



Venice Summer Institute 2005

Workshop on “Recent Developments in International Trade:
Globalization and the Multinational Enterprise”

18–19 July 2005

Venice International University, San Servolo



Trade Costs and Foreign Direct Investment

J. Peter Neary

CESifo
Poschingerstr. 5, 81679 Munich, Germany
Phone: +49 (89) 9224-1410 - Fax: +49 (89) 9224-1409
office@CESifo.de
www.cesifo.de

TRADE COSTS AND FOREIGN DIRECT INVESTMENT ^{*,†}

J. Peter Neary

University College Dublin and CEPR

July 11, 2005

Abstract

This paper reviews the theory of foreign direct investment (FDI), focusing on an apparent conflict between theory and recent trends in the globalized world. The bulk of FDI is horizontal rather than vertical, but horizontal FDI is discouraged when trade costs fall. This seems to conflict with the experience of the 1990s, when trade liberalisation and technological change led to dramatic reductions in trade costs yet FDI grew much faster than trade. Two possible resolutions to this paradox are explored. First, horizontal FDI in trading blocs is encouraged by intra-bloc trade liberalisation, because foreign firms establish plants in one country as export platforms to serve the bloc as a whole. Second, cross-border mergers, which are quantitatively more important than greenfield FDI, are encouraged rather than discouraged by falling trade costs.

JEL: F13

Keywords: Cross-border mergers and acquisitions; Export platform FDI; Foreign direct investment; International trade policy; Trade liberalisation.

*Address for Correspondence: Department of Economics, University College Dublin, Belfield, Dublin 4, Ireland; tel.: (+353) 1-716 8334; fax: (+353) 1-283 0068; e-mail: peter.neary@ucd.ie.

†Prepared for a CESifo Workshop on "Recent Developments in International Trade: Globalization and the Multinational Enterprise," Venice International University, 18 -19 July 2005. This research is part of the International Trade and Investment Programme of the Geary Institute at UCD. For helpful discussions I am very grateful to Frank Barry, René Belderbos, Ben Ferrett, Keith Head, Dermot Leahy, Philippe Martin and Stephen Yeaple.

1 Introduction

Foreign direct investment (FDI) is one of the key features of the modern globalized world. While some traders maintained international links in the late medieval and early modern periods, and multinational firms became important in many industries in the late nineteenth century, the period since the Second World War and in particular since about 1985 saw an explosion in FDI, both in absolute terms and relative to the levels of trade and GDP.¹

Matching these real-world developments, an extensive economic literature has developed in recent years which attempts to explain the nature, causes and consequences of FDI.² The central plank of the now-standard theoretical framework used in this literature is the so-called proximity-concentration trade-off. This suggests that FDI occurs when the benefits of producing close to a foreign market outweigh the loss of economies of scale from producing exclusively in the firm's home plant. As we will see, there is much to be said for this model, and a lot of empirical evidence in support of it. However, it makes a key prediction which seems to run counter to the experience of the 1990s. If FDI is driven primarily by the proximity-concentration trade-off, then falls in trade costs should discourage it, as the benefits of concentrated production increasingly outweigh the gains from improved market access. It should be emphasized that trade costs include both tariffs and transport costs, and that measuring them in even the simplest contexts poses major conceptual and practical problems.³ Yet it seems incontrovertible that both technical and policy-induced barriers to trade fell dramatically during the 1990s. This is especially true within Europe, where trade barriers were dramatically reduced under the EU Single Market programme, even as FDI boomed.

This apparent paradox is the organizing principle of the paper, which presents a selective overview of the theory and empirics of FDI. The paper first outlines the simplest case of horizontal FDI, and then considers in turn vertical FDI, export-platform FDI and cross-border mergers and acquisitions. Throughout I try to present results in the simplest possible way, paring the models down to their essentials to focus attention on the key assumptions, and making use of diagrams where possible. As we will see, there are many ways of explaining the paradox, but their relative empirical importance remains to be determined.

