link between location advantages and ownership advantages
Asafu-Adjaye (2000)	Effect of FDI on Indonesian economic growth	FDI has a significant positive effect on growth
Jarolim (2000)	Role of FDI in the economic transition of the Czech Republic	FDI's spillover effect is statistically insignificant
Henneberger and Ziegler (2000)	Effect of Swiss FDI on employment	Negative correlation between variations in levels of domestic and foreign employment
Leahy and Montagna (2000a)	The welfare implications of using union legislation to attract FDI	The host government may ban unions in the short run to extract higher rents in the future
Barrel and Holland (2000)	Effects of FDI on manufacturing sector in central Europe	FDI has led to increasing labour productivity in most manufacturing sectors
Figlio and Blonigen (2000)	Effects of FDI on local communities in the USA	FDI raises local real wages more than domestic investment but lowers per capita local government expenditure
Berthelemy and Demurger (2000)	Relationship between FDI and growth in China	FDI plays a fundamental role in provincial growth
Zhang (1999a)	Relationship between FDI and economic growth in Asian countries	FDI enhances growth in the long run
Chen and Ku (2000)	Effect of FDI on firm growth	FDI is beneficial to the survival of firms
Braunerhjelm and Oxelheim (2000)	Substitutability between FDI and domestic investment	Substitutability exists for R&D intensive production
Djankov and Hoekman (2000)	Relationship between FDI and productivity in Czech enterprises	FDI has a positive impact on total factor productivity of recipient firms
Hsu and Chen (2000)	Effect of FDI on labour productivity in Taiwan	FDI enhances productivity of small and medium-sized firms. It has negative spillover on large firms
Zhang (1999b)	Effect of FDI on economic growth in China	Long-run link and two-way causality between FDI and growth
Walkenhorst (2000)	Spillovers from FDI to related industries in transition economies	FDI brings not only capital but also managerial and technological skills

Table A3.1 (Continued)

Study	*Issue under investigation*	*Findings*
Bosworth and Collins (1999)	Implications of financial flows for saving and investment in host country	Little correlation among FDI, portfolio investment and loans. FDI has close to one-to-one effect on investment
Glass and Saggi (1999a)	Consequences of FDI in a general equilibrium setting	FDI raises wages and lowers profits in the host country, and vice versa
Yabuuchi (1999)	Effects of FDI on welfare and unemployment	An increase in FDI leads to an increase in welfare and a decrease in unemployment if capital is also used in the domestic manufacturing sector
Fung *et al.* (1999)	Effects of FDI on national welfare	FDI can affect national welfare positively or negatively
Saggi (1999)	Implications of licensing and FDI for technology transfer	Relative to licensing, FDI limits technology spillovers to local firms, but dissipates more rents in the product market
Bonelli (1999)	Links between FDI and industrial competitiveness in Brazil	FDI has contributed to increased productivity and competitiveness in Brazil
Roling (1999)	German job export through FDI	Empirical basis for German job export is weak
Driffield (1999)	Employment consequences of inward FDI in the UK	FDI generates employment substitution away from local firms
Okamoto (1999)	Effect of FDI on production efficiency	FDI has a positive effect through the enhancement of competitive pressure and technology transfer
Chuang and Lin (1999)	Effect of FDI on productivity	FDI has a positive spillover effect on productivity
Elahee and Pagan (1999)	The role of FDI in Asia and Latin America	FDI plays an important role in fostering economic growth
De Andrade-Castro and Teixeira (1999)	FDI, technology transfer and growth	FDI may have a positive effect on long-run growth, eventually helping the recipient country to catch up with the investing country

Aitken and Harrison (1999)	Effect of FDI on domestic firms in Venezuela	FDI affects the productivity of domestic firms negatively. Net impact of FDI is small
De Mello (1999)	Direct investment-led growth	The extent to which FDI is growth-enhancing depends on the degree of complementarity and substitution between FDI and domestic investment
Glass and Saggi (1999b)	FDI and technology	The role FDI plays in technology transfer depends on whether substitute channels are available for transfer to the host country
Heinrich and Konan (2000)	Impact of PTAs on horizontal FDI	PTA welfare increases regardless of changes in FDI
Stone and Jeon (2000)	Relationship between FDI and trade in Asia-Pacific economies	Significant and positive relationship between FDI and trade
Mucchielli *et al*. (2000)	Relationship between intra- or inter-firm trade and FDI	Complementarity for global trade is explained by complementarity for intra-firm trade and substitutability for inter-firm trade
Castilho and Zignago (2000)	Relationship between FDI, trade and regional integration	Positive link between FDI and trade flows mitigated by the impact of integration on FDI
Ellingsen and Warneryd (1999)	FDI and protectionism	An import-competing industry may not want maximum protection because it may encourage FDI, which could be less desirable
Wilamoski and Tinkler (1999)	The effect of FDI on exports and imports	FDI leads to increased exports and imports
Gopinath *et al*. (1999)	FDI and trade	Small substitution effect between foreign sales and exports
Chen (2000)	Relationship between FDI and intra-industry trade	Positive and strong link between FDI and intra-industry trade

5 Country Risk and Political Risk

Country risk represents a potentially adverse impact of a country's environment on the cash flows generated by an FDI project. Country risk analysis is important for a number of reasons. First, the MNC can use it as a screening device to avoid investing in countries with excessive risk. A second reason is that it can be used to monitor countries where the MNC is currently engaged in international business. In this case, a decision to divest (which may involve a change of the location of the production facilities) may be taken if it is felt that country risk has become excessive. A third reason why the study of country risk is important for MNCs is the need to assess particular forms of risk for a proposed project considered for a foreign country. These forms of risk may be general, such as economic risk and political risk, or they may be more specific, such as the risk of a take-over by the host government. Of course, all these problems can be avoided by keeping away from international business, but this strategy would preclude potentially profitable opportunities. Furthermore, it is the antithesis of being a multinational firm.

Apart from the general risks associated with any kind of investment (domestic or foreign), such as market risk and credit risk, investing outside a country's national frontiers entails more and different kinds of risk. These additional risks, which include currency risk and country risk, arise as a result of the nature of the foreign market. Investing abroad generally implies additional costs apart from the normal costs resulting from lack of knowledge about the foreign market. For example, currency risk arises because foreign investment necessarily means that the underlying assets are denominated in a foreign currency, whereas country risk arises because of economic and political factors that are specific to the country where the investment takes place. The result is a greater exposure to a more diversified set of risks.

COUNTRY RISK, SOVEREIGN RISK AND POLITICAL RISK

Considerable conceptual confusion surrounds the idea of country risk. The concepts of 'country risk', 'sovereign risk' and 'political risk' are

often regarded, wrongly, as being interchangeable, when in fact they mean different things. Country risk is wider than either sovereign risk or political risk. We shall now describe the three concepts, but political risk will receive a more extensive treatment in subsequent sections because of its special relevance to FDI.

Country risk may be defined as exposure to a loss in cross-country transactions, caused by events in a particular country that are, at least to some extent, under the control of the government, but definitely not under the control of a private enterprise or individual. Country risk is a broader concept than sovereign risk. Any cross-border transaction is subject to country risk, whether this transaction is conducted with the government, with a private enterprise, or with an individual. Sovereign risk is limited to transactions (mainly lending) to the government of a sovereign country. Since this book is about FDI, not about bank lending, we shall not consider sovereign risk any further.

Only events that are, at least to some extent, under the control of the government can lead to the materialization of country risk. For example, a possible bankruptcy of a subsidiary is country risk if the bankruptcy is caused by a mismanagement of the economy by the government; but is commercial risk if it is the result of mismanagement by the firm. Consider also the case of national calamities. If they are unforeseeable, they cannot be considered to be country risk. But if past experience shows that they have the tendency to occur periodically, then the government can make preparations for such a contingency in order to minimize their harmful effects.

Country risk arises from political risk factors, and economic or financial risk factors. Hence, political risk is a subset of country risk arising from political risk factors. These include war, occupation by a foreign power, riots, disorder, attitude of consumers in the host country, attitude of the host government, changes in the rules and regulations governing FDI, blockage of fund transfers, currency inconvertibility, and bureaucracy. All these factors can obviously have adverse effects on the cash flows arising from FDI projects. In general, political risk refers to potential losses to a firm resulting from adverse developments in the host country. Events whereby political risk materializes range from the outright expropriation of assets to unexpected changes in the tax laws that hurt the profitability of FDI projects.

Economic factors pertain to the current and potential state of the economy. An MNC that exports to a country or establishes a subsidiary in a country should be concerned about the demand for its

products in that country. This demand, naturally, is influenced strongly by the state of the country's economy. A recession in the country could reduce demand significantly because lower income leads to lower demand for all goods and services, including those that are imported. However, one has to be careful about the contribution of recession to economic and country risk. If the recession is worldwide and the government applies appropriate countercyclical policies, it is not country risk. But if the recession occurs in one country in isolation, or if it is aggravated by improper government policies, then it is country risk.

