The relationship between foreign direct investment and international trade
Substitution or complementarity?
A survey

Rosa Forte

CETE – Centro de Estudos de Economia Industrial, do Trabalho e da Empresa
THE RELATIONSHIP BETWEEN FOREIGN DIRECT INVESTMENT AND INTERNATIONAL TRADE

SUBSTITUTION OR COMPLEMENTARITY? A SURVEY

Rosa Forte*
CETE**, Faculdade de Economia – Universidade do Porto
March 2004

ABSTRACT

Theoretically, the explanatory approaches of foreign direct investment (FDI), as for example, the internalization theory and the eclectic paradigm, and general equilibrium trade models that incorporate horizontal multinational firms (MNEs), sustain the existence of a substitution relationship between FDI and international trade. Models of vertical FDI and considerations concerning demand, for their part, support a complementarity relationship. Empirically, however, it is difficult to find substitution between the two variables. This work presents a review of the existing theoretical and empirical literature, highlighting the reasons that underlie the apparent incongruity between theory and empirical works, and drawing attention to gaps that should be corrected in future works.

JEL Classification: F12; F14; F23.
Keywords: Foreign direct investment, multinational firms, international trade.

* Address: Faculdade de Economia do Porto, Rua Dr. Roberto Frias, 4200-464 Porto, Portugal. E-mail: rforte@fep.up.pt. Financial support from Fundo Social Europeu (PRODEP) is gratefully acknowledged. The author is grateful to Prof. Ana Paula Africano and Prof. Francisco Castro for helpful comments and suggestions. Thanks also to two anonymous referees of the 29th EIBA (European International Business academy) Conference.

** CETE - Research Center on Industrial, Labour and Managerial Economics - supported by Fundação para a Ciência e a Tecnologia, Programa de Financiamento Plurianual through the Programa Operacional Ciência, Tecnologia e Inovação (POCTI) of the Quadro Comunitário de Apoio III, which is financed by FEDER and Portuguese funds.
1. Introduction

Traditional models of international trade, such as the neo-classic model of Heckscher-Ohlin (HO), assumed factor immobility among countries, a very restrictive hypothesis that is to a certain extent dissociated from the existing economic reality, which is characterized by increasing international factor mobility, mainly under the form of FDI.1

The growing importance of FDI is reflected in the values of international production, which has expanded strongly in the last two decades and is presently of considerable importance in the world economy. International production consists of the production located in a country but controlled by a MNE with headquarter in another country, and is mostly financed through FDI.2 Thus, its expansion is demonstrated by the evolution of two variables, the Gross Product of the foreign subsidiaries and the Sales of foreign subsidiaries throughout the world, which grew more rapidly than the Gross Domestic Product (GDP) and Exports, respectively. According to UNCTAD (2000), in 1999, the Gross Product of the foreign subsidiaries represented about 10% of the world GDP, while in 1982, it represented only 5%. In 1999, Sales of foreign subsidiaries were about twice as high as Global Exports, while in 1982, they were practically the same, which means that FDI grew more rapidly than trade. This information corroborates one of the stylized facts referred in Markusen (2000): although slackening at the beginning of the nineties, the FDI flows and stocks as well as the Sales of foreign subsidiaries continued to grow significantly faster than GDP and Exports.

FDI can be of two types, horizontal and vertical. Horizontal FDI consist of the production of the same goods and services in different locations while vertical FDI consist of the geographical fragmentation of the production process by stages. In fact, horizontal FDI seems to be the most prevalent type of FDI since a large part of FDI is a two-way investment between similar developed countries (countries with high income, similar per capita incomes and similar relative factor endowments, and with relatively low trade barriers), although with some moderation in the nineties, a period when

---

1 FDI involves the transfer to a foreign country of a group of assets such as financial capital, technology, know-how, management techniques, etc., where the investor controls the use of the transferred resources.
2 The main economic agents that carry out FDI are MNEs. Thus, the terms "multinational firm" and "foreign direct investment" are generally used indistinctly, which is also the case in the present work.
inward FDI directed at developing countries grew substantially. This aspect represents another stylized fact referred in Markusen (op. cit.). In accordance with Markusen (1997), these stylized facts suggest that the main reasons to engage in FDI do not reside in factor remuneration differences nor in aspects related with trade barriers. FDI is motivated by the benefits of obtaining control over production and foreign markets.

The growing globalization of production referred above drew greater attention to the matter of the relationship between FDI and international trade. This relationship is relatively complex since FDI supplies one of the means through which a MNE assures control over international production and, in order to assure that control, the transfer of capital resources is just one part of a wider package. Hence, FDI can substitute or create trade. Theoretically, there are reasons that suggest both substitution and complementarity effects but, empirically, the results almost always point to a positive relationship. The precise nature of this relationship is, consequently, a controversial subject, although the literature is unanimous in recognizing its importance.

Theoretically, the substitution relationship is sustained by FDI explanatory approaches, and by general equilibrium trade models that incorporate horizontal MNEs. The complementarity relationship, for its part, is supported by models that admit vertical MNEs and by considerations concerning demand. Trade models that admit both types of MNEs (vertical and horizontal), usually designated knowledge-capital models, support the two relationship types. Empirically, however, most of the existing works point to a complementarity relationship, in particular between FDI and exports (most studies do not include the imports side). According to Head and Ries (2001), the difficulty in finding, empirically, a substitution relationship can be due to the existence of potential sources of spurious positive relationships between the two variables, such as endogeneity (use of endogenous variables, such as the foreign subsidiaries’ sales/production, as FDI indicator) and aggregation bias (bias that results from the use of aggregate data).

---

3 According to UNCTAD (op. cit.), in 1980 the developing countries presented a ratio of the inward FDI stock relatively to GDP equal to 5.1%, while in 1998 the same ratio was 20%. For the developed countries this ratio was 4.7% in 1980 and 12.1% in 1998.

4 In a quantitative sense, the international trade of goods and FDI are complementary/substitutes when an increase in trade volume is going to increase/decrease the level of FDI or vice versa.

5 This occurs in studies not only at the aggregate level of countries and industry but also at a more disaggregated level, such as firm level studies.
Blonigen (2001), in a study at the product level, shows that the use of aggregate data can have an important role. Thus, Markusen (2000) mentions that, although FDI and trade seem complementary at a superficial level, recent empirical works suggest that they are substitutes at a fine disaggregation level.