¹See UNCTAD (2000), Markusen (2002, Chapter 1) and Barba Navaretti and Venables (2004, Chapter 1) for summaries of the stylized facts about FDI.

²Markusen (2002) and Barba Navaretti and Venables (2004) give overviews of the theory of FDI and multinational corporations.

³See Anderson and van Wincoop (2004) and Anderson and Neary (2005).

2 The Proximity-Concentration Trade-Off

I begin with the simplest framework in which the proximity-concentration trade-off can be illustrated.⁴ Consider a single potential multinational which is the monopoly supplier of a product and seeks to determine the optimal mode of serving a foreign market. The assumption that the firm is a monopoly can be related to the “O” in the “OLI” or Ownership-Location-Internalization framework of Dunning (1973): the firm possesses unique advantages in terms of product quality, marketing, organisation or R&D, which give it an ownership advantage over other potential firms. It is also consistent with models of monopolistic competition: many firms compete against each other, each producing a symmetrically differentiated product, but from the perspective of an individual firm the demand function it faces is given. Of course, the assumption is not consistent with perceived interdependence between oligopolistic firms, which seems *a priori* likely to characterise the markets in which many multinational corporations operate. However, the main points I want to make do not require an oligopolistic setting and I postpone consideration of oligopoly until Section 5. I also concentrate throughout the paper on a single industry in partial equilibrium. Embedding such an industry in general equilibrium is essential for a complete analysis, and much recent research in the theory of FDI has done just this. However, the points I wish to highlight can be adequately addressed in partial equilibrium.

The operating profits which the firm earns in the foreign or host-country market depend on many factors, some under its control (such as output and advertising) others not. Assume in this section that these factors are independent of the manner in which the firm serves the market. In particular, there is no comparative advantage reason which makes it cheaper or more expensive to produce in the firm’s home country or in the host country. In that case we can focus on a single key determinant of operating profits, the unit cost of serving the market, denoted by t . Once again, part of this cost is independent of how the market is served: marketing, distribution and after-sales service costs for example. However, for our purposes it makes sense to focus on the incremental cost of serving the market from abroad, so t should be understood as a measure of the external trade barrier, which is zero if the firm locates in the market and otherwise includes both tariffs and transport costs. Hence we can write the firm’s operating profits as a reduced form function of t , $\pi(t)$, where all the other determinants of operating profits, which are independent of how the market is served, are subsumed into the π function. It is easy to check that a rise in t reduces both sales and profits in the market, so π' is negative. (See the Appendix for more details.)

We can now state the firm’s profits from alternative ways of serving the market. If it does so via exports then its total profits Π^X are simply $\pi(t)$. Of course, the firm also incurs fixed costs in its home country: these

⁴The model in this section is standard. See for example, Smith (1987, Section 2) or Markusen (2002, Chapter 2). The analytic properties of Fig. 1 are taken from Neary (2002).

are an important determinant of its willingness to serve the foreign market at all, but they are independent of how it does so, so little is lost by ignoring them. By contrast, investing in a local plant to serve the market will incur additional fixed costs which we denote f . (It is convenient to interpret fixed costs as measured with respect to the size of the domestic market: see the appendix for a justification.) The benefit from this proximity is the saving on trade costs, which boosts operating profits to $\pi(0)$. Hence the total profits from engaging in FDI, which we denote Π^F , equal $\pi(0) - f$. The choice between FDI and exports therefore depends on the trade-cost-jumping gain, which we denote $\gamma(t, f)$:

$$\Pi^F - \Pi^X = \gamma(t, f) \quad \text{where:} \quad \gamma(t, f) \equiv \pi(0) - f - \pi(t) \quad (1)$$

As the signs under the arguments indicate, this gain is increasing in trade costs t but decreasing in fixed costs f .