Because the state of a country's economy depends on several factors, an MNC should consider all these factors. Some of the obvious ones are interest rates, exchange rates, and inflation. High interest rates normally depress the economy and reduce the demand for the MNC's products, while low interest rates have the opposite effects. Exchange rates determine the country's demand for exports as well as the domestic currency value of the MNC's costs and revenues. This is why MNCs tend to establish production facilities in countries where currencies are undervalued, and sell their products in countries where currencies are overvalued. The problem is that the state of undervaluation or overvaluation may change over time. Inflation affects the purchasing power of consumers and can be an adverse factor for economic growth. Moreover, these three factors are intricately related. Interest rates and exchange rates are related, not only because higher interest rates attract capital flows and lead to currency appreciation, but also because uncovered interest parity (UIP) implies a direct relationship between the expected change in the exchange rate and the interest rate differential. Interest rates and inflation rates tend to move together, because they are related by the Fisher equation (stipulating a proportional relationship between the two variables), and because central banks tend to raise interest rates when inflation rises or is expected to rise. Finally, exchange rates and inflation rates are related by the purchasing power parity relationship, so that the higher the domestic inflation rate, the weaker will be the domestic currency.[1]

Consider, for example, an MNC operating in a country with a high inflation rate and a depreciating currency. Currency weakness threatens the long-term hard currency value (that is, the value in hard currency terms) of local profit and cash flows. High inflation, on the other hand, poses difficult operating problems, such as the problem of pricing products, and the inflation-induced financial accounting

problems. The objective behind measuring country risk arising from these factors is to decide whether an investment in a country with these characteristics is worthy of being undertaken, and if it is, what kinds of measures should be taken to reduce their adverse effects. For example, the MNC may choose to export a substantial part of production to combat currency weakness. Another measure is to maximize the local content of the production process to reduce the problems of foreign currency costs. To reduce the effect of inflation, the MNC may choose to minimize cash and other asset holdings subject to the risks of inflation and hence declining real values.

Thus, country risk is broader than calculating the probability that a subsidiary will be taken over by the host government. The latter is one kind of political risk, which is a subset of country risk. The assessment of country risk includes the assessment of political risk and economic risk that can influence cash flows. Furthermore, one has to bear in mind that political risk events are often triggered by economic factors. Political risk is thus not always entirely independent of economic risk. For example, a persistent trade deficit may induce the host country's government to delay or stop interest payments to foreign lenders, erect trade barriers, or suspend the convertibility of the currency, causing major difficulties for MNCs. Severe inequality in income distribution and deteriorating living standards can cause major political disturbances. For example, in 1977, Egypt witnessed popular revolt against a massive rise in the price of bread resulting from the implementation of some IMF-prescribed economic policies that included the abolition of subsidies. In this chapter we deal more extensively with political risk rather than economic risk, since it is more relevant to FDI.

COUNTRY RISK ASSESSMENT

Although there is no consensus on how country risk can best be assessed, some guidelines have been developed for this purpose. The first step is to recognize the difference between macro-assessment and micro-assessment of country risk. Macro-assessment refers to the overall risk assessment of a country without consideration of the specific characteristics of the MNC's business. Micro-assessment, on the other hand, refers to the risk assessment of a country as related to the particular characteristics of the business in which the MNC indulges.

Macro-assessment involves a consideration of all the variables that affect country risk except for those that are unique to a particular industry. This type of assessment is convenient, because it remains the same for a given country, regardless of the firm or industry under consideration. This is the assessment underlying the country risk ratings found in financial magazines such as *Euromoney* and *Institutional Investor*. Thus, macro-assessment of country risk is not ideal for an individual MNC because it excludes relevant information that could lead to an improvement in the accuracy of the assessment. However, macro-assessment could serve as a foundation that can be modified to reflect the particular business in which the MNC is involved. In this case, the macro-assessment may be carried out by an external party, such as *Euromoney* magazine, whereas the micro-assessment is carried out by the MNC.

Macro-assessment involves a consideration of both political and economic indicators of the country under examination.[2] Political factors include, *inter alia*, the relationship between the host government and the MNC's home country's government, the historical stability of the host government, the probability of war, the probability of changing the rules of the game, and so on. The economic factors should include the main macroeconomic indicators, both current and projected, such as economic growth, inflation, the fiscal balance (budget deficit or surplus), interest rates, unemployment, the extent to which the country relies on export income, the balance of payments and its components, and so on.

There is normally some subjectivity in identifying each of the relevant political and economic factors for the macro-assessment of country risk. There is also some subjectivity in determining the weights assigned to each factor. Furthermore, there are some differences in predicting these factors. However, it seems that as far as FDI is concerned, political risk factors are more important than economic ones. A study by Petry and Sprow (1993) has identified the factors with the greatest potential impact on the profitability of large MNCs. In order of importance, these factors are: (i) restrictive practices; (ii) tariffs or regulations; (iii) unstable currencies; (iv) foreign government subsidies; (v) shaky government; and (vi) national debt.[3]

Micro-assessment of country risk involves the evaluation of micro-political risk and microeconomic risk. Micro-political risk can be best illustrated with the following example. Suppose that a country has received a very good score for macroeconomic risk. The government of that country is, however, sensitive to foreign ownership of mining

operations of uranium, but not to other operations, mining or otherwise. If this government is considering some legislation curtailing foreign ownership of uranium mining operations, then there is high (micro) country risk for an MNC considering starting uranium mining operations there. However, other MNCs will not be subject to this kind of risk. Microeconomic risk results from the sensitivity of the MNC's earnings to changes in the economic environment. Consider, for example, two MNCs operating in the same country: one of them produces electricity, while the other produces luxury clothing. Since the demand for electricity is less cyclical than the demand for luxury clothing, the first MNC's earnings will be less sensitive to economic growth and the business cycle than the earnings of the second MNC. Hence, a country with a good evaluation in the macro-assessment may end up with a low overall evaluation when the micro-assessment is taken into account, and vice versa.

Once the macro and micro factors have been identified, a number of techniques can be used to evaluate these factors. These techniques, which can be used in conjunction with each other, include the following:

The Checklist Approach

The checklist approach involves judgement on all the political and economic factors that are believed to contribute to country risk. Some of these factors (such as the current growth rate or inflation rate) are readily available. Others (such as the future growth rate and inflation rate, and the probability of a civil war) have to be assessed on a judgemental basis. These factors are then assigned weights and used to calculate a score.

The Delphi Technique

The Delphi technique involves the collection of independent opinions on country risk with no group discussion by the assessors who provide the opinions. The advantage of this technique is that the individual assessors (who may be the MNC's employees, external consultants, or both) will not be subject to group pressure when they form their opinions. The project leader has the task of collecting the individual views of the participants to come up with a 'consensus' view. The degree of disagreement can also be assessed by measuring the dispersion of opinions.

Quantitative Analysis

Quantitative analysis involves the use of statistical techniques to analyse data on the factors contributing to country risk. For example, discriminant analysis can be used to distinguish between a country with tolerable risk and another with intolerable risk. Regression analysis can also be used for this purpose, since it can measure the sensitivity of one variable against another. For example, regression analysis can give us an idea about the quantitative effect of changes in the growth rate, inflation rate and interest rate on an MNC's earnings. The problem with quantitative analysis is that the conclusions derived from it are based on historical data. Naturally, there is no guarantee that a relationship that has held so far will hold into the future. But it is the future we are interested in, since the outcome of a decision taken now is contingent upon what will happen in the future. In fact, there is no guarantee that the relationship will remain stable during the historical period over which it is estimated.[4] Another problem with quantitative analysis is that some important risk elements are not quantifiable. For more details on the general problem of trying to generate forecasts from a regression model, see Moosa (2000a).

The Old Hands Method

The old hands method amounts to collecting information from diplomats, journalists, economists, financial analysts and other professionals who have some expert knowledge of the country in question. These professionals typically are external to the firm. Two drawbacks of this method are its unsystematic character, and the fact that it is based on the judgement of outsiders. However, Rummel and Heenan (1978) point out that the old hands method is capable of providing firms with an improved understanding of developments in the country under discussion.

Inspection Visits

Inspection visits involve travelling to a country and meeting government officials, business people and consumers. Meeting these parties helps to clarify opinion about the country. The same task can be accomplished via a representative office in the country. Major banks and other MNCs establish representative offices in other countries for the purpose of following up political and economic developments. All the information that is gathered is then analysed and evaluated.

A shortcoming of this method is what Rummel and Heenan (1978) call an 'overdose of selective information'.

QUANTIFYING COUNTRY RISK

Once the political and economic risk factors have been evaluated, two things can be done with them: (i) the construction of a foreign investment risk diagram; and (ii) quantifying overall country risk. We start with a description of the foreign investment risk diagram.

A foreign investment risk diagram is shown in Figure 5.1, where economic risk is measured on the horizontal axis and political risk is measured on the vertical axis. The scale for risk ranges from zero (no risk) to 100 (maximum risk).[5] The firm's risk preferences are represented in the diagram. The firm can accept economic risk up to 70, is not sure for levels of economic risk between 70 and 80, and certainly will not accept a level of economic risk above 80. As far as political risk is concerned, the firm accepts a level up to 60, is not sure about levels between 60 and 70, and definitely will not accept a level above 70. The points marked on the horizontal and vertical axes are connected by

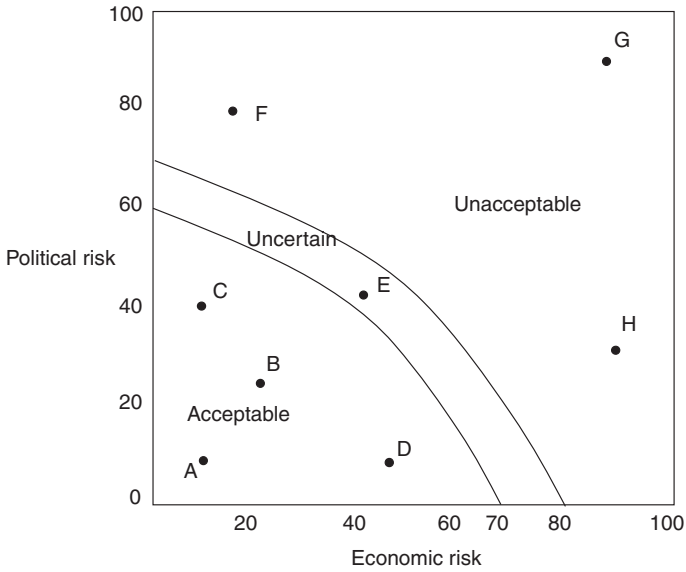

Figure 5.1 Foreign investment risk diagram

concave curves that reflect the trade-off between economic risk and political risk. The diagram is divided into three zones. Countries A, B, C and D fall into the acceptable zone. All these countries are acceptable, but given the choice between countries A and B, the firm will choose country A because it has both a lower political risk and a lower economic risk. The choice between countries C and D is not quite as clear, as it depends on the firm's preferences. Country C has more political risk, whereas country D has more economic risk. Countries F, G and H fall into the unacceptable zone because they have unacceptable political risk (F), unacceptable economic risk (H), or both (G). Finally, there is the uncertain zone where country E falls. This country has a level of economic risk between 70 and 80, and a level of political risk in the range 60–70. This case may be re-examined to find out if country E can be reclassified to fall into the acceptable zone.

