Trabajo realizado por el equipo de la Biblioteca Digital de la Fundación Universitaria San Pablo-CEU

Me comprometo a utilizar esta copia privada sin finalidad lucrativa, para fines de investigación y docencia, de acuerdo con el art. 37 del T.R.L.P.I. (Texto Refundido de la Ley de Propiedad Intelectual del 12 abril 1996)
Theories of the Determinants of Direct Foreign Investment

Giorgio Ragazzi *

A DIRECT FOREIGN INVESTMENT is the amount invested by residents of a country in a foreign enterprise over which they have effective control. The value of direct investment may increase not only through investment of new funds, which may be remitted from abroad or borrowed locally by the foreign investor, but also through reinvestment of earnings, or the sale to the foreign affiliate of nonfinancial assets, such as a license or management services. Changes in outstanding direct investment, therefore, normally differ from international financial flows as recorded in the presentation of most balance of payments. Until the 1960s, direct foreign investment was usually considered as just one form of international capital movement, responding to differences in rates of return on capital. However, in view of the enormous development of direct investment in the postwar period, which is described briefly in Section I, this explanation has appeared to be increasingly inadequate.

While portfolio investment abroad is made to a large extent by individual investors, direct investment is made essentially by corporations. Determinants of the two types of capital flow may thus differ insofar as the objectives and constraints of the two types of investor are different. However, the interrelationship between the two types of capital flow remains a crucial point and is discussed in Section II. Under perfectly competitive conditions, markets for securities would provide a more efficient way to transfer capital internationally than would direct investment, because local enterprises could presumably operate at lower costs in their own country than could foreign firms. The determinants of direct investment must therefore be found in actual deviations from

* Mr. Ragazzi, a graduate of the University of Torino and the University of Virginia, was an economist in the Eastern European Division of the European Department of the Fund when this article was written.

471
perfectly competitive conditions. The main line of modern theory of direct investment, which is reviewed in Section IV, focuses on advantages of superior "knowledge" that allow the foreign firm to obtain higher rates of return than local competitors. Other authors have stressed oligopolistic behavior and the maximization of growth rather than profit as the main determinants of direct investment (Section V). One aspect that has been relatively neglected in the literature is that of imperfections in capital markets, which cause discrepancies between "industrial" risks and rates of return of enterprises and the risks and rates of return implicit in minority holdings of their securities. Insofar as these discrepancies exist, portfolio capital movements may be expected to tend to equalize internationally the rates of return on securities, at equal risk, while direct foreign investment tends to equalize rates of return of enterprises, at equal risk. This argument, which is analyzed in Section III, may help to explain also why the flow of U.S. direct investment in Europe, in the postwar period, has been matched by an opposite flow of European portfolio investment in U.S. corporate securities.

The impact of customs duties and exchange risks on direct investment is analyzed in Section VI, where it is argued that the reserve role of the dollar and its overvaluation compared with most European currencies may have provided substantial incentives for U.S. direct investment in Europe.

I. International Capital Movements and the U.S. Investment Position in the Postwar Period

Until World War I, international capital movements consisted essentially of flows of portfolio capital from a few developed European nations to the rest of the world. These movements were determined mainly by international differences in rates of interest. Direct foreign investment was relatively unimportant and was generally limited to the exploitation of mineral resources in less developed areas. In the inter-

war period, the scope for international movements of capital was limited, and the net value of direct foreign investment actually tended to decline. Since World War II, there have been not only a rapid expansion of international capital movements but also two major qualitative changes: a much greater interdependence between financial markets with large interflows of capital among developed countries, and an enormous growth in direct foreign investment, which has often become the principal vehicle for international capital movements.

It is interesting in particular to consider changes in the net foreign investment position of the United States, since it has accounted for more than two thirds of the total world outflow for direct investment (Diamond [8]), and a large part of total gross capital movements has originated from, or has been directed to, the United States. Between 1950 and 1970, U.S. private investment abroad increased sixfold, from US$19 billion to US$120 billion, while U.S. liabilities to private foreigners increased at a comparable rate, from US$13 billion to US$70 billion. The composition of U.S. assets, however, is substantially different from that of U.S. liabilities: direct foreign investment is by far the largest asset item, whereas most U.S. liabilities are corporate securities and short-term liabilities. During the 1960s, U.S. direct investment abroad accounted for 65 per cent of the total increase of U.S. assets abroad, while foreign direct investment in the United States accounted for less than 15 per cent of the increase of the U.S. liabilities to private foreigners (Table 1). These figures suggest that direct foreign investment cannot be considered simply as one channel through which a country, richly endowed with capital, exports capital to other countries, but it must be analyzed within the larger framework of interflows of various types of capital among different countries.

In particular, for a number of reasons, it is useful to concentrate on the net U.S. investment position vis-à-vis Western Europe. First, in the 1960s U.S. direct investment in Western Europe expanded much faster than in other areas, increasing from 14 per cent to 31 per cent of total U.S. direct investment. Second, it consisted mainly of investment in manufacturing, that is, it was not related to the exploitation of local raw materials. Third, international capital movements between the

1 It is estimated that in 1914 investments abroad of the United Kingdom amounted to £4.4 billion. About 40 per cent of this consisted of portfolio shares, 30 per cent was in government and municipal foreign and imperial railways, 10 per cent was in raw material industries, and 8 per cent was in bonds, 10 per cent was in shares in manufacturing, and 6 per cent was in banking and finance. Tugendhat [22]. The numbers in square brackets refer to items listed in the References, pp. 497–98.