All this can be illustrated in (f, t) space as in Figure 1. Profits from exporting, Π^X , are independent of f , decreasing in t , and strictly positive for $t < \tilde{t}$, where \tilde{t} is the threshold tariff at which exports are just profitable, and is defined by $\pi(\tilde{t}) = 0$. By contrast, profits from FDI, Π^F , are independent of t , decreasing in f , and strictly positive for f less than the threshold level of fixed costs $\pi(0)$ at which FDI is just profitable. It is now easy to read off the figure the different modes of serving the market which a profit-maximising firm will choose. If both trade costs and fixed costs are above their threshold values then the firm cannot make positive profits and so will not serve the market at all: this corresponds to the region denoted “ O ”. If only one cost variable exceeds its threshold value then the firm will opt for the one mode of serving the market which yields positive profits. Finally, if both variables are below their threshold values, then both modes are profitable, and the choice between them depends on the sign of $\Pi^F - \Pi^X$ given by (1). Setting this equal to zero defines the boundary between the regions in Figure 1 in which the firm will engage in exports and FDI, labelled “ X ” and “ FDI ” respectively.

This analysis is the foundation of the proximity-concentration trade-off, and its implications are clear. Higher fixed costs favour exporting over FDI, whereas higher trade costs favour FDI over exporting. Furthermore, the same firm never engages in both FDI and trade.⁵ The model is a disarmingly simple one, and it is worth teasing out the riches of its implications. It can be interpreted in either a time-series or cross-section context, and, though it is stated explicitly in terms of a single firm, it can also be applied at the level of sectors or whole countries. Thus for comparisons across time the model implies that falls in trade costs should encourage FDI relative to exports, and vice versa. For comparisons across sectors it

⁵The model has been extended by Helpman, Melitz and Yeaple (2004) to allow firms within the same monopolistically competitive industry to have different efficiency levels. This permits two-way flows of FDI, but still predicts that a given firm will engage in either FDI or trade and not both.

implies that lower trade costs should be associated with more exports relative to FDI and vice versa. And for comparisons across space it implies that closer markets should be served by exports and further ones by FDI.

Since I will spend much of the rest of the paper criticizing these implications, it is only fair to begin by noting that there is considerable though not overwhelming evidence in their favour. Consider first the econometric evidence. Brainard (1993) showed that, as trade and transport costs rise, the *level* of outward FDI from the U.S. (measured by local sales of U.S. affiliates) falls, but the *share* of FDI in affiliate sales plus U.S. exports rises. Thus while the predictions of the theory are not borne out in an absolute sense they are confirmed in relative terms: lower trade costs lead to a substitution away from FDI towards exports. Similar results are found by Carr, Markusen and Maskus (2001) and by Yeaple (2003b). However, results for the effects of distance (which is positively but not strongly correlated with transport costs) are less favorable to the theory.⁶ Both FDI and exports fall with distance, and its effects on the share of FDI is sensitive to the specification used. Of course, distance may be proxying for factors other than trade costs, such as the costs of communicating with foreign subsidiaries, but this is clearly inconsistent with the simple proximity-concentration trade-off.

Over and above this econometric evidence, there is also considerable case-study evidence which is consistent with the proximity-concentration trade-off. Indeed, case studies are an important supplement to econometric estimates and often highlight special features which large-sample econometric studies are likely to miss. Let me mention two. The first is the experience of Ireland in the 1930's, which transformed rapidly from an extremely open economy to a highly protected one following a change of government in 1932.⁷ (See Neary and Ó Gráda (1991) for details and further references.) Despite the small size of the Irish market, the theory would predict a large inflow of FDI. Yet this did not occur until six years later. The reason is simple: protection had been imposed by the new nationalist government as part of a campaign to reduce British influence in Ireland; when British firms responded by trying to set up affiliates in Ireland the Irish government passed new legislation prohibiting their doing so. Only when this legislation was relaxed in 1938 did FDI increase significantly. A second case study is that of Japanese electronics firms in the European Community (EC) in the late 1980s by Belderbos and Sleuwaegen (1998). They concluded that the rapid increase in Japanese manufacturing investments in the late 1980s was mainly induced by EC anti-dumping and other trade restricting measures targeting Japanese firms and that such tariff-jumping investment substituted for exports from Japan. But since anti-dumping duties are a form of contingent protection, the

⁶I am indebted to Stephen Yeaple for these findings.