The foreign investment risk diagram reflects the importance of political risk relative to economic risk. The shallower the curves representing the trade-off between political and economic risk, the lower the maximum amount of political risk and the greater the maximum amount of economic risk the firm is willing to bear. Figure 5.2 shows that a reduction in the level of political risk from P_1 to P_2 is accepted for a large increase in economic risk (from E_1 to E_2).

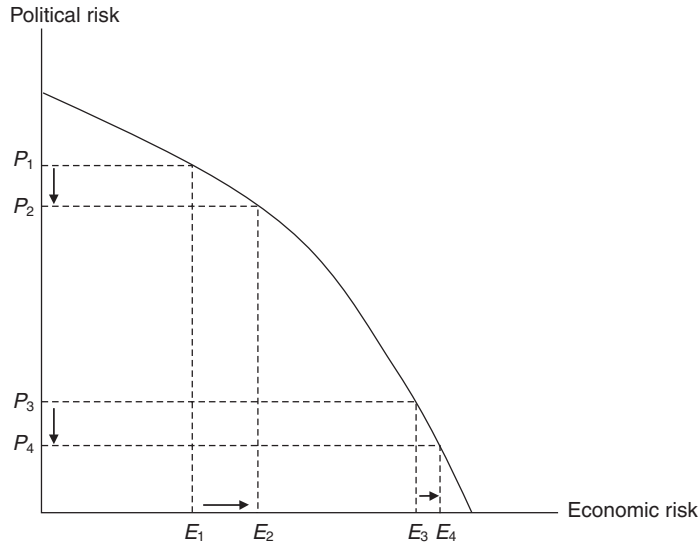

Figure 5.2 The trade-off between political risk and economic risk

However, if we start from a lower level of political risk, say P_3, then the same reduction in political risk as before (from P_3 to P_4, which is equal to the reduction from P_1 to P_2) will be acceptable at the cost of a smaller increase in economic risk (from E_3 to E_4, which is smaller than the reduction from E_1 to E_2). Thus, the foreign investment risk diagram is firm-specific, and this should be the case because the relative importance of the two kinds of risk varies from one firm to another. If an MNC wants to indulge in FDI to attract demand in a foreign country then it must be highly concerned about economic risk. If, on the other hand, another firm has the objective of locating an FDI in the country under consideration to exploit low production costs and export the products, then political risk would be more important. The first firm has more tolerance for political risk than for economic risk, whereas the second firm has more tolerance for economic risk. These two firms have foreign investment risk diagrams that look like diagrams A and B in Figure 5.3. Some firms have high tolerance for both kinds of risk, while others have low tolerance for both. These two cases are represented by diagrams C and D in Figure 5.3. There are also firms that are willing to accept the maximum amount of political risk if they are compensated by low economic risk (diagram E), and those that are willing to accept the maximum amount of economic risk if they are compensated by low political risk (diagram F). Finally, there are firms that are willing to accept the maximum amount of either kind of risk if they are compensated by a lower amount of the other kind of risk (diagram G).

While a foreign investment risk diagram may be useful for an MNC, it does not quantify the overall risk rating for an individual country, which requires the combination of political and economic risk factors. Suppose that there are m political risk factors: $P_1, P_2, \ldots, P_m$, and n economic risk factors: $E_1, E_2, \ldots, E_n$. Let us assign the weights v_1, $v_2, \ldots, v_m$ to the political risk factors, and $w_1, w_2, \ldots, w_n$ to the economic risk factors. The political risk score, P, and the economic risk score, E, are given, respectively, by:

$$P = \sum_{i=1}^{m} v_i P_i \tag{5.1}$$

and:

$$E = \sum_{i=1}^{n} w_i E_i \tag{5.2}$$

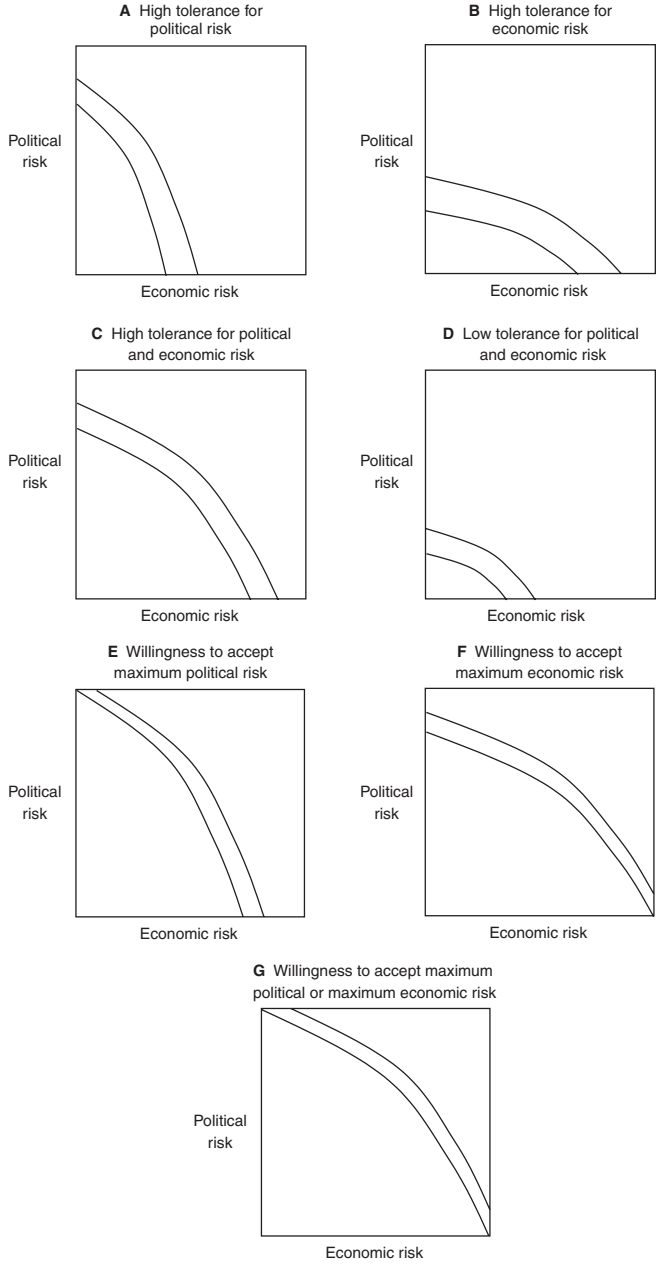

Figure 5.3 Risk tolerance as represented by the foreign investment risk diagram

If political risk and economic risk are assigned the weights λ_P and λ_E, respectively, then the overall country risk score, S, is given by:

$$S = \lambda_P P + \lambda_E E = \lambda_P \sum_{i=1}^{m} v_i P_i + \lambda_E \sum_{i=1}^{m} w_i E_i \tag{5.3}$$

It should be emphasized at this stage that country risk assessors have their individual procedures for quantifying country risk, which may differ from the one described in this section. In fact, it is often the case that the system is designed in such a way that a high score implies low risk, and vice versa. Here, the system is designed so that a high score means higher risk, which is easier to understand. In the following section, we describe a system that is used in practice, whereby the lowest risk country has the highest score, and vice versa.

AN EXAMPLE OF MACRO-ASSESSMENT: THE *EUROMONEY* METHODOLOGY

Periodically, *Euromoney* magazine publishes tables of country risk ratings. These ratings are based on country risk scores ranging from 100 (risk free) to zero (ultimate risk). The overall score is calculated as the sum of the individual scores corresponding to each of nine indicators, that is:

$$S_j = \sum_{i=1}^{9} S_{ij} \tag{5.4}$$

where S_j is the country risk score of country j and S_{ij} is the jth country's score corresponding to factor or indicator i, where $i = 1, \ldots, 9$. The individual scores are calculated by assigning a weight, w_i (expressed as a percentage, not as a decimal), to indicator i, such that $0 \leqslant S_i \leqslant w_i$. The exact formula that is used to calculate S_{ij} is:

$$S_{ij} = w_i - \left[\frac{w_i}{\ell - h}\right](x_{ij} - h) \tag{5.5}$$

where ℓ is the lowest value in the range, h is the highest value and x_{ij} is the value of the ith factor for country j. Table 5.1 reports the nine indicators and the maximum scores (which are equal to the weights, w_i) on each of these indicators. Ramcharran (1999) uses the *Euro-*

Table 5.1 Elements of the *Euromoney* country risk scores

Indicator	*Max* (w_i)	*Description*
Political Risk	25	Defined as the risk of non-payment or non-servicing of payments for goods, services, loans, trade-related finance and dividends, and the non-repatriation of capital.
Economic performance	25	Based on income per capita and economic forecasts
Debt indicators	10	Calculated from three ratios: debt to income; debt service to exports; and current account balance to GNP
Debt in default or rescheduled	10	Based on the ratio of rescheduled debt to debt stock
Credit rating	10	Based on the ratings of Moody's, Standard & Poor's, and Fitch IBCA
Access to bank finance	5	Calculated from disbursements of private long-term unguaranteed loans as a percentage of GNP
Access to Short-term finance	5	Based on the views of the OECD and the US Export Import Bank
Access to capital markets	5	Based on surveys of debt syndicates
Discounts on forfeiting	5	Reflects the average maximum tenor for forfeiting and the average spread over riskless countries

Source: *Euromoney*, September 1999, p. 254.

money risk measures to study the relationship between FDI and country risk over a period of three years.