This work intends to present a survey of the literature on the problem referred above, and is organized as follows. Section 2 will focus on several theoretical contributions, which seek to explain the relationship between FDI and international trade. Section 3 will present a synthesis of the main existing empirical works. Finally, section 4 will present some conclusions, namely in terms of the reasons that justify the discrepancy between theoretical and empirical analyses, as well as indicate paths of research in future works.

2. The relationship between FDI and international trade: theoretical considerations

2.1. FDI explanatory approaches

These approaches consider FDI and exports as alternative manners of supplying a foreign market and, consequently, one substitutes the other. From among the main FDI approaches, we can point out the internalization theory and the eclectic paradigm.

According to the literature, as in Buckley and Casson (1976), the origin of the internalization theory goes back to Coase and his theory of the firm and to later contributions by Williamson. In the context of this theory, firms and markets are considered as alternative forms of organizing production since the intra-firm and market mechanisms exhibit, potentially, different efficiency levels in the execution of different transaction types. The firm’s role is fundamental whenever the costs of using the market mechanism (transaction costs) were larger than the organization costs of the same activities inside the firm. In these conditions the firm will internalize those activities.

The systematized application of the internalization concept to the MNE begun with Buckley and Casson (op. cit.). These authors suggested that a MNE will internalize its activities in a foreign country through FDI if the internalization cost (internal organization costs, such as communication costs, administrative costs, etc.) is inferior to the cost associated with export or to other forms of entry. Thus, the internalization
theory considers that FDI substitutes exports when enough costs exist associated with external transactions.

The eclectic paradigm, developed by Dunning (1980) considers that the firm will prefer FDI rather than other modes of entry if three types of advantages exist: ownership advantages of a firm, location advantages of a market, and internalization advantages of integrating the transactions inside the firm. The ownership advantages refer to the specific assets and qualifications of the firm: to compete with foreign firms in their own markets, MNEs should possess superior assets and qualifications that could have sufficiently high remunerations to compensate the high costs of serving these markets. The location advantages reflect the attractiveness of a specific country, in terms of its market potential (size and growth) and investment risk. Measures of location advantages include similitude in culture, in market infrastructures and the availability of lower production costs. Finally, the internalization advantages are concerned with the costs of choosing a hierarchical way of operation (FDI) instead of an external way. Whenever the three types of advantages are gathered, firms will engage in FDI. If the location advantage does not exist but the firm possesses the others, it will opt to export. The greater the ownership advantages the firm possesses, the greater the incentive to internalize; the greater the attractiveness of an external country relative to the domestic country, the greater the probability of the firm to engage in foreign production. Also, in this case, FDI and export are seen as two alternative entry modes.

2.2. Trade models that incorporate MNEs

Traditionally, the FDI theories and the international trade theories have been developed separately. For several years, the international trade theory was dominated by the HO model, that considered that the basis for trade resided in different relative factor endowments among countries. However, in the 1980s, this model began to be questioned due to the fact that it does not explain the great volumes of trade of similar products among countries with similar endowments. Thus, the industrial organization approach to international trade ("new trade theory") began to incorporate models based on increasing returns, imperfect competition and product differentiation, generating an intra-industry and inter-industry pattern of trade, depending on the differences in the relative factor endowments. This "new trade theory" modeled the firms as national firms (NEs), that is, firms that have a single productive infrastructure and export. However,
industries with strong scale economies and oligopolistic market structures are usually dominated by MNEs that choose endogenously the number and location of their production facilities, making vertical and horizontal investment decisions in foreign markets. Thus, attempts appeared in the sense of integrating FDI and trade theories through the development of trade models incorporating MNEs.

In this type of models, the firms’ activities are subdivided in two categories: firm specific (firm level) activities or headquarters, such as research and development (R&D), publicity, marketing, distribution, administration services, etc., and plant specific activities (activities related with the production process), generating firm specific fixed costs and plant specific fixed costs, respectively. Markusen (1984) emphasizes the fact that headquarters specific activities have, in many cases, the characteristic of public goods since they can be used in multiple facilities with a single firm level investment. Hence, these models assume the existence of multi-plant economies of scale (or firm level scale economies) associated with a firm specific input.

In the first attempts to introduce MNEs in general equilibrium trade models, only one type of investment was admitted: either vertical or horizontal. While models with vertical FDI seem more important to explain one-way investment flows among countries with significant factor endowments differences (normally, investments by developed countries’ firms in developing countries), the models with horizontal MNEs are particularly important in explaining intra-industry investment flows among developed countries. Since the latter seem to be the prevalent type of FDI, the literature has focused essentially on horizontal FDI. Models that integrate both types of investments, designated knowledge capital models, only appeared recently.

**MNEs and vertical FDI (VER models)**

Helpman (1984) introduces MNEs in a trade model whose emphasis is on the differences in the relative factor endowments (the author assumes the absence of transport costs or other obstacles to trade). The model considers two countries, two goods (the good Y – homogeneous, and the good X - differentiated) and two internationally immobile production factors (the factor work – L, and a general input –

---

6 For example, an innovation resulting from R&D investment can be incorporated into any number of additional facilities without reducing its marginal product in the existent ones. Hence, MNEs contribute
It is admitted that Y is intensive in factor L and X is intensive in factor H. This input has to be adapted in order to produce the desired variety of X. Once adapted, it becomes a firm specific asset that can be supplied to facilities located in different countries without additional costs. The differentiated good sector is characterized by monopolistic competition and subject to increasing returns to scale (there are scale economies at the plant level).

Taking into account the above hypotheses, the explanation for the multinational activity considers that firms expand vertically abroad taking advantage of factor prices differences associated with different relative offers of those factors. The MNEs’ activities only appear in a single direction inside an industry, in firms with a single plant (production does not exist in several facilities due to the hypothesis of absence of trade costs combined with scale economies at the plant level). Some firms in the differentiated sector locate the production of the input (adaptation of factor H) in the country with relatively abundant factor H and the production of the final good in the economy, which is relatively abundant in L, becoming a vertically integrated MNE.