⁷"Ireland" is used here to refer to the political unit which was an independent state from 1922 until 1949 though with a constitutionally ambiguous status in international law, reflected in its official designation as the "Irish Free State". It became a republic in 1949.

effect of FDI was not merely to evade tariffs but to ensure that they were not imposed in the first place. The busy econometrician (busy because she has hundreds of other data points to worry about) would in the case of Ireland in the mid-1930s observe protection but no FDI, and in the case of EC affiliates of Japanese electronics firms in the late 1980s observe FDI but no protection. Yet both episodes are fully consistent with the logic of the proximity-concentration trade-off, once it is supplemented by obvious features of the institutional and political context.

Nevertheless, there remains the puzzle noted in the introduction. How can the theory be reconciled with the enormous increase in FDI in the 1990s, especially into the European Union? Clearly tariffs and transport costs fell dramatically, but FDI rose much faster than exports. With Ireland as a prime example of a host country which benefitted enormously from this inflow, how is it that the simple theory explains the Irish experience in the 1930s but not in the 1990s? A clue may come from an old literature which explored the issue of whether exports and FDI are substitutes or complements in competitive factor-endowment models. This literature was initiated by Mundell (1957), who showed that they are perfect substitutes in the textbook two-sector two-country Heckscher-Ohlin model: barriers to trade encourage international capital flows which if unimpeded lead to an equilibrium identical to that which would obtain under free trade. However, extensions by Markusen (1983), Jones and Neary (1984) and Neary (1995) among others showed that trade liberalisation can encourage FDI if the induced capital flows enter the export sector, because countries differ either in technology or in endowments of sector-specific factors. This literature has fallen out of fashion, as its view of FDI as physical flows of a productive factor has given way to an industrial-organisation inspired view of FDI as an intra-firm transfer of intangible assets by multinational corporations. But its insights can help explain the anomalies we have found in the predictions of the proximity-concentration trade-off hypothesis. In the next two sections we turn to two such approaches.

3 Vertical versus Horizontal FDI

The first framework in which FDI may encourage rather than substitute for exports is when it is vertical rather than horizontal. The theory of vertical FDI originated with Helpman (1984), who showed in a Heckscher-Ohlin model that, when stages of production differ in their factor intensities, international differences in factor endowments give incentives for vertical disintegration by firms. More generally, it can arise from any comparative-advantage reason which makes it more profitable to locate one or more stages of production outside the market where the final good is sold.

The simplest example of vertical FDI, and one that is easily linked to the model of the last section, is where the firm has two stages of production: “headquarter services” (in Helpman’s phrase) provide internal

public goods to the firm, are located in the parent country and incur fixed costs only, while the production stage can be located wherever it is most profitable to do so. Assuming for simplicity that each unit of output requires a single unit of labour, we can write the operating profits of serving the parent market as $\pi^*(c)$, where c includes both factor costs and market access costs. Ignoring demand in the host country for the present, the firm now has two options. If it remains a domestic firm and supplies its market from its parent plant, where w^* is the local wage rate, its profits will equal $\pi^*(w^*)$, which we can denote Π^D . Alternatively, it can engage in FDI and locate a new plant in the host country which exports all its output back to the parent country, incurring a trade cost of t^* . In that case, it incurs a plant-specific fixed cost f as in the previous section, and earns operating profits of $\pi^*(w + t^*)$, where w is the host-country wage. The relative profitability of FDI now becomes:

$$\Pi^F - \Pi^D = \mu(w + t^*, w^*) - f \quad \text{where:} \quad \mu(w + t^*, w^*) \equiv \pi^*(w + t^*) - \pi^*(w^*) \quad (2)$$

The new element is $\mu(w + t^*, w^*)$ which we can call the “offshoring gain.” Crucially it is *decreasing* in the parent country tariff, so trade liberalisation will encourage FDI. Figure 2 illustrates. Now the FDI region corresponds to low values of *both* fixed costs and trade costs.