Obviously, this system is different from the system described in the previous section, not only because a high score implies low risk, but also because other factors are taken into account as well as political risk and economic risk. Indeed, political risk and economic risk combined command half the total weight, whereas measures of indebtedness comprise the other 50 per cent of the weight. These measures of indebtedness are important for banks whose transactions are mainly cross-border lending, but not so important for FDI. Indeed, even political risk is defined (for the purpose of this measure) as the risk

Table 5.2 Country risk scores

Country	S_1	S_2	P	E	$S(P)$	$S(E)$
Switzerland	24.74	23.05	1.04	7.80	3.07	5.77
Japan	23.26	19.31	6.96	22.76	11.70	18.02
Greece	18.61	13.61	25.56	45.56	31.56	39.56
Egypt	13.08	8.04	47.68	67.84	53.73	61.79
Brazil	10.37	8.35	58.52	66.60	60.94	64.18
Vanuatu	10.09	9.85	59.64	60.60	59.93	60.31
Senegal	6.54	6.70	73.84	73.20	73.65	73.39
Belarus	3.06	3.81	87.76	84.76	86.86	85.66
Somalia	0.59	4.52	97.64	81.92	92.92	86.64
Cuba	2.02	3.97	91.92	84.12	89.58	86.46

Source: *Euromoney*, September 1999, pp. 251–3.

of non-payment and non-servicing of payment, which is rather different from what we have seen so far.

But let us for the time being assume that political risk is political risk no matter how it is defined. We can then convert the scores obtained from the *Euromoney* system to the scores that can be obtained from the system suggested in the previous section, which is based on economic risk and political risk only. Table 5.2 illustrates an example for ten countries with varying degrees of country risk.

Let us now figure out what Table 5.2 tells us. The first two columns report the political and economic country scores, S_1 and S_2, according to the *Euromoney* system as at September 1999.[6] These scores are out of 25, with a high number indicating low risk. Thus, Switzerland has both the lowest political risk and the lowest economic risk, Somalia has the highest political risk, and Belarus has the highest economic risk. P and E are the political and economic risk scores calculated by converting the *Euromoney* scores to match the system described in the previous section. The conversion is based on the formulae $P = 4(25 - S_1)$ and $E = 4(25 - S_2)$. Now, the higher the score, the higher the risk.

We can use Equation (5.3) to calculate the country risk scores by assigning values to the weights λ_P and λ_E. We obtain the country risk scores under $S(P)$ if we let $\lambda_P = 0.7$ and $\lambda_E = 0.3$, that is for MNCs assigning greater weight to political risk. Alternatively, we obtain the country risk scores under $S(E)$ if we let $\lambda_P = 0.3$ and $\lambda_E = 0.7$, that is for MNCs assigning greater weight to economic risk. According to $S(P)$ and $S(E)$, Switzerland has the lowest country risk, whereas Somalia has the highest country risk.

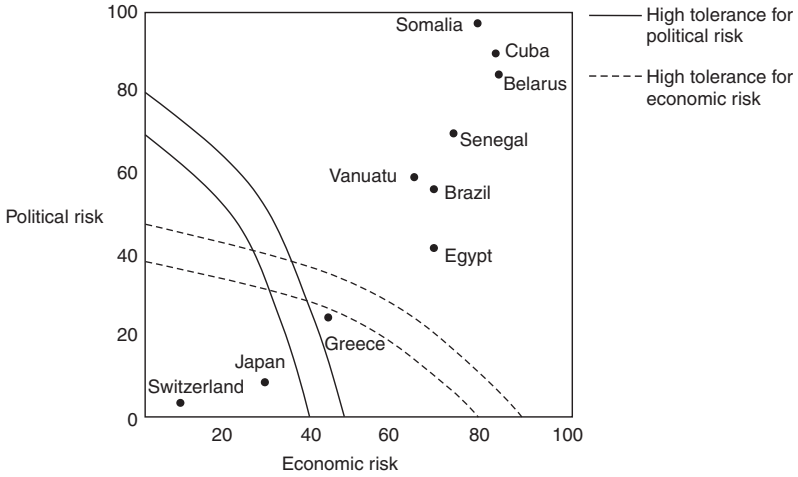

Figure 5.4 Foreign investment risk diagram with actual scores

We can use the data on P and E to draw investment risk diagrams for a firm with a high tolerance of political risk and for a firm with a high tolerance of economic risk. These are shown in Figure 5.4. The firm with the high tolerance of political risk accepts a maximum amount of political risk of 70 with an uncertain range of 70–80, while accepting a maximum amount of economic risk of 40, with an uncertain range of 40–50. The firm with the high tolerance for economic risk accepts a maximum amount of economic risk of 80 with an uncertain range of 80–90, while it accepts a maximum amount of political risk of 40, with an uncertain range of 40–50. For the first firm, only two countries fall in the acceptable zone: Switzerland and Japan. For the second firm, a third country is acceptable, namely, Greece.

THE USE OF COUNTRY RISK ANALYSIS

Once the degree of country risk has been determined, the next step is to decide whether or not the risk is tolerable. If it is felt that the risk is too high then the firm does not need to analyse the feasibility of a project to be undertaken in that country. Of course, it may be argued that no risk is too high if the rate of return on the underlying project is high enough to compensate for the risk. However, in some cases, the

risk is considered to be so high that the country is deemed to be off limits. This would be the case if, for example, the country has a tendency to experience civil war or kidnapping of foreign personnel for ransom. This would also be the case if the probability of confiscation is too high as judged from historical experience.

Country risk analysis can also be incorporated in capital budgeting analysis, as suggested by Robock (1971), Kobrin (1979), Raddock (1986), Sethi and Luther (1986) and Clark (1997). One way to do this is by adjusting the discount rate or the cash flows. The higher the country risk, the higher the discount rate applied to the project's cash flows. If, for example, blocked funds are anticipated, then the discount rate may be raised from 10 per cent to 13 per cent. The problem with this procedure is that there is no precise formula for adjusting the discount rate for country risk, which makes adjustment rather arbitrary. The use of a shorter payback period can be used for the same purpose. However, Haendel (1979) argues that neither of these two methods provides a detailed examination of the risk involved, or a true reflection of the investor's fear. This is why it may be preferable to incorporate country risk analysis by adjusting the cash flows, as suggested by Shapiro (1992).

Suppose that a project is analysed under three scenarios derived from country risk analysis: (i) that nothing will happen; (ii) that the host country will block a certain percentage of the funds to be transferred to the parent firm; and (iii) that the project will be confiscated after few years. Suppose also that these scenarios produce three net present values (NPV_1, NPV_2 and NPV_3) with probabilities of p_1, p_2 and p_3 respectively. The NPV of the project in this case should be calculated as the expected NPV, which is the weighted average of NPV_1, NPV_2 and NPV_3, where the weights are the probabilities. Hence:

$$NPV = \sum_{i=1}^{3} p_i NPV_i \tag{5.6}$$

Levi (1990) has suggested the following formalization of the process whereby country risk is allowed for in capital budgeting. Let X_t be the cash flow expected to arise from a project in the absence of country risk. Assume that country risk is present, so that the project would cease to exist (for example, because of a take-over by the host government) at year t with a probability p. Hence the probability that a cash flow arises in each individual year is $1 - p$. This means that the probability that cash flows arise for t years is $(1 - p)^t$. The expected

value of the cash flow in year t is $X_t(1-p)^t$. If X_t is constant so that $X_t = \bar{X}$, then the present value of the cash flow is given by:

$$PV = \sum_{t=1}^{n} \frac{(1-p)^t X_t}{(1+k)^t} = \bar{X} \sum_{t=1}^{n} \frac{(1-p)^t}{(1+k)^t} \tag{5.7}$$

As $n \to \infty$, we obtain:

$$\sum_{t=1}^{n} \frac{(1-p)^t}{(1+k)^t} = \frac{1-p}{k+p} \tag{5.8}$$

Hence:

$$PV = \frac{\bar{X}(1-p)}{k+p} \tag{5.9}$$

In the absence of country risk, we have:

$$PV = \frac{\bar{X}}{k} \tag{5.10}$$

It is obvious that the present value of the cash flows in the absence of country risk is greater than that obtained when country risk is present.

Agmon (1985) has suggested a workable and comprehensive way to integrate country risk into capital budgeting. Agmon's method is based on the proposition that the potential dependence of a project on the external environment is divided into two components: vulnerability and cost. Vulnerability is expressed in terms of the probability that an event that is likely to affect the project (such as tax changes) will occur. Vulnerability is also defined in terms of a probability distribution, and for simplicity it is assumed that the distribution can be described fully by its first and second moments. Cost, on the other hand, is measured as the actual impact on the cash flows of the project if a given event occurs. The distinction between vulnerability and cost is crucial, because firms do not need to be concerned with all possible events. Only the non-trivial effects on cash flows have to be weighted by the probabilities that they will take place.