2 Although these changes do not correspond to capital flows as recorded in the balance of payments statistics, conceptually the two may be reconciled, for instance, for direct investment, by considering reinvested earnings as a current account receipt matched by an equal outflow of capital (Table 3).
Table 1. United States: International Investment Position, Excluding U.S.
Government Assets and Liabilities and Official Reserve Assets, 1960-70
(In billions of U.S. dollars)

<table>
<thead>
<tr>
<th>Private U.S. Assets Abroad</th>
<th>U.S. Liabilities to Private Foreigners</th>
</tr>
</thead>
<tbody>
<tr>
<td>World total</td>
<td></td>
</tr>
<tr>
<td>Direct investment</td>
<td>31.9</td>
</tr>
<tr>
<td>Bonds</td>
<td>5.6</td>
</tr>
<tr>
<td>Corporate stocks</td>
<td>4.0</td>
</tr>
<tr>
<td>Other long-term assets</td>
<td>3.1</td>
</tr>
<tr>
<td>Total, long-term</td>
<td>44.6</td>
</tr>
<tr>
<td>Short-term assets</td>
<td></td>
</tr>
<tr>
<td>Banks</td>
<td>3.6</td>
</tr>
<tr>
<td>Other</td>
<td>1.2</td>
</tr>
<tr>
<td>Grand total</td>
<td>49.4</td>
</tr>
</tbody>
</table>

Western Europe

| Direct investment          | 6.6   | 14.0  | 24.5  | 4.7   | 6.1   | 9.5   |
| Bonds                      | 2.2   | 0.9   | 0.5   | 7.3   | 10.5  | 12.6  |
| Corporate stocks           | 1.1   | 2.4   | 2.0   | 1.0   | 1.1   | 4.3   |
| Other long-term assets     |       |       |       |       |       |       |
| Total, long-term           | 9.9   | 19.2  | 29.6  | 13.0  | 18.3  | 31.7  |
| Short-term assets          |       |       |       |       |       |       |
| Banks                      | 1.3   | 1.2   | 1.4   | ...   | 3.5   | 9.1   |
| Other                      | 0.9   | 1.8   | 1.8   | ...   | 1.9   | 4.2   |
| Grand total                | 11.2  | 21.3  | 32.8  | ...   | 23.7  | 44.9  |

Sources: Department of Commerce, Survey of Current Business, various issues.

*Year-end figures.

1 Excluding international and regional organizations.

United States and Western Europe were not subject to major limitations, as in Japan, nor did they enjoy special advantages, as in Canada.

The U.S. deficit position of short-term assets (excluding U.S. Government assets and liabilities and official reserve assets) vis-à-vis the rest of the world is accounted for essentially by Western Europe (Table 1). On long-term assets, the United States has a large net surplus position vis-à-vis the rest of the world, mainly Canada and Latin America, but a small deficit vis-à-vis Western Europe. Western Europe is thus a creditor of the United States on both long-term and short-term capital, even without considering official holdings of dollar balances. During the 1960s, Western Europe's net creditor position on long-term assets vis-à-vis the United States remained almost unchanged, but there was a strik-

In order to explain the structure of the international investment position of the United States, particularly vis-à-vis Europe, in the mid-1960s, Kindleberger [12] and others suggested that international flows of capital between the two areas were determined essentially by differences in the term structure of interest rates. They argued that, owing to a higher preference for liquidity in Europe, short-term interest rates were lower and long-term rates higher than in the United States. Thus, the liquidity deficit of the United States reflected not a fundamental disequilibrium but rather the fact that the United States was acting as a banker for Europe by providing the European countries with the short-term assets that they preferred in exchange for long-term capital. This view, however, was challenged by Triffin [21], who pointed to the fact that most of the large short-term European holdings in the United States consisted of official monetary reserves, and, more importantly, that total long-term European investment in the United States matched U.S. investment in Europe (at the end of 1964), although the former consisted mainly of portfolio capital and the latter mainly of direct investment.

Insofar as different liquidity preferences are a result of different attitudes toward risk, one would expect interflows between two financial markets of assets that have different risk, not necessarily different maturity. For instance, if European investors require a higher premium for bearing risks than U.S. investors, one would expect (in the absence of exchange risks) that Europeans would purchase relatively riskless U.S. financial assets, for which they may obtain a better (expected) return than for comparable assets in Europe, and U.S. investors would purchase relatively more risky European assets. Since there may well be stocks that are considered less risky than some bonds, and long-term

9 In theory, the existence of different interest rates according to maturity for debentures that are assumed to be fixed in money value and free of default risk is explained essentially by the risk of capital gains or losses caused by possible future changes in interest rates, Tobin [20]. In reality, debentures are usually not free of default risk, and, in addition to debentures, a large part of total financial assets consists of stocks. If the market required a premium for bearing risk, whatever its causes may be, all securities would be priced so that higher expected rates of return are associated with higher estimated risks, independent of the maturity, which is only one of the various components of risk, Linnekin [16].
bonds that are considered less risky than others with shorter maturity, interflows of the same type of financial asset (stocks, debentures with certain maturities) may be explained by the fact that investors adjust their portfolios according to their risk preferences. An additional reason for such interflows of portfolio capital may also be the desire to reduce overall risks through international portfolio diversification, Levy and Sarnat [15]. These considerations cannot explain, however, why most of the capital outflow from the United States consists of direct foreign investment instead of portfolio securities, and in general why capital movements toward countries with developed capital markets take the form of direct investment instead of purchases of securities.

II. Direct Investment Vis-à-Vis Portfolio Investment

For the purpose of classifying capital flows in the balance of payments, direct investment is normally distinguished from portfolio investment on the grounds that the former entails control of a local enterprise by a foreign resident, while the latter does not. From this point of view, the nature of the foreign investor is irrelevant. This is, however, of great relevance in understanding the economic factors that determine one or the other type of investment. Direct investment is made essentially by corporations. Examples of direct foreign investment by individuals are relatively rare, and normally of short duration. In practice, individuals invest abroad mainly in securities (or bank deposits); the choice open to them is confined largely to portfolio investment. On the contrary, corporations invest abroad mainly in the form of direct investment, but they also have the choice of making portfolio investment. The identity of the investor is of great relevance, inasmuch as the behavior and objectives of individuals and corporations differ substantially.