As for trade flows, it is verified that the trade pattern created is varied, a combination of intra-firm, intra-industry, and inter-industry. On the one hand, the country with relatively abundant factor H tends to export headquarter services to the country abundant in L (the headquarters export services of input H to the affiliates, therefore, generating intra-firm trade) in exchange for varieties of a final differentiated good (intra-industry trade) and a homogeneous good (inter-industry trade), and not just exporting the differentiated good, as would happen if factor endowment difference were not significant.

Hence, FDI creates trade and, therefore, a complementary relationship exists. Helpman (op. cit., pp.467) highlights that "(…) in some sense the larger the difference in relative factor endowments the larger is the volume of trade".

Markusen (1984), for his part, developed a model based on the concept of multi-plant economies, intending to provide a reason for the fact that firms engage in FDI instead of portfolio investment. The existence of multi-plant economies also justifies the fact that

to an increase of the world efficiency since they avoid duplication in R&D, necessarily involved in independent NEs.

Although some authors refer that it is a model with horizontal MNEs (the author himself sustains that the MNE will choose production facilities in both countries, becoming a horizontally integrated MNE),
the MNE chooses, by definition, to carry out at least one activity type in each of several countries.

Contrarily to Helpman, Markusen assumes that the countries have an identical factor endowment, in order to show clearly how the multi-plant economies of scale affect the production and trade pattern. Additionally, he considers that the firm specific activities involve a centralization characteristic (apart from the ‘public good’ nature mentioned above), which gives a vertical dimension to MNEs. MNEs tends to disperse the productive activities geographically and to centralize headquarter specific activities in a particular location. As Markusen (op. cit., pp. 208) argues "The total output of two scientists working independently may, for example, be less than their output working cooperatively in the same location. Similarly, communication among different managerial and technical departments is more efficient in centralized location". On the other hand, Markusen admits that a sector is subject to increasing (but reduced) returns, assuring that the monopolist maintains facilities in the two countries (the firm becomes a horizontal MNE) instead of trying to supply the markets from a single facility.

Markusen admits two equilibrium types. First, a duopoly between two national enterprises (NEs) producing the good in each one of the two countries. In this situation, trade does not exist since the output, the goods, and factor prices are the same in the two countries (due to the hypotheses of identical preferences, technology and factor endowment). Second, a multinational monopoly, with the production of the good being monopolized by a MNE, with two plants (one in each of the two countries). If the MNE considers that it is efficient to concentrate certain activities (headquarter specific activities) in the domestic country, the two identical countries will specialize in different activities and will produce different groups of goods. In this case, a MNE can lead to the creation of trade, i.e., the multinational activity can become a cause of trade.

Similarly to the model of Helpman, Markusen’s model assumes an exogenously specified market structure (multinationality is assumed). On the contrary, later works, namely models with horizontal MNEs, consider that the market structure is

the model also validates a vertical dimension for the MNE. Markusen also acknowledges this aspect when considering that the countries will be partially specialized in the activities that they develop in the production of the good, due to the centralization characteristic of the firm specific activities. Taking into account that the emphasis given to the vertical dimension is, in fact, high we opted to include it in this type of models.
endogenously determined as a result of the plant location decisions on the part of the firms, and, in equilibrium, two-way investment could exist.

**MNEs and horizontal FDI (HOR models)**

Horstman and Markusen (1992) and Brainard (1993) developed HOR models that admit the existence of scale economies at the firm level and scale economies at the plant level. These models assume the existence of two identical countries (similar in size, factor endowments and technology). In this way, no trade results from the comparative advantages. They also assume the existence of transport costs among countries that increase with distance. In these models, the choice between exporting or investing abroad (engage in multi-plant production) depends on the trade-off between proximity advantages (such as saving on transport costs) and advantages of scale resulting from the concentration of production in a single plant. The firm compares the additional fixed costs of a second facility and the transaction costs of serving a foreign market through exports. Even though Brainard (*op. cit.*) considers the case in which firms produce differentiated products and Horstman and Markusen (*op. cit.*) consider the case of homogeneous goods, the results are surprisingly similar: for intermediate levels of transport costs, MNEs exist, in equilibrium, when the firm level fixed costs, tariffs and transport costs are high relative to plant specific fixed costs. If the transport costs reached zero, only NEs exporting to each other’s markets would exist (no firm would incur in the fixed costs of a second plant). If the transport costs reached very high levels, only MNEs would exist (with two facilities), which would present lower fixed costs per market, and, consequently, they would remove the NEs (that face prohibitive export costs).

In particular, given the technology characteristics (with firm specific fixed costs and plant specific fixed costs) and Cournot-Nash behavior on the part of the firms, Hortsman and Markusen (*op. cit.*) identify the existence of three equilibrium types. First, an export duopoly constituted by two NEs with a single facility (the most familiar in the trade literature), that tends to appear when the plant specific costs are high relative to the firm specific costs and transport costs. Second, a multinational monopoly (a MNE with two plants, one in each market), that tends to exist when the firm specific costs and transport costs are increased to a point where the duopoly generates negative profits. Finally, a multinational duopoly constituted by two MNEs, both with two
plants. This will be obtained by decreasing the plant specific costs in such a way that the multinational duopoly is lucrative and dominates the export duopoly. In this way, it is verified that the market structure is endogenously determined by technology.

Brainard (op. cit.), for her part, considers the existence of differentiated goods produced under increasing returns to scale by an industry of monopolistic competition. In a simple production process with 2 stages, Brainard’s model presents three possible equilibrium types. First, pure trade equilibrium, constituted only by NEs, with a single plant located in the same market of its headquarters. In this case, there exists two-way balanced trade (intra-industry trade) in differentiated final goods (the volume of intra-industry trade is a decreasing function of the transport costs). Second, pure multinational equilibrium, constituted only by MNEs that carry out productive activities and sell abroad. In this equilibrium, the two-way trade in headquarter services substitutes completely the trade of goods in the differentiated sector. Finally, mixed equilibrium where MNEs coexist with NEs. In this equilibrium, two-way trade in final goods as well as in headquarter services occurs. The resulting type of equilibrium depends on the relative size of the transport costs and firm level scale economies relative to the plant level scale economies. The first equilibrium is more probable as transport costs and the firm level scale economies relative to the plant level scale economies are smaller, while the second equilibrium type is obtained in the opposite circumstances. Mixed equilibrium is maintainable for the intermediate interval of the parameters values. Thus, it is more probable that firms expand their production abroad horizontally when transport costs are larger and scale economies at the plant level relative to the scale economies at the firm (corporate) level are smaller.  