If the host country market is non-negligible then we get a combination of vertical and horizontal motives. Now the choice between FDI and staying at home (labelled “DX” since it involves producing for both domestic sales and exports) depends on the sum of the tariff-jumping and offshoring gains:

$$\Pi^F - \Pi^{DX} = \gamma(w, w^* + t, f) + \mu(w + t^*, w^*) \quad \text{where:} \quad \gamma(w, w^* + t, f) \equiv \pi(w) - f - \pi(w^* + t) \quad (3)$$

The tariff-jumping gain function γ is identical to that in the last section. Hence the two tariffs have opposite effects on FDI: falls in the host country tariff tend to discourage it while falls in the parent country tariff tend to encourage it. If both tariffs are reduced in equal proportions the effect is ambiguous, and depends on the relative sizes of the two markets and the differences in wages and trade costs.

Faced with these theoretical ambiguities, it is natural to look at the empirical evidence for guidance on the relative importance of the two motives. At a purely descriptive level, Brainard (1997) and Markusen (2002, Chapter 1) note that foreign affiliates of U.S. firms export relatively small amounts of their output back to the U.S., between 13% and 15% depending on the year, with affiliates in Canada as a noteworthy but not unexpected exception. Turning to econometric evidence, Brainard (1997) finds that FDI is high in industry-country pairs with high transport costs and low plant scale economies, while international differences in relative factor abundance have little effect on FDI. All of this is consistent with the view that FDI is

primarily horizontal rather than vertical. In the same vein, Markusen (2001) finds evidence that bilateral flows of FDI are encouraged by similarities in market size and in relative endowments of skilled and unskilled labour between countries, and interprets this as evidence against the importance of vertical FDI. On the other hand, Yeaple (2003b), while confirming the importance of the proximity-concentration motive, finds that, other things equal, U.S. multinationals invest more in skill-scarce countries than in skill-abundant countries in the least skilled-labour-intensive industries. He notes that this is consistent with a comparative advantage or vertical view of FDI. We can conclude that the case against the vertical FDI model is not proven but that the case in favour is not strong enough to explain the paradox of trade liberalisation coexisting with FDI growth.

4 Export-Platform FDI

In this section I consider a different way to resolve the paradox noted in Section 2, drawing on my work on export-platform FDI in Neary (2002).⁸ Suppose that the model is identical to that in Section 2, except that the host country is one of two identical countries in a potential economic union. The previous analysis still holds when intra-union barriers are equal to t , with the added implication that the FDI option implies establishing two plants, one in each union country.

Now, suppose that intra-union barriers are reduced to a level τ which is less than the common external trade cost t . Clearly this does not affect the profits from exporting to both countries from the firm's country of origin: these continue to equal $\pi(t)$ for each destination country as in Section 2, so the total profits from exporting, Π^X , equals twice this, $2\pi(t)$. However, the profits from locating a plant in one of the union countries are now greater: in addition to the net profits of serving the host-country market $\pi(0) - f$, there is an additional gain from serving the partner-country market, $\pi(\tau)$. Hence the total profits from FDI, Π^F , equal $\pi(0) + \pi(\tau) - f$, and the relative attractiveness of FDI is now:

$$\Pi^F - \Pi^X = \gamma(t, f) + \chi(t, \tau) \quad \text{where:} \quad \chi\left(\underset{+}{t}, \underset{-}{\tau}\right) \equiv \pi(\tau) - \pi(t) \quad (4)$$