The main determinants of individuals' investment decisions are the expected rate of return and the risk of the investment. Since most investors are risk averse, securities associated with a higher risk nor-

---

4 During the last century, direct investment in the United States by European individuals was common but normally, after some time, the investor found it convenient to become a U.S. resident and the investment lost its foreign classification.

5 In addition to real estate, which, however, cannot be regarded as a direct investment.

6 Another important determinant may be liquidity, but for securities, liquidity can be regarded as one aspect of total risk (see footnote 3).
Flows of portfolio capital occur to countries where expected rates of return for securities of a given class-risk are higher. Such flows thus tend to equalize internationally the rates of return for securities of equal risk. It is pointed out by most authors that portfolio investment under perfectly competitive conditions would be a more efficient way to move capital in response to international differences in rates of return than would direct investment. One of the main reasons is that, in the absence of special advantages, subsidiaries of foreign companies are likely to operate at higher costs than are local competitors, because of the lack of institutional knowledge of the local market, difficulties in adapting to local customs, and the cost of international communications with the parent company. The main requirements for perfectly competitive conditions in this context may be summarized as follows: (a) the rate of return and risk of foreign equities effectively reflect the rate of profit and risk of foreign enterprises; (b) enterprises of one country have no special advantages that allow them to operate subsidiaries in another country more profitably than local enterprises; (c) the objective of both individuals and enterprises is the maximization of profit in competitive markets; and (d) individuals and enterprises attach the same premium to exchange risks and are equally able to cover themselves against such risks.

Under these assumptions, all international capital movements would presumably take the form of portfolio capital. The determinants of direct investment may thus be analyzed in the light of actual deviations from those assumptions. Some authors (Balassa [4], Hymer and Rowthorn [10]) have concentrated on deviations from assumption (c), stressing oligopolistic behavior and the fact that modern corporations give higher priority to growth than to profit (Section V). According to this line of argument, flows of direct investment would be determined mainly by variables other than international differences in rates of return to capital and may not be related to flows of portfolio capital.

The line of theory developed by Hymer [9] and Kindleberger [13] concentrates on deviations from assumption (b). According to their view, for a firm to undertake direct investment, it must have an advantage over firms of the foreign country. The advantage may be in terms of superior technology, management skills, knowledge of markets, possible economies of scale, etc. (Section IV). The economic justification for the direct investment would be that the firm is able to earn, through its foreign subsidiary, a higher rate of return than are competitors in the foreign country. Possible differences in average rates of profit in the two countries are regarded as irrelevant for direct investment, on the grounds that any such difference would cause flows of portfolio capital rather than direct investment. Flows of capital under one or the other form could thus occur in the same or in opposite directions.

Relaxing assumption (a) leads to the consideration of imperfections in the capital markets that make rates of return and risks associated with foreign securities—in particular, equities—substantially different from the rates of profit and risks of the respective foreign companies. Portfolio investment may occur only through the purchase of existing securities. In the extreme case where no organized market for securities exists, as in many less developed countries, capital inflows obviously may occur only in the form of direct investment. Even when securities are available, inefficiencies in capital markets, as in many European countries, may increase the risks of minority investors far above the level of the "industrial risk" inherent in the operations of the company. The additional risks borne by the portfolio investor result mainly from lack of updated information on the company's operations and narrowness in the market for securities, which cause fluctuations in the price of securities much larger than justified by fluctuations in the operating results of the company (Section III). Owing to these inefficiencies in the market for securities, portfolio investors may be deterred from buying shares of a foreign corporation, even if the expected rate of profit of that corporation is higher than that of comparable domestic firms. Shortfalls in securities markets, however, may be avoided through direct investment. Even in the absence of oligopolistic behavior or of technological advantages of domestic firms, capital outflows may thus occur in the form of direct investment into a foreign country where the average rate of profit is higher but where portfolio capital inflows are impeded by inefficiencies in the market for securities. Here, flows of direct investment would tend to equalize internationally the rates of return on capital, whereas flows of portfolio capital would follow a different pattern, dependent mainly on the development of securities markets. One may conclude that it is difficult to say, a priori, which of the two forms of capital is likely to be more closely related to international differences in rates of return to capital.

As was mentioned earlier, portfolio investment abroad may be made not only in response to higher expected returns at equal risk but also
in order to reduce the overall risk through international portfolio diversification. This economic function, however, may also be served by a domestic firm through international diversification of its investment abroad. Moreover, with regard to the shifting of short-term capital for exchange rate speculation, companies with large interests abroad are likely to do substantially better than individual investors, Aliber [2]. In this respect, direct foreign investment may reduce the incentive for portfolio investment abroad. Individual investors, both domestic and foreign, may find it more convenient to concentrate their funds on securities of large multinational corporations, which represent a package of investment in many different countries and are likely therefore to be more stable than stocks of smaller corporations operating in various national markets.

In general, even if the determinants of portfolio and direct investments are different, flows of the two forms of capital are likely to be partly substitutes. Thus, an outflow of portfolio capital, for instance, will discourage direct investment abroad (and might encourage an inflow of foreign direct investment) to the extent that: (a) it reduces the supply of risk capital to local firms and increases that supply to foreign firms; (b) it pushes up the market value of securities of foreign firms, increasing the cost of take-overs of foreign firms by domestic firms; and (c) it pushes up the exchange rate of the foreign country, increasing the cost of direct investment in that country.

III. Imperfections in Securities Markets

Imperfections in markets for securities may be an important determinant of direct investment abroad. Even in the absence of oligopolistic behavior or of technological advantages, direct investment may be attracted toward areas where average rates of profit are higher when such rates are not equalized internationally by portfolio capital flows owing to inefficiencies in securities markets. This argument seems to be relevant in explaining the expansion of U.S. direct investment abroad, particularly in Europe.