Markusen and Venables (1998), for their part, developed a model that is an extension of the models of Horstman and Markusen (1992) and Brainard (1993), since they explicitly consider the role of the asymmetries among countries. More precisely, they examine the interval for the parameters where the convergence in country size and in relative factor endowment leads or not to more MNEs relative to NEs. Similarly to the two previous models, they conclude that MNEs tends to be found in equilibrium when

---

8 This is what Brainard (1997) designates "proximity-concentration" hypothesis. Having tested empirically this hypothesis, she obtained results that point to the fact that it is reasonably robust. In this way, the prediction that intra-industry trade in headquarter services substitutes the trade in goods depends on the trade-off "proximity-concentration" in each production stadium.
the firm level scale economies and transport tariffs/costs are high in relation to the plant 
level scale economies. Additionally, their conclusion (designated "convergence 
hypothesis") points to the fact that MNEs become more important relative to NEs as 
countries become more similar in size and relative factor endowment, and as the world 
income grows, decreasing international trade. Beginning with different countries, the 
convergence of their characteristics leads, first, to an increase of the trade volume and, 
then, to a reduction, as MNEs begin to displace the NEs. It is worth noting that the 
"new trade theory" concentrates on the competition between NEs of similar countries 
and Markusen and Venables (op.cit.) demonstrate that this is precisely the place where 
it is expected that the activity be dominated by MNEs and not by NEs.

**Knowledge-capital models of MNEs (KK models)**

Theoretical models combining vertical and horizontal FDI are analytically difficult. 
Hence, attempts that propose an integrated treatment of these two investment types only 
recently appeared, allowing that firms have the option of multiple plants or separate the 
headquarter and a single plant geographically. These models have been designated 
knowledge capital models and they are based on three fundamental hypotheses. First, 
firm specific (headquarter) activities, such as R&D, can be geographically separate 
from production; second, headquarter activities are intensive in qualified work relative 
to production; third, headquarter activities have a ‘public good’ characteristic, in the 
measure that they can be used simultaneously by several facilities.

The latter hypothesis creates scale economies at the firm level and creates reasons for 
horizontal FDI while the others create reasons for vertical FDI, locating the headquarter 
activities where the qualified work is cheaper and production where the non-qualified 
work is cheaper. In this models, several combinations of vertical MNEs, horizontal 
MNEs, and NEs can appear endogenously as a function of the parameters values (trade 
costs, differences among countries in terms of factor endowment, investment barriers).

---

9 Barrios et al. (2000) tested empirically the "convergence hypothesis". Using a group of panel data 
relative to pairs of countries for the period 85-96, they estimated the equation that relates the 
multinational activity between two countries (measured in employment terms) with some variables that 
measure the convergence of those countries having obtained results that, to a certain extent, support the 
referred hypothesis. In fact, the authors concluded that differences in market size tend to reduce the 
bilateral multinational activity while the existence of a common language increases that activity 
significantly. However, in some cases, the results indicate that transport costs are a negative determinant, 
contrarily to the convergence hypothesis.
Such is the case of the model by Markusen et al. (1996), with extensions in Markusen (1997), Markusen (2000) and Carr et al. (2001). Since Carr et al. test the model econometrically, the following presentation is based on these authors' article.

This model includes two countries (h and f), two internationally immobile homogeneous production factors (non-qualified work – L, and qualified work - S) and two homogeneous goods (X and Y). Y is intensive in L and produced under constant returns to scale in a competitive industry while X is intensive in S, exhibiting increasing returns to scale and subject to Cournot competition.

Taking into account the above hypotheses, this model admits the existence of six types of firms. On the one hand, type H_i (i=h, f), constituted by horizontal MNEs that maintain facilities in both countries, with the headquarters located in the country i. This type of firms is dominant when the countries are relatively similar in size and in relative factor endowment, and the transport costs are high. Similar countries interact through direct investment with the firms type H_f and H_h invading each other’s market (intra-industry direct investment). On the other hand, type N_i (i=h, f) include NEs that maintain a single plant and headquarters in country i (in the larger country). These firms can or cannot export to the other country (depending on the transport cost). This is the prevalent type of firms when the countries are similar in the relative factor endowment but very different in size. That is particularly true if the larger country is also abundant in qualified work. Finally, type V_i (i=h, f), constituted by vertical MNEs that maintain their headquarters in country i and a single plant in the other country. These firms can or cannot export to country i, depending on the transport costs. This type of firms dominates when the countries are similar in size but very different in terms of factor endowment. In this case, incentive exists to concentrate headquarters in the country abundant in qualified work and production in the other country, unless the transport costs are high. In this way, it is verified that these six types of firms emerge as special cases for certain parameters values. The KK model explains, therefore, the volume of production of the MNEs foreign affiliates with headquarters in a country (origin

10 If plant level scale economies exist, the location of production will also be influenced by the market size.
11 If the countries are different in one of these aspects, a country will be favored as production location and headquarters or in terms of one of these two activities.
12 In this case, NEs located in the larger country will be favored because they avoid more costly production in the smaller market.
country) as a function of the characteristics of the two countries: the origin country and the country where the affiliates are located (or host country).\textsuperscript{13}

Carr et al. \textit{(op.cit.)} tested econometrically the theoretical predictions of the KK model, having obtained results that support this model. However, the same does not happen with Markusen and Maskus (2001). These authors sustain that a problem of the Carr et al. \textit{(op. cit.)} estimation, and in fact a chronicle difficulty in the generality of the empirical works, is that an explicit alternative hypothesis to the model to be tested does not exist. Thus, Markusen and Maskus \textit{(op. cit.)} try to solve this problem, testing the three models mentioned above: the HOR model, the VER model and the KK model. The econometric results obtained strongly support the HOR model and reject the VER model.\textsuperscript{14} Model KK, on the other hand, although presenting a weak performance, rated better than the VER model. Thus, the authors stress the importance of considering an alternative hypothesis: if the VER model was separately estimated, the investigator could conclude that it obtained solid support, but if the HOR model is considered an alternative, that support quickly disappears. These results suggest that horizontal FDI are much more important in the world economy than vertical FDI, especially vertical investments motivated by factor endowment differences.