In general, there are three approaches to the integration of country risk into capital budgeting. The first is the approach suggested by Shapiro (1992), which involves the adjustment of the expected cash flows to account for losses caused by country risk. Buckley's (1996) approach is a refinement of Shapiro's approach in the sense that

Buckley attempts to devise a general formula for integrating country risk into capital budgeting. The second approach has been suggested by Clark (1997). This approach involves the measurement of the effects of country risk on the outcome of investment as the value of an insurance policy that reimburses all losses resulting from an event. The third approach, which has been suggested by Mahajan (1990), is based on the use of option valuation theory to derive the pricing of country risk, particularly the risk of expropriation. Wafo (1998) justifies the use of this approach by the following argument: if the likelihood of expropriation depends on project outcomes, then the only proper valuation technique is contingent claims analysis. Wafo also argues that standard valuation approaches, such as the adjustment of future cash flows, are adequate only when the probability of expropriation is independent of the value of the project.

POLITICAL RISK: THE DEFINITION ONCE MORE

Political risk is the part of country risk that is normally more relevant to FDI. Most of the literature on FDI refers to political risk, and not to country risk, which may be more relevant to international lending by banks (see, for example, Nagy, 1979). Because of the special importance of political risk, the rest of this chapter is devoted to a more elaborate treatment of this kind of risk.

The concept of political risk has been around for some time. Baskin and Miranti (1997) suggest that attempts were made in the seventeenth century to deal with various sorts of risk, including political risk. The concept of political risk started to appear in the literature when it became important in the 1960s as newly independent countries tried to overcome their capital shortage problems by taking over the MNCs. Thus, expropriation and nationalization became critical concerns in the 1960s for companies with foreign operations. After the ousting of the Shah of Iran in 1979, political instability was added to confiscation, nationalization and expropriation as a source of political risk. For example, four months before the removal of the Shah, a country risk assessor for Gulf Oil detected severe political pressure building within Iran. As a result, the company began planning to deal with the subsequent loss of Iranian oil, which at one time amounted to 10 per cent of its crude supplies.

The increasing realization of the importance of political risk led to attempts to quantify it. Rummel and Heenan (1978) conducted the

first study that dealt with the assessment of political risk, in which they proposed a method for converting political instability into probabilistic terms, thus providing a scientific definition of political risk. Another attempt to define and assess political risk is represented by the development, by Theodore Haner, of the business environment risk information index (BERI) as a quantitative guide to political risk ratings. In 1979, William Copli and Michael O'Leary began to develop the Political Risk Services (PRS) evaluation system, which was used widely by MNCs in the 1980s. Since then, new approaches to political risk have been developed, with the aim of quantifying political risk and integrating it into the decision-making process of an enterprise.

Wafo (1998) summarizes the historical evolution of the concept of political risk as follows:

1. In the period from antiquity to the 1960s, there was neither an elaborate concept of political risk nor a precise political risk consciousness, although the manifestations of political risk were taken into account.
2. In the 1970s, the birth of the concept of political risk was linked to the spread of collective doctrines (nationalism and Marxism).
3. In the period since the 1980s, the concept of political risk has assumed increasing importance for academics and professionals alike. This period has also witnessed the birth of quantitative risk assessment methods, the probabilistic interpretation of political risk, and the systematic use of these quantitative approaches on the corporate level by the professionals.
4. The period since 1990 has witnessed the scientific refinement of the concept of political risk through the contribution of other fields of research such as political science, sociology, decision theory and psychology.

Let us now turn to the task of finding a precise definition for political risk. Before analysing political risk it must be defined, and it must be defined precisely. However, despite the widespread coverage of the subject, political risk has not received a clear-cut consensus definition. Indeed, Rummel and Heenan (1978) refer to political risk as 'one of the most misunderstood and misinterpreted aspects of multinational operations'. The literature contains various definitions of political risk. It can be defined as 'unpredictable demands raised by foreign state or society on the assets, returns or cash variable of shareholders from international investment' (Wafo, 1998). Robock and Simmonds

Table 5.3 Sources of political risk

Source	*Manifestation*
Restrictions on entry of foreign investors	Restrictions on entry of foreign investors Restrictions on the types of enterprises that foreign investors may undertake Restrictions on ownership
Systems for controlling FDI inflows	Outright ban on selected industries Vague criteria on official approval of FDI High taxes and weak incentives Equity limits Local content requirement rules
Limits on foreign exchange transfers	Restrictions on capital and profit repatriation Long repatriation deals Limits on repatriation by net worth Limits on repatriation by foreign exchange earnings
Government intervention	Price controls Setting prices for natural resource expropriation Regulating monopolies Large state enterprises sector
Social instability	Fragile political structures Weak organizational level in the society Corruption
Political violence	Crime Civil war Civil disobedience Riots
Government incapability	Inability to regulate the economy and conduct reforms Lack of democratic institutions and spirit
Turbulent relations with international organizations	Turbulent relations with the IMF and the World Bank Turbulent relations with the UN
Lack of commitment to international environment and labour rules	
Turbulent relations with foreign investors in the previous five years	Lack of commitment to bilateral investment rules Subtle expropriation of gained returns
Hostile attitude of elites and society towards FDI	Hostile declarations of parties Hostile programmes
Hostile attitudes towards foreigners	Violence against foreigners Intolerance towards foreigners Restriction on expatriate labour
Reluctance of host countries to reveal reliable information	Lack of transparency Secrecy with respect to most political and economic decisions

Source: Wafo (1998).

(1973) view political risk as existing in international investment when discontinuities occur in the business environment as a result of unanticipated political change. Another definition is presented by Haendel (1979), who defines political risk as the risk or probability of occurrence of some political event(s) that will change the prospects for the profitability of a given investment.

A simple and straightforward definition of political risk is that it is that part of country risk related to political factors. We have already stated a number of political factors that give rise to political risk. Now, we look at a comprehensive list of political risk factors as presented by Wafo (1998). The political risk factors or sources of political risk are classified under thirteen different headings, as shown in Table 5.3. The terms are mostly self-explanatory.

Political risk may be classified according to two criteria: (i) incidence; and (ii) the manner in which political events affect the investing firms. If incidence is used as a criterion for classification, then political risk can be classified into macro risk and micro risk. Macro risk pertains to the situations when FDI is affected by adverse political developments in the host country, irrespective of the nature of the investor or the underlying project. Micro risk, on the other hand, implies that only selected areas of FDI or particular firms are affected. For example, a military take-over of the government in the host country is a macro political risk, but new legislation restricting foreign ownership in the mining sector is a micro risk, affecting only firms involved in the mining industry.

If the manner in which firms are affected is used as a criterion for classification, then political risk can be classified into three kinds:

1. Transfer risk, which arises from uncertainty about cross-border flows of capital, payments, know-how and so on. One example of transfer risk is the unexpected imposition of capital controls, and withholding taxes on dividends and interest payments.
2. Operational risk, which is associated with uncertainty about the host country's policies affecting FDI. One example of operational risk is changes in environmental policies and minimum wage legislation.
3. Control risk, which arises from uncertainty about the host country's policy regarding ownership and control operations. An example is the nationalization of the local operations of MNCs. This happened, for example, when Mao Zedong took power in China in 1994, and when Castro took over in Cuba in 1960. In the

1950s, Gamal Abdul Nasser nationalized the Suez Canal, which was controlled by British and French interests, eventually leading to the invasion of Egypt by Britain, France and Israel in 1956.

POLITICAL RISK INDICATORS

Political risk indicators are designed to give the corporate manager some idea about the political environment in a country. They are designed to forecast the future political evolution of a particular country. Two kinds of political indicators are generally available: general indicators and partial indicators. General political risk indicators are based on all available political variables, designed to describe fully the political situation in a country. Partial political risk indicators are based on one political variable such as democracy, political violence or government stability.

Rummel and Heenan (1978) describe the general political risk indicators as aiming at anticipating changes in government policies and personnel, and at providing qualitative evaluation of the positions

Table 5.4 Political risk elements in the PRI

Category	*Factor*
Internal causes of political risk	Fractionalization of the political spectrum and the power of these fractions
	Fractionalization by language, ethnic and/or religious groups and the power of these fractions
	Restrictive (coercive) measures required to retain power
	Mentality, including xenophobia, nationalism, corruption, nepotism, willingness to compromise
	Social conditions, including population density and wealth distribution
	Organization and strength of forces for a radical left-wing government
External causes of political risk	Dependence on and/or importance to a hostile major power
	Negative influence of regional political forces
Symptoms of political risk	Social conflict involving demonstrations, strikes and street violence
	Instability as perceived by non-constitutional changes, assassinations, and guerrilla wars

Source: Wafo (1998).

to be taken by opposition parties or pressure groups. One of the most frequently used general political risk indicators is the business risk environment information index (BERI). In constructing this index, socio-political changes are measured by the political risk indicator (PRI) and constitute only one part of the country risk assessment. The PRI is based on ten elements classified under three headings, as shown in Table 5.4.

The PRI takes values ranging between 0 (the lowest rating, and thus the highest risk) and 100 (the highest rating, and hence the lowest risk). A country with a PRI of 70–100 has a low risk in the sense that political changes will not lead to conditions that are seriously adverse to business. A PRI in the range 55–69 indicates moderate risk in the sense that political changes that are seriously adverse to business have occurred but, under the government in power, there is a low probability that such changes will occur in the future. A PRI of 40–54 implies high risk, in the sense that political developments with serious consequences for business, exist or could materialize in the future. Finally, a PRI of 0–39 implies prohibitive risk in the sense that political conditions severely restrict business operations and that a loss of assets is possible. This scaling is represented diagrammatically in Figure 5.5. Another general indicator of political risk is the world political risk forecasts (WPRF). This indicator is heavily based on expert opinion which is used to build up a predictive political view of the underlying country.