A number of factors may contribute to making the holding of portfolio shares in European companies unattractive, even in companies

* Numerous U.S. firms derive 30 per cent to 40 per cent, and in certain cases up to 60 per cent, of their total earnings from foreign investment.

with a high expected rate of return. A major factor is the lack of information about the company’s affairs. In most European countries, with a few exceptions, such as the United Kingdom, public auditing of corporations is not as well developed as in the United States, and shareholders receive much less, if any, information about the current position of their company. For portfolio investors, this increases the risk of possible deviations from the expected rate of return. The direct investor, however, being in control of the company, has immediate and direct access to all information; his risk is thus limited to the “industrial” risk inherent in the operations of the company. Another factor is that the market for stocks (and the ratio of stocks that are normally traded in the market to total stocks of single companies) is much smaller in most European countries than in the United States, and this may cause much larger fluctuations in the market price of stocks, both for speculative reasons and in relation to fluctuations in the rate of return of the company, than in the United States. Since normally a portfolio investor is interested mainly in the day-to-day value of his stock, while a control (direct) investor is interested mainly in the medium-term and long-term profitability of the company, widespread fluctuations in the value of a stock have a higher negative weight for the former than for the latter. Table 2 shows, for instance, that in the period 1951–67 the rate of return on stocks (measured as the average annual percentage change in the dollar value of the index of common stocks) of the United States was not too different from that of other major industrial countries, while the standard deviation in annual rates of return was definitely much lower in the United States than in all other countries (except Belgium and the United Kingdom).

Of course, U.S. purchases of European portfolio securities have also been impeded by institutional factors, particularly the Interest Equalization Tax introduced in 1963. But this could not explain the opposite massive flow of European portfolio purchases of U.S. securities. This
flow suggests that the greater efficiency of the U.S. stock market acts as a formidable incentive to portfolio investment in U.S. securities.\textsuperscript{12}

Disadvantages of inefficient capital markets may be avoided, however, through direct investment. A U.S. corporation may thus have an incentive to take over a more profitable European company (or to establish a subsidiary in Europe), even if portfolio investors prefer to buy securities issued by the former rather than by the latter.

<table>
<thead>
<tr>
<th>Table 2. Major Industrial Countries: Mean Rates of Return and Standard Deviations of Common Stocks, 1951–67</th>
</tr>
</thead>
<tbody>
<tr>
<td>(In per cent)</td>
</tr>
<tr>
<td>Rate of Return 1</td>
</tr>
<tr>
<td>Japan</td>
</tr>
<tr>
<td>Germany</td>
</tr>
<tr>
<td>United States</td>
</tr>
<tr>
<td>Italy</td>
</tr>
<tr>
<td>France</td>
</tr>
<tr>
<td>Netherlands</td>
</tr>
<tr>
<td>Canada</td>
</tr>
<tr>
<td>United Kingdom</td>
</tr>
<tr>
<td>Belgium</td>
</tr>
</tbody>
</table>


The argument may be illustrated graphically. Lines $UU$ and $EP$ (in Chart 1) represent the investment opportunity loci for portfolio investors in U.S. and European stocks, respectively. The U.S. stock market offers lower risks at equal rates of returns (line $UU$ is higher than line $EP$ for most of the relevant range of choices), and this explains the large net outflow of portfolio investment from Europe to the United States. The U.S. financial market is assumed to be so "efficient" that there is no additional advantage in control stocks over portfolio stocks; hence, there is only one market and one opportunity investment locus ($UU$).\textsuperscript{12} In Europe, on the contrary, institutional factors result in substantial additional advantages of control over portfolio stocks, mainly in the form of a lower risk that the former bear for the same expected rate of return. There are thus two separate markets: $EP$ is the investment opportunity locus for portfolio investors, and $EC$ is the opportunity locus for control investors. Since European portfolio investors cannot acquire control shares in Europe (switching from locus $EP$ to locus $EC$) because the size of investment that is required constitutes an insurmountable threshold, they opt for portfolio investment in the United States. U.S. companies, however, are large enough to become control investors in Europe and through direct investment can take advantage of the higher rates of return in European companies.\textsuperscript{13}

\textsuperscript{11} It is not suggested here that one should always find a correlation between inflows of direct investment and inefficiency of stock markets, since clearly many other factors influence actual direct investment. In this respect, the value of U.S. direct investment relative to gross national product (GNP) is much larger in the United Kingdom than in other European countries, although the United Kingdom has the most developed stock market in Europe. However, figures for 1957–64 (Aliber [1]) shows that the outflow of direct investment from the United Kingdom largely exceeded the inflow, and that the United Kingdom had the largest net outflow of direct investment after the United States. Estimates for 1964 (Behrman [5]) indicate that the United Kingdom was the only European country except the Netherlands (whose stock market is more developed than those of other member countries in the European Economic Community) with a large net surplus position for direct investment. Therefore, if net instead of gross flows are considered, the United Kingdom's example does not contradict the view that inflows of direct investment may be stimulated by the inefficiency of domestic stock markets.

\textsuperscript{12} Political and fiscal considerations also certainly play a role, but it is difficult to see how large portfolio capital inflows could occur in the absence of an efficient market, such as the New York Stock Exchange.

\textsuperscript{13} Mariner (U.S.) vehicle, April 1963, in 85 per cent of the total outstanding corporate stocks.
markets may thus explain both the flow of U.S. direct investment in Europe and the opposite flow of portfolio capital.

European companies compensate in part by incurring a higher debt/capital ratio than U.S. companies, but the possibility of raising funds through borrowing (essentially from banks) has obvious limits. Insofar as the inefficiency of capital markets impedes the raising of risk capital from portfolio investors, the total supply of risk capital for new industrial undertakings is limited by the concentration of wealth within a few hands, and thus becomes a matter of national distribution of wealth. This explains why in many instances foreign (U.S.) investment is the only practical alternative to government investment, even for countries that export substantial amounts of capital.