2.3. The role of demand in the explanation of the relationship between FDI and international trade

Lipsey and Weiss (1984) mention that some theoretical channels exist that influence positively the firms’ foreign demand, and in that way, allow for the existence of a positive relationship between FDI and international trade (namely exports).

If the firm produces a single product, it makes a choice between exporting or producing in the target market basing on the cost and, as a result, export and foreign production are substitutes. The only way its total sales in that market could be positively related with production in that market would be if production allowed the firm to move its demand curve to the right. In accordance with Lipsey and Weiss \textit{(op. cit.)}, foreign demand can

\textsuperscript{13} It should be noted, however, that the hypotheses concerning the size and composition of the fixed costs (firm level scale economies and plant level scale economies) are crucial for the model’s predictions.

\textsuperscript{14} Remember that other empirical works, like that of Brainard (1997), that empirically test the proximity-concentration hypothesis, and Barrios et al. (2000) that test the convergence hypothesis, also lead to results that provide a solid support for the theoretical predictions of the horizontal models.
be stimulated through the supply of important after-sale services. Head and Ries (2001, p.111) corroborate this idea: “A sales office may provide valuable services to foreign consumers that cannot be delivered efficiently through contracting with local agents”. On the other hand, the evidence of a permanent commitment with the market (as is the case of FDI) represents a strong attraction for the customers (particularly in some products, like machinery), while the lack of such a commitment works as an impediment to sales.

If the firm produces final goods and intermediate components or materials used in those goods, it is possible that foreign production be associated with higher levels of exports of intermediate goods that substitute the exports of final goods. Additionally, if the firm produces several final goods, the establishment of a productive unit for one of its products in a foreign market can increase the demand, and, consequently, the exports of other final goods for that market. Quoting Aharoni, Lipsey and Weiss (op. cit., pp.305) mention "(…) It is also a reaction. We make [product x] in India. People see our name, they know us, so they buy our [products y and z] that are imported from the United States". Blonigen (2001) further adds that the productive presence of the firm in a foreign market with a product can increase the total demand for all its products through a faster and more efficient distribution and delivery.

3. The relationship between FDI and international trade: empirical evidence

3.1. Country level studies

Grubert and Mutti (1991) evaluated the relationship between FDI and international trade (exports and imports), using data from 1982 for 33 countries that have commercial relationships with the United States (U.S.). In order to avoid endogeneity problems, the authors sustain that the relationship between FDI and trade is more correctly analyzed using exogenous indicators of the relative attractiveness of operating abroad, such as the average effective tax rate. In particular, if trade and FDI are complementary, then as the cost of operating in a certain country decreases (measured by the rates that firms pay in that country), the level of exports to that country will increase.\(^\text{15}\) The authors obtained

\(^\text{15}\) This method is based on the hypothesis that those rates influence investment but they are not directly related with trade flows.
results that support the existence of complementarity in a bilateral perspective: the U.S. seem to import more from countries where FDI is more accentuated (countries with lower taxes); also, they export more to those countries. However, they consider that a more complete analysis of the relationship between FDI and trade requires a multilateral perspective and not just a bilateral one (as is the case of most of the empirical studies). In fact, according to this perspective, an appearance of complementarity can occur if the American exports to a certain foreign country increase when operations are established in that country. However, if these exports have been displaced from another foreign country, the total exports cannot increase (they may even decrease if exports to third countries decrease). In accordance with Grubert and Mutti (op. cit., pp.286) *"The separate analysis of local sales and sales to other countries indicates that the potential displacement of exports to third country markets can be significant".*

Clausing (2000) examines the relationship based on two groups of panel data for the period 1977-1994. The first contains data relative to the operations of American MNEs in 29 host countries and data about American exports (intra-firm and arm-length), seeking to study the relationship between American exports and American FDI. The second includes data relative to operations in the U.S. accomplished by affiliates of MNEs with headquarters in 29 countries and data on American imports (intra-firm and arm-length), aimed at analyzing the relationship between American imports and FDI in the U.S.. Clausing estimates specifications that relate the trade flows with variables that reflect the FDI and with typical variables, such as the exchange rate, income of the countries involved in the exchange, the distance between countries and trade barriers.16

As a measure of FDI, Clausing uses the affiliates’ net local sales, that is, excludes the affiliates’ sales to another countries, thus adopting a bilateral perspective, and also excludes the imports from the parent firm. The results found demonstrate that FDI positively influences trade. A strong complementarity exists between intra-firm trade and FDI (multinational activity may stimulate exports of parts or related products), and a weaker complementarity exists between inter-firm trade and FDI, since some exports may be displaced. This occurs also for the relationship between imports and FDI.17

16 The equations that include these variables are known as “gravity equations”.
17 This conclusion is also supported by the analysis of the relationship between price variables that measure the cost of operating abroad and exports, similarly to Grubert and Mutti (op. cit.), although in this case the results for the relationship between imports and the inward FDI are less conclusive.
3.2. Industry level studies

Lipsey and Weiss (1981) examine the relationship using cross-section data, for 1970, relative to American exports, by industry (and of 13 other great exporter countries) for 44 destination countries. The authors relate these exports with the characteristics of the destination countries (size, whether they are members of the EEC or not, distance from the U.S. and from Germany) and with the production in those countries by affiliates of firms with headquarters in the U.S. and in the other 13 countries (FDI indicator). In 14 industries that represent most of the American productive investment abroad, Lipsey and Weiss (op. cit., pp. 494) "(...) find quite consistently that the level of activity of U.S. manufacturing affiliates is positively related to U.S. exports and, in less developed country markets, negatively related to exports by 13 other countries. The number of foreign-owned manufacturing affiliates is positively related to exports by foreign countries (...)". As a result, the authors conclude that the activity of the foreign affiliates tends to promote the exports of the countries where the parent firm is located, existing therefore a complementarity relationship (in a bilateral perspective). Simultaneously, the results obtained seem to indicate that the production of the American firms’ affiliates substitutes the exports from the other 13 countries and the production of the affiliates of MNEs headquartered in these 13 countries substitutes American exports. Hence, in a multilateral perspective, the relationship can be negative.