Now there are two sources of gain from FDI. As before, $\gamma(t, f) \equiv \pi(0) - f - \pi(t)$ is the trade-cost-jumping gain as the host-country market is served from a local plant rather than from exports. In addition, $\chi(t, \tau)$ denotes the gain from serving the partner-country market facing the intra-union trade cost τ rather than the higher common external trade cost t . We can call this the *export-platform gain*. Two implications are immediate. First, FDI is now more attractive relative to exporting. Unlike the trade-cost-jumping gain which

⁸For other discussions of this topic, see Motta and Norman (1996), Yeaple (2003a), Ekholm, Forslid and Markusen (2003) and Grossman, Helpman and Szeidl (2004).

can be positive or negative, the export-platform gain is always non-negative. This implies for example that the size of a host-country market is not a crucial determinant of locating a new plant there. What matters is the size of the trade-cost-adjusted market which can be served from that plant. Second, and crucial for the theme of this paper, the export-platform gain is decreasing in the intra-union trade cost τ whereas both it and the trade-cost-jumping gain are increasing in the external trade cost t . Hence, in striking contrast to the simple horizontal FDI model of Section 2, intra-bloc trade liberalisation tends to encourage FDI, once we recognize the importance of the export-platform motive. Note that a three-country perspective is essential for this result: the external trade cost t continues to exert a positive effect on FDI; indeed, it enhances both the trade-cost-jumping and the export-platform motives.

Some further implications of this model can be deduced from Figure 3. The region labelled "*FDI*" in Figure 1 is one where high trade costs and low fixed costs justify building a plant in both of the union countries. Such a region still exists when τ is less than t , provided τ is strictly positive: in Figure 3 it is denoted "*FDI* (2)". However, it is reduced in size by the emergence of a new region labelled "*FDI* (1)": this gives the combinations of f and t where (for given τ) it is profitable to establish a plant in only one market serving both. Thus, as τ falls, the export platform motive favours plant consolidation by firms already engaged in FDI with multiple plants. The new region also expands at the expense of the X region, as implied by equation (4). Finally, and more surprisingly, it expands at the expense of the O region, since export-platform FDI is profitable for any fixed cost less than $\pi(0) + \pi(\tau)$ and greater than $\pi(0)$, provided the common external trade cost is above the threshold level \tilde{t} . Thus, as τ falls, the export platform motive not only favours FDI over exporting but it also (for parameter values in the shaded region) favours FDI over not serving the market at all. A final implication of the model is that the same firm engages in both exports and FDI, albeit not across the same frontier: the firm engages in FDI into the host country and also in exports from there to the partner country. Hence exports and FDI become complements rather than substitutes in the aggregate data.

This approach avoids many of the criticisms of the simple horizontal FDI model of Section 2. How does it relate to the empirical evidence? It is clearly consistent with the experience of the European Union in the 1990s, when the dismantling of non-tariff barriers on internal trade under the Single Market programme coincided with a huge inflow of extra-EU FDI especially from the U.S. The Irish economy in particular exemplified this pattern, with many firms locating giant plants far larger than need to service the Irish market, causing both FDI and exports to rise in tandem. The model therefore reinforces the view that FDI, attracted by the deepening of the EU Single Market, was a major cause of the "Celtic Tiger" boom, which saw double-digit growth rates in GDP for much of the 1990s. (See Barry (1999).) As for econometric evidence, most of the literature looks only at bilateral flows of FDI, but two recent papers present evidence

that supports the importance of the export-platform motive. Head and Mayer (2004) study Japanese FDI in European regions and find that it is encouraged by market potential, which they measure using both host-region GDP and the GDP's of adjacent regions. Blonigen et al. (2004) use spatial econometric techniques to measure distance effects beyond adjacent countries. They find evidence against agglomeration effects: higher U.S. FDI in neighboring countries reduces the amount of U.S. FDI into individual European countries; but in favour of the export-platform hypothesis: higher GDP in neighboring countries increases U.S. FDI.