To the extent that direct investment is determined by the causes just mentioned, one would expect it to have the following characteristics: (a) The company acquired would, in general, be a "healthy" one, with a higher expected rate of return, at equal industrial risk, than comparable U.S. companies. (b) The management of the foreign subsidiary would remain substantially autonomous, and the participation of the foreign company could be limited to a relatively low share of total capital. (c) The foreign subsidiary would not necessarily have to operate in the same sector as the parent company.

IV. Advantages of Superior "Knowledge" and Economies of Scale

According to Kindleberger [13] and others, direct foreign investment is determined essentially by advantages that allow a firm to operate a subsidiary abroad more profitably than local competitors. These advantages may be classified in two broad categories: superior "knowledge" and economies of scale. The term "knowledge" includes production technologies, managerial skills, industrial organization, knowledge of product, and factor markets. A common aspect of all advantages of superior knowledge is their character of public goods, that is, the marginal cost of exploiting them abroad through direct investment is practically nil for the firm that owns the knowledge, or at least much lower than the cost that the local firm would incur in developing comparable knowledge. In fact, many authors (e.g., Johnson [11]) regard the transference of knowledge as the core of the problem of direct investment.15

Superior knowledge as such, however, is not enough to justify direct investment; the latter must also provide to the firm the highest return among all alternative ways of exploiting the superior knowledge in the foreign market. These alternatives are, essentially, (a) exporting products that embody the knowledge; (b) selling the knowledge to local producers in the foreign market; and (c) producing abroad (through direct investment) products that embody the knowledge.

As mentioned earlier, operating a subsidiary in a foreign market involves additional costs compared with those of local producers, owing to lack of institutional knowledge of the local market, costs of communications, etc. Therefore, when transportation costs, customs duties, and other factors make local production less costly than imports (see below), the foreign firm would be better off selling the knowledge to local producers, who operate more efficiently in their own market. Direct investment may thus be justified only when market imperfections do not allow the foreign firm to obtain the full rent value of its superior knowledge from local producers.16

Thus, the choice between direct investment and sale of knowledge depends on the additional costs of doing business abroad, on the one hand, and the cost and feasibility of selling the knowledge, on the other hand. The United States, in addition to direct investment abroad, sold knowledge on a large scale during the 1960s in the form of licenses, royalties, management services, etc.17 Many types of knowledge, however, cannot be sold, mainly because they cannot be embodied in a license, as is true for managerial expertise, industrial organization, knowledge of markets, etc. Even when the knowledge can be embodied in a license, the local producer may not be willing to pay its full rent

15 Also, oligopoly, with product differentiation, can be treated as superior knowledge, since the marginal cost of exploiting the differentiated product in a foreign market is practically nil; the advantage here consists in the knowledge of the product, possibly protected by a brand name.

16 It is implicitly assumed here that the objective of the firm is to maximize profits. In oligopolistic markets, or when the firm seeks to maximize growth rather than profits (see Section V), direct investment may be undertaken even when its return is lower than the possible income from the sale of the license.

17 Between 1964 and 1969, U.S. gross receipts from sale of knowledge (royalties, license fees, and rentals) to foreign subsidiaries of U.S. companies and to nonaffiliated foreign firms, respectively, amounted to US$2.6 billion and US$2.4 billion (the latter including also management and service fees), U.S. Department of Commerce [33], "Policy Aspects of Foreign Investment by U.S. Multinational Corporations," p. 37.
value because of uncertainties about its utilization, or because the superiority of the foreign firm consists not only in accumulated knowledge but also in a continuous production of new knowledge that the local producer can utilize fully only through a permanent or institutional agreement with the foreign firm.

The additional costs of direct investment are largely in the nature of fixed costs. This helps to explain why it is often observed that a foreign firm first exports a new product into the local market, then sells the license to a local producer, and finally establishes its own subsidiary for local production. The reasons for this sequence have been illustrated by Aliber [1]. As the local market grows to a size that makes local production cheaper than imports (given the possible economies of scale, the relative levels of production costs, the costs of transportation, and existing customs duties), a local producer may offer to pay a higher rent value for the license than the foreign firm could obtain by exporting, while direct investment still remains too costly in relation to the volume of sales in the local market. However, a further growth of the market reduces the relative importance of fixed costs of direct investment, and unless the local producer is willing to pay close to the full rent value of the license, which normally will not be so, at some point the foreign firm is able to obtain a higher return by producing directly through a subsidiary rather than by selling the license.

The fixed nature of most additional costs of direct investment may also explain why it is undertaken mainly by large companies, which are able to mobilize a volume of resources greater than the minimum below which direct investment is unprofitable. Another reason why small companies normally opt for sale of the license rather than for direct investment is that their superiority usually consists of a one-shot innovation that would leave them with an unprofitable foreign subsidiary after the innovation becomes obsolete, whereas large companies may rely upon a continuous production of new knowledge, Caves [7].

Attributing the cause of U.S. direct investment abroad mainly to the advantages of superior knowledge leaves open the problem of why U.S. companies happen to be in such a privileged position. Institutional factors, such as the volume of research financed by the U.S. Government or the larger size of U.S. companies, undoubtedly contribute to this predominance. However, a theoretically more satisfactory explanation is that suggested by the theory of the product cycle (or "industry life cycle") developed mainly by Vernon [24].

Whereas the classical theory of international trade assumes that technology is constant, the theory of the product cycle regards technological innovations as the main determinants of the structure and development of world trade and the distribution of production among countries. The theory may be extended along the following lines to explain the predominance of U.S. direct investment abroad. In most countries, a rise in real income causes similar changes in consumption patterns, with the progressive diffusion of new products. The United States, having a large market and the highest per capita income, is normally the place where new products are introduced first. In the initial stage when the product is not yet standardized, development of a new product requires close contact with the market, and production is, therefore, concentrated in the United States even if production costs in other countries are lower. When the product is established in the U.S. market, the U.S. company starts exporting it to other countries where demand initially is inelastic. Later, however, as demand in foreign countries grows, and production may be easily located outside the United States because techniques are standardized, the U.S. company is induced to invest abroad both to exploit lower local productions costs and to avoid the danger of losing the local market to local producers. According to this interpretation, U.S. direct investment abroad would be essentially of a "defensive" rather than an "offensive" nature; that is, U.S. companies would invest abroad to avoid losing their markets (not only abroad but in the United States as well) rather than to gain new markets.18

However, theories based on the product cycle per se cannot explain direct investment because they do not explain why, when producing abroad becomes more advantageous, U.S. firms establish subsidiaries...