Pfaffermayr (1996) analyzes the relationship using data relative to seven Austrian industries during the period 1980-1994, and using the accounting value of FDI stock as a measure of multinational activity. The author argues that outward FDI and exports should be considered endogenous variables with common determinants, such as the intensity in capital, work, qualifications and R&D (it is worth noting that in spite of the recognized interdependence between FDI and exports, few empirical works analyse both simultaneously).

Basing on this endogenous framework, the methodology followed by the author consisted of the estimate of a simultaneous equations system, having found a significant and stable complementarity relationship between FDI and exports. An increase in FDI influences significantly (in a positive way) the exports while the positive impact of an increase of exports in FDI is confirmed for lower significance levels.
Brainard (1997) tests the relationship between trade and FDI on the cross-section data of 1989 relative to 63 industries and 27 countries. The author's analysis is confined to the U.S.’s bilateral relationships: outward FDI (sales of the foreign affiliates of MNEs headquartered in the U.S. and American exports) and inward FDI (sales by American affiliates of MNEs headquartered in other countries and American imports). Similarly to previous studies, Brainard (op. cit.) uses instrumental variables in order to avoid the simultaneity problems between the affiliates’ sales and trade. Thus, as an indicator of FDI, she used the affiliates’ employment level and their net assets, and obtained results that point to the existence of a positive relationship between FDI and trade.

3.3. Firm level studies

Lipsey and Weiss (1984) tried to improve their previous study by disaggregating the data considerably more. As they argue, "By comparing U.S.-owned production and trade across countries within industries we avoided some of the bias that might result from the operation of industry comparative advantages that promoted both direct investment and export " (Lipsey and Weiss, op. cit., pp.304). Using cross-section data (for 1970) of individual firms, the authors related the exports of each firm for each one of five areas (composed by developed countries) with the parent firm's characteristics (parent firm size measured by its sales in the U.S.) and with the output of its foreign affiliates (affiliates’ sales minus imports from the U.S.) and with the size of the market included in each area (measured by the income of that area - GDP). The results obtained indicate the existence of complementarity between the affiliates’ production and the parent firm's exports to the area in which the production takes place: the larger the production of an American firm affiliate in a certain foreign area, the larger, in general, the parent firm's exports to that area. This relationship is further emphasized between the foreign production and the export of intermediate goods.¹⁹

Mucchielli et. al. (2000), using data relative to 421 French firms for the year 1993 and using the number of workers of the affiliate as proxy/indicator of FDI, obtain results that support the existence of complementarity between global exports and FDI (as well as between imports and FDI, although less significant). However, analyzing the volume of trade of the French firms that invest abroad, either with their own affiliates (intra-

---

¹⁸ Thus, this study presents the referred problems related with the subject of endogeneity.
firm trade) or with other firms of foreign countries (inter-firm trade), the conclusion points to a strong complementarity between FDI and intra-firm trade and a substitution between FDI and inter-firm trade (for exports and for imports). In terms of global trade, since FDI seems to influence relatively more exports than imports, the trade balance of the French firms that invest abroad is globally positive.

Head and Ries (2001) use a group of panel data containing 932 Japanese firms during a 25-year period (1966-90) to investigate the effects of FDI on exports. They distinguish productive affiliates from distribution affiliates, thus increasing the chances of obtaining a negative effect of productive investment in exports. As an indicator of FDI, they use the number of investments in production (m - manufacturing) and distribution (d), having lagged these variables a year in order to remove the influence of shocks that affect contemporaneously FDI and exports (this means that the FDI variables are pre-determined relative to exports). The results obtained allowed them to conclude that the firms that increase their investments abroad also increase their exports, that is, in the full sample of firms, the authors found complementarity. However, the relationship varies across firms. In fact, for a group of firms that are not vertically integrated (that are unlikely to ship intermediates to overseas production affiliates) the foreign productive facilities seem to substitute their own exports. Thus, the authors conclude that that a source of complementarity is the sales promotion of intermediate goods.20

3.4. Product level studies

According to Blonigen (2001), several advantages exist in the use of this disaggregation level. On the one hand, it permits that the complementarity effect resulting from vertical production relationships and the substitution effect of exports by affiliates production be modeled and tested separately in the same equation. On the other hand, as it is centered on a single product, the demand complementarities between products are not disguised by the data. The author's analysis based on data relative to Japanese production in the U.S. and to exports to this country of two types of products: automobile components (that have a strong vertical relationship with automobile production) and final consumption products (data of 1978-1991), using data on employment levels as proxy

---

20 It should be noted that, similarly to Lipsey and Weiss (1981), this work raises several problems, since it does not take into account the endogeneity problem.
for the affiliates’ production. As expected, the results found by Blonigen (op. cit.) demonstrate a strong positive relationship between the Japanese production of automobiles in the U.S. (the industry that uses the inputs) and the Japanese exports of automobile components to the U.S. (vertical production relationships are associated with strong complementarity between exports and foreign production). In terms of the Japanese production of automobile components in the U.S., it was expected that the Japanese exports of components decrease. The results obtained for ten specific components confirm the existence of high substitution effects (even without controlling for the potential increase in demand for these products due to the increase of the Japanese production of automobiles in the U.S.). The analysis for 11 products of final consumption suggests equally the substitution of exports for local production for most of the products. Hence, product level data show a strong substitution effect, unlike previous studies at a more aggregate level.

3.5. Summary

In table 1, we present a summary of the existing empirical works, highlighting the analysis level, the sample type, the variables used to measure the multinational activity and the main results obtained.

It is clear from the table that the generality of the empirical works highlight a complementarity relationship between FDI and international trade (namely exports), although we can find some exceptions. The fact that the empirical results were mixed is not surprising, when taking into account that the studies use different samples, different variables as FDI indicators and take place at different levels of analysis (country, industry, firm, product). As noted by Blonigen (2001, pp.82) "One might expect that the more disaggregated nature of the firm-level data would be more likely to yield net substitution, yet almost all of these studies find net complementarity as well".