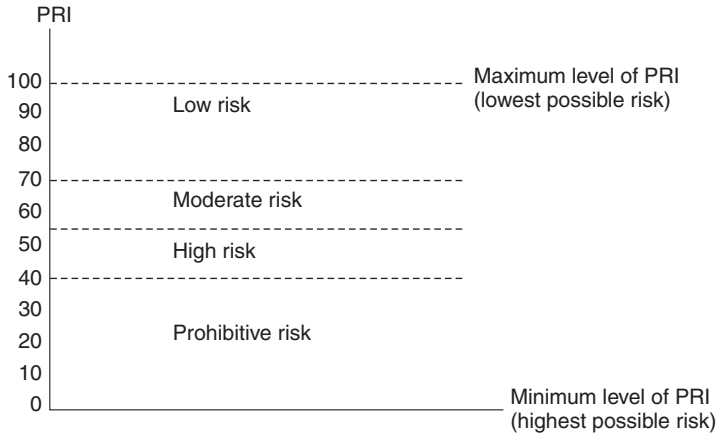

Figure 5.5 Interpretation of the values of PRI

Partial political indicators pertain to, or are based on, one part of the political system. Examples of partial political indicators are the government stability indicators and those measuring democracy, political violence, political volatility, and corruption. There are, for example, two methods for calculating indicators of government stability from published data. The first is based on the average, over a given period of time, of such incidents as coups and government changes. The second approach is based on the probability of a government change, estimated from a probit regression.

Political risk indicators are used to monitor political risk after a project has been taken. The political environment in any country is dynamic in the sense that it changes over time. Changes taking place over time include changes in the laws and regulations affecting FDI in general, or the underlying project in particular, as well as changes in public opinion. This is why the monitoring of political risk after a project has been implemented is crucial. Haendel (1979) suggests that the monitoring of political risk can be accomplished by using Bayesian analysis, which serves to revive (on the basis of new information) the prior probabilities used initially to calculate the NPV. Schweitzer (1976) describes how the CIA used Bayesian analysis to predict political changes in the Middle East.

Monitoring political risk also means explaining the nature of the relationship between the host country and the investor. In this sense, FDI decisions are seen as resulting from negotiations between the host country and the investor. One way to understand the nature of the relationship between the two parties is by resorting to Vernon's (1968) theory of 'obsolescing bargaining'. This theory essentially reflects the concerns of foreign investors in developing countries over the possibility that government policies might easily change *ex post* (that is, after the implementation of the project) in ways that are detrimental to the profitability of the investment. The theory refers to changes in the relative bargaining power of the country and the investor over time. A variety of considerations suggest that the bargaining position of the host country generally strengthens over time relative to that of the investor. One explanation for this proposition is based on informational asymmetry, as the host country has access to all information needed by the investor, to which the investor does not have access. However, given the financial power of MNCs and their full control over technology transfer, any increase in the host country's bargaining power will most probably be marginal.

THE EFFECTS AND MANAGEMENT OF POLITICAL RISK

Research has shown that the main determinants of the decision to invest abroad are market prospects and risk factors. Political risk is obviously part of the risk considerations, and this is particularly the case with investment in less developed countries. Burgmann (1996) concludes that some factors contribute to the increased risk of MNCs compared to domestic firms, and these include foreign exchange risk, political risk and tax uncertainty. Ramcharran (1999) examined the effect of political and economic risk on FDI for twenty-six countries using cross-sectional data for the years 1992, 1993 and 1994. He found a significant relationship between FDI and these risks.

The management of political risk includes the identification and assessment of political risk, its quantification, the anticipation of losses and the prevention or reduction of the incidence of losses. Kelly (1974) identifies the following five steps in the management of the political risk associated with FDI:

1. The identification and analysis of loss exposure.
2. The measurement of exposure associated with these risks.
3. The development of alternative techniques for treating each exposure.
4. The reduction and implementation of the best technique or combination of techniques to forecast each exposure.
5. The evaluation of results in an effort to improve the procedures of identification, measurement and treatment.

There are in general two major ways of dealing with political risk: (i) risk avoidance or transfer; and (ii) risk negotiation. The first approach is based on the utilization of existing financial and non-financial methods as well as legal arrangements to avoid risk or to transfer it to a third party. The second approach is based on negotiating a deal with the host government. The two approaches are not mutually exclusive and can be employed simultaneously.

Political risk can be managed via some operating strategies and with the help of some international institutions. One of the most important operating strategies is carrying out FDI via joint ventures with a partner from the host country that can be a government body. This is the business form that is most used frequently in strategic alliances (Mariti and Smiley, 1983). The setting up of joint ventures with parties from the host countries is particularly useful when a firm

wants to enter a country with a hostile government and/or restrictive legislation. Wafo (1998, p. 58) concludes that the literature recognizes a close relationship between the entry mode (a joint venture or otherwise) and political risk. Buckley and Casson (2000b) identify three conditions under which the establishment of joint ventures becomes attractive, the third of which refers to political factors.

The second operating strategy for managing political risk is based on the choice of the capital structure and the source of capital. The objective of this strategy is to transfer risk to a third party. This can be done by financing FDI projects with funds from the host government, other governments, international development agencies, overseas banks, and customers. Debt service in this case will be met out of production, not out of the parent firm's funds. Should a take-over by the host government materialize, the net worth that is lost in the project would be minimal. Other strategies will be discussed later, with special reference to the risk of take-over by the host government.

International organizations, such as the World Bank and the OECD, are instrumental in the management of political risk. They have been active in their attempts to establish an international investment law. These two organizations have been considering four ways to reduce political risk in developing countries. These are:

1. Measures that aim at influencing the host government's investment policy.
2. Measures that deal with the creation, and setting up, of a new multilateral insurance agency for political risk.
3. The creation of an organization for arbitration and conciliation of investment disputes.
4. Attempts to reduce the effect of political risk by extending the rules of the World Trade Organization (WTO) to the international investment domain (such as the transparency of investment rules and the application of the concept of the 'most-favoured nations').

The World Bank has three agencies working in this field. One of these organizations is the International Finance Corporation (IFC), which began operations in 1956. Its objective is to promote economic development in Third World countries by supporting the private sector. The investment promotion activities of the IFC include an advisory service on policies and institutions that can assist members in attracting and regulating foreign investment. There is also the International Centre for Settlement of Investment Disputes (ICSID), created in

1956. Its function is to assist in the conciliation and arbitration of investment disputes between member countries. The third agency is the Multinational Investment Guarantee Agency (MIGA). The proposal to establish the MIGA was launched at the Seoul meetings of the World Bank and the International Monetary Fund in 1985. This agency provides insurance to cover the types of political risk that are beyond the control of the host government. Specifically, the MIGA offers long-term (up to 15 or 20 years) political risk insurance coverage to eligible investors for qualified investments in developing member countries. The particular kinds of risks covered by the MIGA are: (i) transfer restrictions; (ii) expropriation; (iii) breach of contract; and (iv) war and civil disturbances.

There are also multilateral agreements between countries that deal with political risk. Wafo (1998) presents a list of reasons that drive the desire to establish multilateral agreements: (i) an increased number of host countries having a more significant stake in these agreements; (ii) a more positive attitude towards FDI by many host countries since the 1980s; (iii) a body of bilateral treaties on FDI; and (iv) the presence of institutions that could serve as an acceptable forum in which to develop an international investment agreement, such as the World Bank.

The Multilateral Agreement on Investment (MAI), which was negotiated at the OECD, is the most important proposal of this kind. The negotiations ceased in 1999 without reaching any consensus on the form the agreement should take. The objective behind this agreement was to create a strong and comprehensive multilateral legal framework for FDI. It would also reduce barriers to FDI and increase legal security for international investors. The question that arises here concerns the need for an MAI, given that investment flows have been growing rapidly. The answer is that the unevenness of investment regimes causes foreign investment to flow to relatively few countries, particularly developing countries. One objective of the MAI is therefore to encourage FDI to flow to countries typically regarded as being off limits because of a higher level of political risk.

There is, however, another view of the MAI that portrays it as being biased against the host countries in favour of the foreign investors. Specifically, the objectives of the MAI, according to this view, are the following:

1. Opening all sectors of economies to foreign companies.
2. Requiring countries to treat foreign investors in the same way as local companies.

3. Banning performance requirements and measures of accountability.
4. Setting binding dispute settlement rules, allowing investors to challenge laws and seek financial compensation.

This sceptical view of the MAI raises the following concerns:

1. The MAI was negotiated with little public input. The wide range of concerns about international capital mobility, labour rights and financial stability were largely ignored.
2. Developing countries had no role in drafting the MAI.
3. Pre-emption of local state and national laws.
4. Corporate investors would get new rights but there are no mechanisms to hold them accountable to the social and environmental concerns of the source or recipient countries.
5. The MAI would contain weak, non-binding environmental provisions.

In short, the sceptical view of the MAI can be summarized as follows. The MAI would bring about an enormous transfer of power from governments to the boards of directors of the world's largest corporations. It would pose serious challenges to environmental and other social goals, including democratic governance. The MAI would increase the rights and opportunities of MNCs without increasing their responsibilities. And while the MAI would result in an increase in FDI, it would also result in a redistribution of benefits from the host countries to MNCs.