18 A recent study by the Harvard Business School, "U.S. Multinational Enterprises and the U.S. Economy," in [23], found that a number of case studies confirmed this general description of an industry life cycle. It found also that between 1950 and 1970, for a large sample of industries, the share of world production made in the United States declined sharply, while that of U.S. companies' total production, made in both the United States and abroad, declined much less markedly. The thesis that U.S. direct investment abroad is of a defensive nature has also been advanced by Hymer and Rowthorn [10], although their argument is that U.S. firms invest abroad in order to maintain their world market share at a time when GNP in the United States is growing less rapidly than in other developed countries (see Section V). Various empirical studies have found a positive correlation between inflows of direct investment and the rate of growth of GNP, Spittel [19].
abroad instead of selling the license to more efficient local producers. To explain that, one must consider again the imperfections in the license market or assume that there is oligopolistic behavior by the firms (see below).

Besides superior knowledge, a determinant of direct investment may be the opportunity of achieving economies of scale. Economies of scale may be internal or external to the firm; the former normally lead to horizontal investment, and the latter to vertical investment. Foreign investment in vertically related stages of production is common mainly in industries producing and processing minerals and other raw materials. The main advantage of direct investment here consists in reducing the costs and uncertainties that exist when subsequent stages of production are handled by different producers by coordinating decisions at various stages within one firm.

An increase in production through horizontal investment may permit a reduction in the unit cost of certain general services, such as financing, marketing, or technological research, that have the nature of fixed costs, but this case is qualitatively similar to that of superior knowledge. It is different when internal economies of scale are achieved through an international integration of production "by components"; that is, each affiliate produces those parts of the final product for which local production costs are lower, and subsequently the final product is assembled for marketing in various countries. In this way, the firm may benefit from local advantages in production costs while achieving maximum economies of scale in production of single components. Such an international integration of production would be much more difficult through trade among different producers because of the need for close coordination between different phases of production and new plant investment.

In a different case in which each plant in the various countries produces the same product in its entirety, horizontal foreign investment may have the advantage of allowing the firm to even out the effects of business cycles in various markets by shifting the direction of sales of its subsidiaries more efficiently than independent producers could do, and by reducing risks of overproduction by planning new investment on a world-wide basis.

V. Oligopoly and Maximization of Growth

In oligopolistic markets, the main determinant of direct investment may be simply that of increasing profit rates by reducing competition, irrespective of whether the investor is a more efficient producer than the firm that is taken over. In fact, the main purpose of vertical integration is often to raise barriers to the entrance of new competitors, and horizontal investment may be undertaken simply to prevent small local producers from developing into future competitors. For U.S. investment in Europe, however, particularly in manufacturing, this does not seem to be a relevant factor. Indeed, the entry of U.S. companies often increased competition by breaking the established oligopolistic equilibrium in national markets.

A somewhat different approach is that derived from theories that stress growth rather than profit as the main objective of the firm.\(^\text{20}\) In this framework, Bela Balassa [4] argued that when a mature oligopolistic structure has been established in the domestic market, the firm may be induced to invest abroad because efforts at increasing its share in the domestic market would meet retaliation from other oligopolists. In spite of all the additional costs of foreign investment, expansion in foreign markets is thus less costly than in the domestic market.

In addition to undertaking foreign investment, another alternative open to the firm in order to expand rapidly without costly fights with its competitors is that of taking over other local firms, both within and outside the same sector. The development of conglomerates and the large wave of mergers that has taken place in the United States over the past two decades seem to support the view that foreign investment was part of a more general drive for expansion of U.S. firms. Antitrust legislation that hindered take-overs of competitors within the same sector was certainly an additional stimulus to foreign investment. A study of a number of firms by Bain [3] shows, however, that the degree of industrial concentration is, in general, much higher in Europe than in the United States. This reduces the validity of the argument, suggesting that other factors besides the degree of oligopolistic concentration must account for the exceptional expansion of U.S. investment abroad.

Hymer and Rowlthorn [10] found that the share of total world production accounted for by U.S. corporations has declined, despite their

\(^{20}\) Profit is regarded as a constraint rather than as a target in itself, in the sense that the achievement of a minimum rate of profit is necessary to finance further growth and to retain control of the firm.
expanded foreign investment, and concluded that U.S. firms invested abroad not to increase their share of world markets but to avoid the decline implicit in the slower rate of growth of the U.S. economy compared with rates in the Common Market and Japan. This argument could be extended to explain the rapid development of British direct investment abroad at a time when the British economy was stagnating and the market for exports in the Commonwealth was shrinking. A generalization of the argument would be that companies of all countries strive to achieve a similar rate of growth and that they invest abroad whenever growth in their own markets lags behind that in the rest of the world. Direct foreign investment would be a sort of antidote to domestic stagnation, which appears a somewhat paradoxical view.

VI. The Impact of Tariffs and Exchange Rates

So far, the assumption has been that exchange rates are constant and fixed. Actually, the theories of direct foreign investment based on oligopolistic behavior or technological advantages "lack elements of 'foreignness' in the sense that the explanatory variables do not include any of the factors that distinguish national economies, including participation in different customs areas, currency areas, and tax jurisdictions." 21 In other words, these theories can be applied to explain direct foreign investment as well as the growth of the firm and the flow of investment among regions of the same country. This does not necessarily reduce their relevance for explaining the flows of direct investment in the post-war period. However, in order to catch those aspects that are characteristic of direct foreign investment, it is necessary to consider the impact of those factors that separate the economies of different countries, mainly customs duties and exchange rates and exchange risks.