---

20 It is worth noting, however, that the authors did not test directly this proposition since they did not have data on the parent firm’s exports of intermediate goods. Nevertheless, the expectation is that vertically integrated firms are those that will most probably export intermediate goods.
Table 1: Synthesis of the existing empirical works

<table>
<thead>
<tr>
<th>Level of analysis</th>
<th>Author/ Year</th>
<th>Sample type</th>
<th>Multinational activity indicator (FDI)</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country level studies</td>
<td>Clausing (2000)</td>
<td>Panel data</td>
<td>Net local sales</td>
<td>Complementarity (stronger in the case of intra-firm trade)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Taxes paid by affiliates operating abroad (1)</td>
<td>Complementarity</td>
</tr>
<tr>
<td></td>
<td>Grubert and Mutti (1991)</td>
<td>Cross-section</td>
<td>Average effective tax rate (1)</td>
<td>Complementarity in a bilateral perspective (2)</td>
</tr>
<tr>
<td>Industry level studies</td>
<td>Brainard (1997)</td>
<td>Cross-section</td>
<td>Employment level and net assets</td>
<td>Complementarity</td>
</tr>
<tr>
<td></td>
<td>Lipsey and Weiss (1981)</td>
<td>Cross-section</td>
<td>Productive affiliates’ sales (less imports from the origin country)</td>
<td>Complementarity in a bilateral perspective (2)</td>
</tr>
<tr>
<td></td>
<td>Pfaffermayr (1996)</td>
<td>Panel data and cross-section</td>
<td>Accounting value of the FDI stock</td>
<td>Complementarity</td>
</tr>
<tr>
<td>Firm level studies</td>
<td>Head and Ries (2001)</td>
<td>Panel data</td>
<td>Nr° of production and distribution investments</td>
<td>Complementarity (3)</td>
</tr>
<tr>
<td></td>
<td>Lipsey and Weiss (1984)</td>
<td>Cross-section</td>
<td>Productive affiliates’ sales (less imports from the origin country)</td>
<td>Complementarity</td>
</tr>
<tr>
<td></td>
<td>Mucchielli et. al. (2000)</td>
<td>Cross-section</td>
<td>Foreign affiliates’ workers</td>
<td>Complementarity of the intra-firm trade Substitution of the inter-firm trade</td>
</tr>
<tr>
<td>Product level studies</td>
<td>Blonigen (2001)</td>
<td>Panel data</td>
<td>Foreign affiliate workers</td>
<td>Substitution (4)</td>
</tr>
</tbody>
</table>

(1) Indicator of the cost of foreign operation (exogenous indicator)
(2) Possibility of substitution in a multilateral view
(3) Substitution in the case of firms not vertically integrated
(4) Complementarity in the case of vertical production relationships

The existence of complementarity even for less aggregate data, such as firm level data, is apparently surprising. However, since most MNEs are multi-product firms, the strong complementarity results found can be due to the existence of demand complementarities and/or vertical production relationships between the firm’s products, which are difficult to identify even with firm-level data. As Blonigen (op. cit., pp.84) mentions, even firm-level data "(…) does not allow one to separately identify a substitution effect to the extent that the firm is multi-product, which is quite likely given that the companies examined are large MNCs". Furthermore, complementarity can also result from the existence of potential sources of spurious positive relationships between FDI and trade.

The existing empirical works explored two types of sources: aggregation bias and endogeneity. Aggregation bias refers to the resulting bias from the use of aggregate data. Blonigen (op. cit.) provides evidence that aggregation bias can have an important
role, showing that the substitution effects are easier to identify in disaggregate data at the product level. The subject of endogeneity refers to the fact that the methodology applied in the empirical studies usually uses the foreign affiliates’ sales/production as proxy/indicator of FDI. However, this FDI indicator is an endogenous variable in the measure that it is influenced by factors that also affect exports (for example, the size of the market). Thus, a simultaneous increase of exports and foreign affiliates’ sales can reflect only an exogenous increase of the host country demand, and it is incorrect to interpret this simultaneous increase as a complementary effect.21

To deal with this problem, several authors, like Grubert and Mutti (1991) used instrumental variables (exogenous indicators of the relative attractiveness of operating abroad), continuing, however, to find complementarity. Hence, this subject does not seem to be as important as aggregation bias. However, one problem with this solution resides in the choice of the appropriate instrument, that is to say, variables that have a direct effect on FDI but do not belong to the export equation. On the other hand, for studies at the firm level, the endogeneity problems are even larger, because it is probable that a firm that possesses good-quality products or efficient production techniques achieves high export volumes and invest more abroad, without which FDI would stimulate exports.

Thus, studies at the firm level "(...) must also address the issue of unobserved heterogeneity amongst the sources of exports and overseas investment " (Head and Ries, op. cit., pp.109). Although these authors have take into account heterogeneity among the firms (through the use of fixed effects), the same did not occur with other studies at the firm level.

21 Cantwell (1994, p.305) argues "Although much of the empirical literature on the association between international trade and production has proxied MNC activity on international production by using FDI data, it is in this respect potentially misleading from a theoretical point of view". Hence, it is inadequate the association that is usually made between international production and FDI. In fact, the growth of international production does not necessarily require the growth of FDI, since as the affiliate evolves it is increasingly probable that its financing sources become local.
4. Conclusion

The relationship between FDI and international trade is complex since there are several aspects that must be taken into account. On the one hand, headquarters’ exports can be substituted by the affiliates’ local sales (in the host country), but foreign production can use input imports from the parent firm. On the other hand, inter-firm trade (between the home and host country) can also change. To these bilateral effects, we can add a reduction of the exports from the home country to third countries. Therefore, in a multilateral perspective, the relationship becomes even more complex.

Theoretically, FDI explanatory theories and models of horizontal FDI uphold a substitution relationship while models of vertical FDI and considerations concerning demand sustain the existence of complementarity. Taking into account that the prevalent type of FDI seems to be horizontal, we would expect that a substitution relationship prevailed in empirical works. However, it is possible to verify that most of the empirical works found complementarity. In this way, there are apparent incongruities between the theory and the empirical work. However, the strong complementarity demonstrated in the empirical works can, at least in part, be explained by the multi-product nature of the firm (which means that demand complementarities and/or vertical production relationships can exist between the products of the firm), as well as by the existence of sources of positive spurious relationships between FDI and trade, as is the case of endogeneity and aggregation bias. In particular, Blonigen (2001) showed that the bias resulting from the use of aggregate data can have an important role. However, the data used by Blonigen include multiple firms, and consequently, some of them can serve the market only with local production while others can only be exporters. This fact suggests that future works should try to use data at the firm product level.

It is clear, therefore, that despite the fact that recently some progress has been made in the study of the relationship between FDI and international trade through the use of data at a more disaggregated level, some aspects need further study.