Political risk can also be managed via bilateral agreements. These agreements can take two forms: bilateral government treaties, and bilateral investment treaties. Bilateral government treaties are concluded with the objective of guaranteeing a commitment by the host government with respect to FDI. Bilateral investment treaties, on the other hand, are designed to improve the conditions for investment by firms of each signatory country in the other country. Bilateral investment treaties cover several issues of interest to the investor: treatment of the investor after an investment has been made in the host country, expropriation, compensation, and dispute settlement.

Furthermore, many countries have tried to deal with political risk by setting up government guarantee companies. An example is the US Investment Guarantee Program, managed by the semi-governmental Overseas Private Investment Corporation (OPIC). Typically, OPIC insures American investors against expropriation loss, inconvertibility

of currency, war, riots, and insurrection. Commenting on the impact of OPIC's activity, US Senator Frank Church is quoted as having said 'once the government assumes the insurance of the company, the company's interest and that of the government become identical, and the company can threaten to fall back on the U.S. government, whenever it deals with a foreign government' (US Congress Committee on Foreign Relations, 1973).[7] A similar function is performed in Japan by the Ministry of Economy, Trade and Industry (formerly, the Ministry of International Trade and Industry), in the UK by the Export Credits Guarantee Department (ECGD), and in Canada by the Export Development Corporation.

The most severe political risk is a host government take-over, which may take the form of confiscation or expropriation (without and with compensation, respectively). There are several strategies to deal with exposure to host government take-over. To start with, pre-negotiation with the host government is important. Areas of pre-negotiation include the size and form of the initial investment, the amount of local labour employed in the project (local versus imported), tax burden, ease and cost of local financing, remittance amount and composition (dividends, interest and fees), and protection given to the original investment and accumulated profits. One strategy is to use a short-term horizon, in which case the emphasis is placed on recovering cash flows quickly. Should a take-over take place, losses will be minimized. The MNC will also exert a minimum effort to replace worn-out equipment at the subsidiary. Another strategy is to use unique supplies or technology. In this case, the host government will have no incentive to take over the business, since the project will not be operational without the help of the investing firm. Schnitzer (1999) shows how investors can use their control rights to protect their investments from what he calls 'creeping expropriation' once the investment is sunk. The third strategy is to hire local labour. If local employees of the subsidiary are affected by the take-over, they could put pressure on their government to avoid such an action. A similar result would materialize if the project was financed by borrowing local funds. In this case, the local banks would exert pressure on the host government to avoid the take-over. The last protective measure would be to buy insurance from the OPIC and similar agencies.

Several studies have been conducted to measure the effect of political risk on FDI. These studies typically follow two approaches to the empirical quantification of the effect of political risk on FDI. The first is the genuine political-orientated approach, which stresses

the importance of political factors as determinants of FDI. The second approach, which is economic-orientated, considers political factors in conjunction with economic factors. In both of these approaches, FDI is viewed as being supply-determined (that is, determined by the decisions of the MNCs), whereas demand in the host countries is considered to be infinitely elastic. We have already come across some of these studies (Chapter 2).

Empirical studies based on surveys consistently found that executives consider host country political instability to be a major determinant of FDI project location decisions (see, for example, Basi, 1963; Aharoni, 1966; Root, 1968; Green, 1972 and Frank, 1980). However, studies based on the statistical analysis of data revealed mixed, if not contradictory, results: some have found significant effects of political risk on FDI flows, whereas others have not. For example, Brewer (1983) found a very weak correlation between government instability and restrictions on international funds associated with FDI projects. The relationship between expropriation and government instability also is not clear, as found by Kobrin (1984). And Levis (1979) found that political variables are of secondary importance as compared with economic factors in determining FDI flows. Finally, Schneider and Frey (1985) used a politico-economic model that incorporates a large number of possible determinants of FDI and concluded that FDI is determined simultaneously by economic and political factors.

A FINAL REMARK

In this chapter we examined the measurement and management of country risk and political risk, as well as the effects they exert on FDI-related decisions. While the empirical evidence on the effects of country risk and political risk is mixed (as is invariably the case with empirical testing), it is intuitively sound to imagine that an MNC will shy away from countries characterized by macroeconomic mismanagement, erratic government decision-making, frequent kidnapping of foreign business executives, and so on. In Chapter 6, another important factor for FDI will be examined: international taxation.

10 Summary and Conclusions

RECAPITULATION

This book has taken us on a grand tour of the economic, financial, accounting and management aspects of FDI. We have explored the characteristics, determinants and effects of FDI. We have also gone through the procedures used to determine the financial feasibility of FDI projects and investigated the effects on this feasibility of country risk, taxation and the cost of capital. Then we dealt with the critical issue of transfer pricing, and finally we examined the control and performance evaluation functions in MNCs. It is rather difficult to write a short but a comprehensive summary of what has been discussed, but it is possible to state the following points as some sort of a recapitulation:

1. FDI is the process whereby residents of one country acquire ownership of assets for the purpose of controlling the activities of a firm in another country. Interest in FDI results from (i) its rapid growth; (ii) the concern it raises over the causes and consequences of foreign ownership; and (iii) the fact that FDI has become an important source of funds for developing countries.
2. FDI can be horizontal, vertical or conglomerate. It can also be classified into import-substituting, export-increasing and government-initiated. Another classification is that of expansionary versus defensive FDI.
3. A common sequence that firms use to develop foreign markets for their products consists of (i) exports; (ii) licensing; (iii) foreign distribution; and (iv) foreign production. Steps (iii) and (iv) involve FDI.
4. Theories explaining FDI determination comprise (i) theories assuming perfect markets; (ii) theories assuming imperfect markets; (iii) other theories; and (iv) theories based on other variables. The empirical evidence on these theories is mixed, but the effect of certain individual variables has been established without any doubt (for example, the size of the host economy).

5. There are also theories designed to explain entry modes; that is, the choice among exports, licensing, franchising, subcontracting, M&As, greenfield FDI and joint ventures.
6. FDI gives rise to costs and benefits for the source and host countries. But there is a fundamental disagreement on what constitutes costs and benefits. The division of welfare gains between the two countries is influenced by the bargaining power over the terms of the agreement governing a particular FDI project.
7. The effects of FDI on the host country can be classified into economic, political and social effects. Whether these effects are favourable or adverse is a controversial matter, as they are conditions-specific.
8. International capital budgeting is the process used to evaluate the financial feasibility of FDI projects. Several criteria are used for this purpose, including the NPV, APV, IRR, accounting rate of return, the payback period, and the profitability index.
9. Country risk is exposure to loss in cross-country transactions. It consists of economic risk and political risk, the latter being more relevant to FDI.
10. Taxation significantly affects the location of FDI, the organizational form of the foreign establishment, financing decisions, remittance policy, transfer pricing policy, working capital management, and the capital structure policy.
11. The cost of capital is crucial for capital budgeting, since it determines the discount rate used in project evaluation. MNCs tend to have different cost of capital and capital structures from those of purely domestic firms.
12. Transfer pricing is the pricing of goods and services that are bought and sold between members of a corporate family. The transfer pricing policy followed by an MNC depends on tax considerations, global regulation, performance evaluation, fund positioning, market considerations, risk, government policy, interest in joint ventures and the negotiating power of the subsidiary. Transfer prices can be set on the basis of market prices or costs.
13. Control in MNCs consists of planning, implementation, evaluation and correction of performance in order to achieve organizational objectives. It involves several issues, including centralization versus decentralization. Performance evaluation is an essential part of the control system, comprising the comparison of results with pre-determined objectives.

This brief itemization of the subject matter covered in this book is inevitably superficial, but it serves the purpose of reminding ourselves what was dealt with in the previous nine chapters. Now, we turn to a brief summary of the reasons behind the initiation of FDI.

WHAT MOTIVATES FDI?

From one perspective, FDI can be viewed as an activity that satisfies basic business needs. Hence, there are firm-specific motivations for indulging in FDI. The following is a general outline of the reasons why FDI is undertaken. We came across these motivations when we studied the theories of FDI in Chapter 2.

The Need for Markets

When the growth of sales is limited in the domestic market, a firm considers expanding overseas. Under special circumstances, expansion takes place by establishing production facilities abroad, which allows the firm to exploit the economies of scale. If markets are segmented and there are markets that offer higher profit margins, then firms will move into these markets. Many MNCs give primacy in their business strategies to seeking markets, and a large number of cross-border investments are prompted by the need to expand and diversify markets on a global basis. This is particularly valid for the MNCs engaged in the production of consumer non-durable goods such as food processing, beverages, tobacco, and soap and toiletries. These MNCs tend to spend a relatively high percentage of sales revenue on advertising, and rely on a strong and distinct marketing effort to maintain or increase market position (for example, Coca-Cola sells its products in more than 160 countries).

The Need for Production Efficiency

The need for production efficiency motivates companies to produce in countries where resource inputs are relatively cheap. This is why American companies invest in Mexico and South East Asia for labour-intensive production. Nike, for example, carries out the bulk of its production of training shoes in Asia, paying workers very low wages.

The Need for Raw Materials

To avoid the transportation costs associated with importing raw materials that are unavailable in the home country, a firm will set up a production facility close to the source of raw materials in the foreign country. This is particularly the case if the foreign country is a market for the finished products. This factor underlies FDI in many industrial sectors such as petroleum, metals mining, forest products, and plantation activities. Countries that have served as hosts for this kind of FDI are Canada, Australia, Malaysia and Chile.

The Need for Information and Technology

The need for information and technology has motivated FDI in US companies manufacturing computer software and hardware, as well as pharmaceutical companies.

The Need to Minimize or Diversify Risk

The risk factor has led to FDI in countries being considered more stable and offering low political risk. This explains why some Hong Kong-based MNCs sought to diversify their operations because of uncertainties regarding the return of Hong Kong to China in 1997. Market risk can be reduced by investing in several countries.