Customs Duties

Customs duties are often regarded as a major cause of direct investment. Other things being equal, the higher the rate of duty, the greater is the incentive for a foreign company to produce inside the customs area rather than to export into it. The impact of the duty must be seen, however, in relation to the size of the market and particularly in relation to the possible diseconomies of scale connected with the decentralization of production in a foreign affiliate. 22 Thus, an increase in the rate of duty may not be enough to stimulate an import-substituting foreign investment if the market is small and diseconomies of scale more than offset the cost of the duty, while enlarging the market may attract foreign direct investment even if the rate of duty is not changed. (The establishment of the Common Market, for example, was certainly an incentive to U.S. direct investment in Europe.) The customs duty by itself, however, can be only a complement to theories explaining the determinants of direct investment. Customs duties can actually be treated as transportation costs, and thus can be considered part of the theory of industrial location.

Exchange Rates

The problem of whether the overvaluation or undervaluation of a currency introduces an incentive (even if it is not the main determinant) for direct investment abroad, or for foreign investment in the country, has been largely neglected in the literature, with the exception of an interesting paper by Aliber [1], which focuses on the exchange risk as a determinant of direct investment.

It is convenient to consider separately the impact of the level of the rate and that of the risks of changes in the rate. A currency may be defined as undervalued when, at the current rate of exchange, production costs for tradable goods in the country are, on average, lower than in other countries. 23 This type of undervaluation could conceivably persist over time if, for instance, transfer payments from the country offset the trade surplus. The undervaluation of the currency, as just defined, represents an incentive to the location of production of internationally traded commodities in the country. This alone cannot explain direct foreign investment, since local producers are presumably more efficient than affiliates of foreign firms. However, if both certain local firms and certain foreign firms have some technological advantage over their competitors, the undervaluation of the currency may play an important role, in that it stimulates foreign firms to exploit their advan-

21 Quotation from Aliber [1], p. 20.

22 William Lever, the founder of the Lever Brothers soap empire, is quoted as saying in 1902, "The question of erecting works in another country is dependent upon the tariff or duty ... When the duty exceeds the cost of separate managers and separate plants, then it will be an economy to erect works in the country." Tugendhat [22], p. 14.

23 Lower labor costs abroad are commonly regarded as a major determinant of U.S. direct investment abroad. This view is theoretically invalid unless qualified as in the text that follows.
tage through direct investment in the country where they may benefit from lower production costs, whereas local firms have no incentive to produce outside their own country. In conclusion, if the exchange rate does not equalize production costs among different countries, there is a potential incentive for direct investment to flow to a country with an undervalued currency (and a disincentive for flows in the opposite direction), although, as for customs duties, the undervaluation alone cannot explain the direct investment.

A more intriguing case is that of the impact of exchange risks on direct foreign investment. Aliber [1] argues that when there is a risk of change in the exchange rate the firms of the strong-currency area are at an advantage and are stimulated to invest in the weak-currency area. In summary, his argument goes as follows. Direct foreign investment reflects the fact that the firm in the source-country capitalizes the same income stream of expected earnings (that of the host-country firm) at a higher rate than does the host-country firm. When a change in the exchange rate is expected, capitalization rates on equities, as well as on debt issues, are lower (that is, interest and profit rates are higher) in the weak-currency area. Under perfect market conditions, there would be no incentive for direct foreign investment, because the exchange risk would offset the lower capitalization rate applied to the income stream of the weak-currency firm. However, Aliber argues that the market for equities is biased, in that it does not attach a currency premium to the foreign income of the source-country firm. The latter may thus issue equities in its market (at a higher capitalization rate) and buy the host-country firm, whose income stream is capitalized by the market at a lower rate owing to the exchange risk.

There are, however, three main objections to Aliber's argument: (1) It is not clear why the existence of a currency premium should cause the interest and profit differential to exceed the expected change in the exchange rate. (2) There is no convincing reason why the market should capitalize the additional income to the source-country firm, deriving from the acquisition of the host-country firm, without discounting it for the exchange risk. (3) If the market does not attach a currency premium to the foreign income of the source-country firm, it also should not attach a currency premium to the foreign liabilities of the host-country firm. The latter could then increase its income stream by borrowing at lower rates in the strong-currency area, and thus offset any advantage that the foreign firm might have.

Empirically, Aliber's conclusion that direct investment tends to flow from strong-currency areas to weak-currency areas seems to be contradicted by the continuing inflow of U.S. foreign investment to Europe in recent years, when most European currencies were considered "stronger" than the U.S. dollar. The rapid increase of U.S. direct investment in the United Kingdom in the late 1950s and early 1960s, when the dollar was stronger than the pound, cannot be taken to support Aliber's argument, since, in that same period, the outflow of U.K. direct foreign investment was substantially larger than the inflow. In fact, net direct foreign investment of the United Kingdom increased rapidly at a time when sterling was weak. In contrast, in the same period, countries with strong currencies, such as Germany, had a large net inflow of direct investment, Diamond [8].

In fact, it is possible to argue, contrary to Aliber, that firms in weak-currency areas have an advantage in investing in strong-currency areas if the interest rate differential underestimates the exchange risk, and, in particular, if borrowing in the weak-currency areas by firms in the strong-currency areas is restricted. Similar conditions seem to have prevailed since the first half of the 1960s when the dollar started to be regarded as a weak currency, and may indeed have contributed to the inflow of U.S. direct investment to Europe.

Even under perfect market conditions, the difference between interest rates in different currency areas may not exactly reflect the expected change in the exchange rate owing to the "currency premium" that the market demands for bearing the exchange risk. This does not necessarily mean, however, that the difference in interest rates will exceed the expected change in the exchange rate as suggested by Aliber. Indeed, one may normally expect the opposite to be true, especially if the authorities of the weak-currency area are willing to accept some loss of reserves in order to avoid excessively high domestic interest rates.25

Assume that there are only two types of securities: riskless debentures and equities of equal risk; expected returns on the two are called, respectively, interest and profit. Assume also that the same premium is

24 In this section, consideration is given only to direct investment in industries producing mainly for the local market. As pointed out by Rhomberg [17], changes in the exchange rate have opposite effects on the rate of return of export-oriented industries, that is, industries whose payments in local currency exceeds revenues in local currencies.