As for the prediction that falling intra-bloc trade barriers should encourage plant consolidation, Pavelin and Barry (2005) address this in a study of the geographical diversification of 290 leading European firms between 1987 and 1993. Contrary to the hypothesis, they find that diversification increased substantially over this period, which roughly coincides with the deepening of the Single Market. However, as they note, their sample does not include many U.S. firms; it covers only very large firms, which are likely to be multi-product and even multi-industry; and they measure the geographical diversification of a firm by the number of countries in which it operates rather than by the variance of its production or sales across countries. By contrast, evidence in favour of the plant consolidation hypothesis can be found in Belderbos (1997, pp. 33, 77), who notes that Japanese electronics firms followed a strategy of locating VCR plants in many EC countries in the early 1980s, but divested many of these multiple plants and concentrated on "best" locations in the late 1980s and early 1990s. Clearly more work is needed on this topic.

5 Cross-Border Mergers and Acquisitions

So far, I have assumed that all FDI involves constructing a new plant in a foreign country, the so-called "greenfield" case. However, the most important form of FDI in reality is not the greenfield type but rather cross-border mergers and acquisitions (M&As), where a foreign firm purchases an existing firm in the host country. UNCTAD (2000) documents the importance of M&As in the world economy, noting that they grew rapidly in the 1990s, both absolutely and relative to greenfield FDI: the share of M&As in the total value of world FDI exceeded 80% in 1999, and cross-border M&As were particularly important in FDI flows between developed countries.⁹ By contrast, the enormous scholarly literature on FDI has concentrated on the greenfield case. In this section I review some recent work which attempts to redress this balance, and discuss its implications for the effects of trade costs on FDI.

In discussing M&As, it seems very desirable, for the first time, to adopt an explicitly oligopolistic approach. One reason is empirical: according to UNCTAD (2000), the principal difference between greenfield

⁹UNCTAD caution that the data on greenfield and M&A FDI are not fully comparable. The data on total FDI come from balance of payments statistics, those on cross-border M&As from Thompson, a consultancy group, so the latter is not a proper subset of the former. See Head and Ries (2005) for further discussion. However, the importance of cross-border M&As is uncontroversial.

and M&A FDI is a persistent concentration effect: markets with more mergers and acquisitions are more concentrated. A second reason is conceptual. The theory of industrial organisation emphasizes two broad motives for M&As: a strategic motive, as acquiring firms gain from a reduction in competition, and an efficiency motive, as acquisitions lead to synergies through internal technology transfer, economies of scale, and coordination of production and marketing decisions. Of these, the former can only be considered in a model which explicitly allows for perceived interdependence between firms, and both suggest that firms are large relative to the markets in which they operate. Hence in this section I present a model of cross-border mergers in oligopolistic markets introduced in Neary (2003, 2004).¹⁰ This concentrates on the strategic motive for M&As. As we shall see, it also throws light on the theme of the paper, how trade liberalisation can encourage rather than discourage FDI.

The setting is an industry with n home firms, all with unit cost c , and n^* foreign firms, all with unit cost c^* . Assume that the two countries have a completely integrated market. Write the total profits of a home firm as a function of all these variables: $\pi(c, c^*; n, n^*)$. (I abstract from firm-level fixed costs, since they would provide a trivial justification for mergers.) It is plausible to assume that profits are decreasing in own costs c , increasing in foreign costs c^* , and that the own effect dominates the cross-effect. Hence I will confine attention to this case.¹¹ In the absence of foreign rivals, the threshold level of home costs consistent with breaking even is \tilde{c} , defined as the price which drives demand to zero.¹² With active foreign rivals, the threshold level of home costs consistent with breaking even is defined as an implicit function of the rivals' costs by $\pi(c, c^*; n, n^*) = 0$. Given my assumptions on the derivatives of π , this yields a locus in $\{c, c^*\}$ space for given numbers of home and foreign firms, which is upward-sloping though less steeply-sloped than the 45° line, as illustrated in Figure 4. Hence, with symmetric assumptions concerning foreign firms, the diagram is divided into four regions as shown. In region O it not profitable for any firms to serve the market; in regions F and H only firms from foreign or home respectively are profitable; while in region HF both types of firm are profitable. The latter region, or “cone of diversification” is the most interesting case, since low- and high-cost firms coexist there. The pattern of trade is clear: at all points above the 45° line, home firms are at a cost disadvantage and so are smaller in size than foreign firms in the same sector. Hence there is a presumption that the home country is the importer for points above the 45° line, and this holds exactly if