First, it is necessary to introduce the imports side in the analysis. If the main purpose is to find out how international trade is affected by FDI, the fact that most of the studies
do not include the imports side means that it is not possible to obtain rigorous conclusions on the type of existing relationship.

Second, it is necessary to conduct investigations not only at the industry level and firm level but also at the product level in order to generalize results. The reduced number of available studies, particularly at the product level, does not allow for that generalization.

Furthermore, if possible, future works should use disaggregated data at the level of the product per firm, although this does not mean that it is the appropriate disaggregation level for the study of the relationship between FDI and international trade. In fact, Clausing (2000) mentions that there are some advantages in the use of more aggregate data. For example, it can be true that foreign production substitutes exports of some firms while indirectly stimulating exports on the part of other firms, and consequently the use of desegregate data does not allow one to apprehend such indirect complementarities. Therefore, the ideal situation consists in using several types of data.

Finally, it is necessary to conduct analyses that can differentiate clearly between horizontal and vertical FDI. Markusen (2000) paid attention to this matter. He also suggests the need to try to obtain data on the intra-firm trade in services of intangible assets (management services, technicians, engineering, marketing, etc).

To sum up, a lot of work remains to be done in order to clarify the type of existing relationship between FDI and international trade.

References


<table>
<thead>
<tr>
<th>No</th>
<th>Author(s)</th>
<th>Title</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>139</td>
<td>Sandra Silva</td>
<td>On evolutionary technological change and economic growth: Lakatos as a starting point for appraisal</td>
<td>March 2004</td>
</tr>
<tr>
<td>138</td>
<td>Maria Manuel Pinho</td>
<td>Political models of budget deficits: a literature review</td>
<td>March 2004</td>
</tr>
<tr>
<td>137</td>
<td>Natércia Fortuna</td>
<td>Local rank tests in a multivariate nonparametric relationship</td>
<td>February 2004</td>
</tr>
<tr>
<td>136</td>
<td>Argentino Pessoa</td>
<td>Ideas driven growth: the OECD evidence</td>
<td>December 2003</td>
</tr>
<tr>
<td>135</td>
<td>Pedro Lains</td>
<td>Portugal's Growth Paradox, 1870-1950</td>
<td>December 2003</td>
</tr>
<tr>
<td>134</td>
<td>Pedro Mazeda Gil</td>
<td>A Model of Firm Behaviour with Equity Constraints and Bankruptcy Costs</td>
<td>November 2003</td>
</tr>
<tr>
<td>133</td>
<td>Douglas Woodward, Octávio Figueiredo and Paulo Guimarães</td>
<td>Beyond the Silicon Valley: University R&amp;D and High-Technology Location</td>
<td>November 2003</td>
</tr>
<tr>
<td>131</td>
<td>Aurora Teixeira and Natércia Fortuna</td>
<td>Human Capital, Innovation Capability and Economic Growth</td>
<td>July 2003</td>
</tr>
<tr>
<td>130</td>
<td>Jorge M. S. Valente and Rui A. F. S. Alves</td>
<td>Heuristics for the Early/Tardy Scheduling Problem with Release Dates</td>
<td>May 2003</td>
</tr>
<tr>
<td>129</td>
<td>Jorge M. S. Valente and Rui A. F. S. Alves</td>
<td>An Exact Approach to Early/Tardy Scheduling with Release Dates</td>
<td>May 2003</td>
</tr>
<tr>
<td>128</td>
<td>Álvaro Almeida</td>
<td>40 Years of Monetary Targets and Financial Crises in 20 OECD Countries</td>
<td>April 2003</td>
</tr>
<tr>
<td>127</td>
<td>Jorge M. S. Valente</td>
<td>Using Instance Statistics to Determine the Lookahead Parameter Value in the ATC Dispatch Rule: Making a good heuristic better</td>
<td>April 2003</td>
</tr>
<tr>
<td>126</td>
<td>Jorge M. S. Valente and Rui A. F. S. Alves</td>
<td>Improved Heuristics for the Early/Tardy Scheduling Problem with No Idle Time</td>
<td>April 2003</td>
</tr>
<tr>
<td>125</td>
<td>Jorge M. S. Valente and Rui A. F. S. Alves</td>
<td>Improved Lower Bounds for the Early/Tardy Scheduling Problem with No Idle Time</td>
<td>April 2003</td>
</tr>
<tr>
<td>124</td>
<td>Aurora Teixeira</td>
<td>Does Inertia Pay Off? Empirical assessment of an evolutionary-ecological model of human capital decisions at firm level</td>
<td>March 2003</td>
</tr>
<tr>
<td>123</td>
<td>Álvaro Aguiar and Manuel M. F. Martins</td>
<td>Macroeconomic Volatility Trade-off and Monetary Policy Regime in the Euro Area</td>
<td>March 2003</td>
</tr>
<tr>
<td>120</td>
<td>Ana Paula Serra</td>
<td>The Cross-Sectional Determinants of Returns: Evidence from Emerging Markets' Stocks</td>
<td>October 2002</td>
</tr>
<tr>
<td>119</td>
<td>Cristina Barbot</td>
<td>Does Airport Regulation Benefit Consumers?</td>
<td>June 2002</td>
</tr>
<tr>
<td>118</td>
<td>José Escaleira</td>
<td>A Procura no Sector das Artes do Espectáculo. Tempo e Rendimento na Análise das Audiências. Um Estudo para Portugal</td>
<td>June 2002</td>
</tr>
<tr>
<td>117</td>
<td>Ana Paula Serra</td>
<td>Event Study Tests: A brief survey</td>
<td>May 2002</td>
</tr>
<tr>
<td>116</td>
<td>Luís Delfim Santos and Isabel Martins</td>
<td>A Qualidade de Vida Urbana - O caso da cidade do Porto</td>
<td>May 2002</td>
</tr>
<tr>
<td>115</td>
<td>Marcelo Cabûs Klötze and Fábio Luiz Biagini</td>
<td>A Restutuição do</td>
<td></td>
</tr>
</tbody>
</table>
Editor: Prof. Aurora Teixeira (ateixeira@fep.up.pt)
Download dos artigos em:
http://www.fep.up.pt/investigacao/workingpapers/workingpapers.htm