25 This has clearly been true for the United States, where losses of reserves, owing to the special position of the dollar, were not a serious constraint on domestic policies.
demanded by each market for bearing the risk associated with equities, so that the differential between rates of profit and rates of interest in the two areas is the same. (This difference is taken to be 4 per cent in the tabulation that follows.) Assume also that firms in both areas have the same expectations regarding future changes in the exchange rate (for instance, in the tabulation, both expect a devaluation averaging 3 per cent a year over the considered time horizon).

Average Annual Rates Over a Given Time Horizon
(In per cent)

<table>
<thead>
<tr>
<th></th>
<th>Hypothesis A</th>
<th>Hypothesis B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Interest rate</td>
<td>Profit rate</td>
</tr>
<tr>
<td>Strong-currency area</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Weak-currency area</td>
<td>3</td>
<td>7</td>
</tr>
</tbody>
</table>

If the securities market also shared the same expectations and these were fully reflected in the interest and profit differential between the two areas, there would be no incentive to move funds in either direction. This is shown by Hypothesis A in the tabulation. However, the interest and profit differential between the two areas may underestimate the change in exchange rate as expected by the firms, for instance, because the authorities of both areas intervene to keep domestic interest rates in line with their domestic policy targets, and equalizing portfolio movements of capital do not occur because the market requires a currency premium for bearing the exchange risk or because portfolio movements are subject to restrictions. Instead, firms move their own funds (raised through sales of either debentures or equities) in response to expected gains from changes in exchange rates more readily than do portfolio investors. Under these conditions (Hypothesis B in the tabulation), firms in the weak-currency area are better off investing in the strong-currency area. Moreover, in doing so, they have an advantage over firms of the strong-currency area, because the expected rate of profit in national currency for the same investment in the strong-currency area is higher for the firms of the weak-currency area than for those of the strong-currency area by the amount of the expected average rate of devaluation. Firms of the weak-currency area may thus issue equities or borrow in their area and acquire equities in firms of the strong-currency area because they may apply a lower capitalization rate than the latter. Firms of the strong-currency area could reduce their disadvantage by issuing debentures (assuming that they cannot issue equities) denominated in the weak currency. If, however, flows of debt capital (or, in general, portfolio capital) from the weak-currency to the strong-currency area are restricted or are limited by the existence of a currency premium, while flows of firms' own capital are not, then firms of the weak-currency area have a substantial advantage over firms of the strong-currency area.

U.S. firms investing in Europe during the past decade appear to have enjoyed some of the advantages just mentioned. During the 1960s, the cost of borrowing in the United States was normally lower than in Europe, despite the fact that many European currencies were often considered stronger than the U.S. dollar. In the early 1960s, some European companies took advantage of lower U.S. interest rates by issuing securities in New York, but this possibility was effectively ended by the introduction of the Interest Equalization Tax in July 1963. In contrast, in the period 1960–64 (as shown in Table 3), U.S. direct investment abroad was financed essentially through capital transfers from the United States (60 per cent) and reinvested earnings (40 per cent). U.S. corporations did not borrow abroad until 1965, when they were pressured to do so by the Voluntary Direct Investment Program introduced at the beginning of that year.

U.S. companies became active in the Euro-bond market in 1965, but their participation reached large proportions only after the institution of mandatory controls on foreign direct investment in January 1968. Despite the Voluntary Program, in the period 1965–67, capital transfers from the United States for financing direct investment abroad

---

26 This is explained largely by the reserve role of the dollar.
27 In 1964 the outflow of U.S. funds for direct investment in Europe jumped to US$1.3 billion, compared with an average of US$0.8 billion in the previous three years; one may suspect that this increase was stimulated by the restrictions on U.S. purchases of European securities.
28 Throughout the text, reference is made only to borrowing by the parent company to increase its equity in the subsidiary (which is what is meant by "direct investment"). While borrowing by the subsidiaries is not considered here, it rose from 38.2 per cent of the total assets of U.S. subsidiaries abroad in 1966 to 41.5 per cent in 1969, Berlin [6].
averaged US$2.9 billion, while new issues of U.S. securities abroad averaged only US$0.4 billion. This fact confirms that it was substantially cheaper to raise capital in the United States than abroad. In 1968 mandatory controls were introduced on the amount that each U.S. firm could invest abroad in each year, including proceeds from reinvested earnings.\footnote{These limits varied according to the geographical areas and were much stricter for developed countries in Western Europe than for other areas.} The Program was not intended to discourage direct investment per se but rather to shift the source of financing from the United States to foreign markets. Indeed, under the Program, proceeds of long-term foreign borrowing can be used to finance direct foreign investment in addition to the firm’s allowable amount. The impact of the mandatory controls was felt immediately: net transfers from the United States declined from US$2.9 billion in 1967 to US$0.9 billion in 1968, while new issues of U.S. corporate securities abroad rose from US$0.4 billion in 1967 to US$2.1 billion in 1968 (Table 3). In that year, non-U.S. borrowers, in order to maintain a share in the market, had to sharply increase issues denominated in deutsche mark, which amounted to paying a higher rate of interest in view of the risk of revaluation of the deutsche mark, Shapiro and Deastlow [18]. Subsequently, however, net transfers from the United States increased again to record levels (Table 3).

In conclusion, it can be argued that the expectation of exchange gains owing to the undervaluation of European currencies contributed to attract U.S. direct investment in Europe, while the special role of their national currency, in addition to other factors, allowed U.S. corporations to secure financing on cheaper terms (taking into account exchange risks) than their European competitors.

REFERENCES