Integrating Operations

Vertical integration is achieved when a firm indulges in various stages of the production process. For example, vertical integration is achieved in the oil industry when a firm engages in operations ranging from exploration and extraction to retailing. It is beneficial because it results in assured delivery between various stages of production. If different stages can be carried out better in different locations, then expansion in foreign countries will take place.

Non-transferable Knowledge

If a firm develops expertise in the production of a certain commodity, and if it is difficult to transfer this knowledge, the firm will be better off expanding overseas.

Protecting Knowledge

A firm may have transferable expertise that it does not wish to transfer. In this case the firm itself will take production overseas. An example is the desire by Kentucky Fried Chicken to protect the secrecy of its recipe.

Protecting Reputation

To protect a brand name or product quality, a firm may decide to carry out production abroad. For example, McDonalds expanded overseas to protect its reputation of producing 'tasty' burgers.

Capitalizing on Reputation

Firms with an international (good) reputation may capitalize on it by expanding overseas. International well-known banks, for example, can attract deposits when they set up branches in foreign countries.

Avoiding Tariffs and Quotas

A good example is when Japanese car manufacturers establish production facilities in the USA. When production takes place in the foreign country it will be a substitute for exports that may be subject to tariffs and quotas.

Exchange Rate Considerations

Firms move into countries with weak currencies because the initial set-up cost is low. In general, firms choose to establish production facilities in countries whose currencies are weak or undervalued, and sell their products in countries whose currencies are strong or overvalued. The extent of overvaluation and undervaluation may be measured by deviations from purchasing power parity (see, for example, Moosa, 1998, Chapter 9).

Relationships with other MNCs

Some MNCs expand overseas because they follow others. This is particularly valid for banks and other financial institutions. If an MNC decides to establish a subsidiary in a particular foreign country, the bank or financial institution serving the parent company at home

may open a branch in the same foreign country to service the banking and financial needs of the subsidiary. This, in fact, was one of the reasons for the growth of international banking.

An Example

Some of these motives have been used to explain certain cases of FDI. Eng *et al*. (1995, p. 451) give the following explanation for the growth of British FDI in the USA since the 1980s:

1. Recognition of the size of the US market, and that the US is a safe haven in a turbulent world.
2. Emergence of non-US MNCs with the ability to compete successfully in the US market.
3. Depreciation of the US dollar against major currencies in the second half of the 1980s.
4. Narrowing of the gap in production costs between the USA and foreign locations.
5. Concern regarding possible US protectionist measures.
6. Relatively non-restrictive US policy towards inward FDI, together with the active promotion of FDI by individual state governments.

FDI AND MNCs: THE ARGUMENTS FOR AND AGAINST

We have by now realized that the effects of FDI and MNCs operating in foreign countries constitute a highly controversial issue. In an attempt to reach a verdict on whether they are good or bad, let us sum up the arguments for and against FDI and MNCs, starting with the arguments *for* them:

- FDI flows continue to expand even when world trade slows down, or when portfolio investment dries up. These flows are less volatile than portfolio investment flows, because FDI represents a long-term commitment to the underlying project.
- FDI is an important source of funds for developing and transition countries.
- It involves the transfer of financial capital, technology and other skills desperately needed by developing countries.
- FDI raises income and social welfare in the host country unless there are distortions caused by protection, monopoly and externalities.

- It contributes to filling the saving and foreign exchange gaps by providing financial capital.
- It provides a vehicle for reviving the domestic capital market through which domestic savings can be channelled to finance domestic investment.
- FDI boosts growth in the host countries through technology diffusion and the transfer of capital.
- FDI can lead to an increase in employment in the host country by setting up new production facilities and by stimulating employment in distribution. It can preserve employment by acquiring and restructuring ailing firms.
- FDI initially leads to an improvement in the capital account of the host country.
- FDI is likely to boost productivity if (i) it is export-promoting; and (ii) the underlying conditions allow the installation of plants designed to achieve economies of scale.
- FDI boosts the skills of local workers through training.
- FDI helps to provide local firms with increased opportunities by establishing links with local suppliers for locally-produced goods.
- FDI boosts competition in the host country's markets.

This list looks very impressive, but the list of arguments *against* FDI and MNCs is equally, if not more, impressive. Let us examine this list:

- FDI symbolizes new colonialism.
- It results in a loss of sovereignty and in compromising national security. There are several examples of MNCs interfering with the politics of the host country.
- MNCs are often in a position to obtain incentives (from the host country) in excess of their needs and perhaps in excess of the benefits they bring to the host country.
- MNCs exist and operate primarily because of market imperfections, which precludes the conditions under which FDI supposedly boosts welfare.
- Even if FDI leads to a gain in world output, it results in some distributional changes between labour and capital.
- The sheer size of MNCs may jeopardize the national independence of the host country.
- FDI creates enclaves and a foreign elite in the host country.
- It introduces adverse cultural changes.

- FDI does not perform the function of providing capital, for three reasons: (i) it is a relatively expensive source of financial capital; (ii) actual capital flows provided by MNCs may not be large, as they may choose to obtain funds from the local capital market; and (iii) the capital contribution of MNCs may take a non-financial form (for example, goodwill). By raising capital locally, MNCs crowd out domestic investment, which is perhaps more suitable than foreign investment.
- The domination of a developing country by an MNC may economically be detrimental to growth through a lower rate of capital accumulation, greater incidence of undesirable practices and adverse effects on competition.
- It is invariably the case that subsidiaries operating in host countries are wholly-owned by the parent MNCs. The host country has no control over the operations of these subsidiaries.
- FDI can reduce employment through divestment and closure of production facilities. The empirical evidence shows that the overall employment effects of the activities of MNCs on the host country are small.
- Outward FDI destroys jobs at the source country because output of foreign subsidiaries becomes a substitute for exports from the home country.
- FDI leads to an increase in wage inequality in the host country.
- FDI is often blamed for its balance of payments effects. The source country faces a sudden deficit when the FDI occurs, whereas the host country faces a perpetual deficit because of profit repatriation.
- MNCs indulge in the production of luxury goods rather than the basic consumer goods needed in developing countries.
- FDI does not play an important role in technology diffusion because (i) inappropriateness of the technology they provide, as it is too capital-intensive; and (ii) the availability of cheaper sources of technology. Moreover, the R&D activities are concentrated in the MNCs' home countries.
- MNCs are very powerful negotiators, likely to strike favourable terms in bilateral negotiations with the government of a poor country.
- The costs of training labour are not large enough to make a significant contribution to the improvement of the skills of local workers. The practices of MNCs may be irrelevant to the host country, in which case training will not be useful and may even be

harmful. Moreover, it is often the case that MNCs reserve key managerial and technical positions for expatriates.

- MNCs worsen income distribution in the host countries.
- They also worsen income distribution worldwide by paying foreign workers low wages, charging ordinary consumers high (sometimes extortionate) prices, and paying 'celebrities' obscene amounts of money to sponsor their products. For example, it is reported that sportsmen Michael Schumacher and Tiger Woods earned more than US$100 million each in the year 2000 alone, most of which was sponsorship money.
- They abuse transfer pricing, depriving host countries of tax revenue. Reportedly, a subsidiary of an MNC operating in a particular country has not made any profit for over thirty years, which makes one wonder why this subsidiary is still operating in the same country.
- They form alliances with corrupt elites in developing countries. There are several examples of billionaires who migrate to developed countries after accumulating massive wealth in extremely poor countries. It is no wonder that the OECD has encouraged MNCs not to indulge in any activity involving corruption and bribes.
- Most MNCs are sufficiently vertically integrated or have incentives to engage in inter-subsidiary transactions that limit the scope for developing strong ties with local suppliers.
- FDI leads to a worsening of market concentration and the possibility of monopolistic or oligopolistic practices.
- Because MNCs have significant financial, political and negotiating power, they avoid blame for a lot of damage to the environment in developing countries.

Obviously, most of the items in this list pertain to developing host countries. Let us now sum up by making an attempt to reach a verdict on FDI and MNCs.

THE VERDICT

It is rather difficult to reach a clear-cut verdict on such a controversial issue. On the one hand, MNCs do not operate in host countries just to help these countries to overcome their economic problems. On the other hand, there is no reason why MNCs operating in host countries

with the objective of profit maximization do not produce any positive externalities (self-interest may lead to public good, according to Adam Smith's *Wealth of Nations*). It is, of course, unthinkable that a country should isolate itself from FDI flows just because of the list of arguments against FDI and MNCs that we have just been through. The truth must be somewhere in between the extreme views on this contentious and politically-sensitive issue.

Whether MNCs are good or bad is a normative question that cannot be value-free and objective. One may present views, on either side of the debate, that may be consistent, lack consistency, or be somewhere in between. One may also produce arguments with varying degrees of power of persuasion. However, it may be rather difficult to prove what is right and what is wrong. Empirical evidence may be called upon in this case, but the problem is that this is Economics and not Physics. Empirical evidence in Economics is invariably mixed and often unreliable. The results of empirical testing in economics are typically too fragile and too sensitive to subjective factors to be reasonably reliable.

One can only say at this stage that there is a role for FDI and MNCs to play in the development process. Perhaps the starting point is for MNCs to follow the extensive set of recommendations proposed by the OECD on how they should deal with the host countries. These recommendations provide a highly useful and positive code of conduct. But, it is undoubtedly the case that anti-MNC and anti-globalization protests will continue in the future as long as the activities of MNCs lead to widening the gap between rich and poor, and to massive environmental damage. This topic will most certainly remain as controversial as ever.