¹⁰Cross-border mergers have been studied in models of large-group monopolistic competition by Barba Navaretti and Venables (2004, Chapter 3) and Nocke and Yeaple (2004), and in a model emphasising the market for corporate control but abstracting from trade flows (and hence from trade costs) by Head and Ries (2005). My approach is closer to the small but rapidly growing literature on cross-border mergers in oligopolistic markets exemplified by van Long and Voudsen (1995), Falvey (1998), Horn and Persson (2001), Bertrand and Zitouna (2003) and Ferrett (2004).

¹¹Exact conditions are given in Neary (2002, Appendix A.2). As often in oligopoly theory, the properties hold provided that demand is not too convex.

¹²The threshold \tilde{c} is independent of n as well as of c^* and n^* . This follows from the typical home firm's first-order condition in the absence of any foreign firms, which is $p(ny) + yp'(ny) = c$. If the demand function has constant elasticity \tilde{c} is infinite, while if it is linear \tilde{c} equals its intercept.

the two countries are symmetric (with equal market size, and with the same number of home firms as foreign firms in each industry). In the perfectly competitive case where entry of new firms into the industry was free, the HF region would collapse to the 45° line, as only low-cost firms would survive, and so specialisation patterns would reflect perfectly the two countries' comparative advantage.

The cone HF is also the only region in which mergers may take place. In the F and H regions all firms are the same, and a classic result in industrial organisation, due to Salant, Switzer and Reynolds (1983), states that bilateral mergers are not profitable in an industry with more than two identical firms. Salant et al. confined attention to cases where firms have identical costs, whereas Neary (2004) shows that provided costs are sufficiently different then bilateral mergers are indeed profitable. To see why, define the gain from a takeover of a home firm by a foreign firm as G_{FH} :¹³

$$G_{FH}(c, c^*; n, n^*) = \Delta\pi^*(c, c^*; n, n^*) - \pi(c, c^*; n, n^*) \quad (5)$$

This consists of two parts. The first is the change in the acquiring firm's profits as the market becomes more concentrated following a takeover which reduces the number of home firms from n to $n-1$: $\Delta\pi^*(c, c^*; n, n^*) \equiv \pi^*(c, c^*; n-1, n^*) - \pi^*(c, c^*; n, n^*)$. Since oligopoly profits are decreasing in the number of firms, this is always positive. The second is the initial profits of the target home firm, $\pi(c, c^*; n, n^*)$, which is the amount that it must be paid to persuade it to sell. Along the boundary between the F and HF regions in Figure 4 both of these terms are zero, since the home firm's output is zero, so both its own profits, and its impact on other firms' profits if it ceases production, are zero. As c^* falls, the changes in the two terms can be deduced with the help of Figure 5, adapted from Leahy and Neary (2005). This shows how the two terms change as the cost competitiveness of the home firm rises, holding c , n and n^* constant.¹⁴ Both increase, but the $\Delta\pi^*$ term increases more rapidly at first. Why? Because the home firm is initially very small, and its profits increase with the square of its output. By contrast, eliminating a small firm raises industry price and so increases the foreign firm's price-cost margin, giving it additional profits on every unit it sells. As the home firm's cost competitiveness continues to rise (moving to the right of point Q in Figure 5), its initial profits increase more rapidly, whereas the change in the foreign firm's profits from eliminating it rise approximately linearly. (See the Appendix, Section 7.2, for details.) At some point, denoted R in Figure 5, the two curves intersect, so the gain from taking over the home firm becomes zero: the foreign firm would benefit greatly from the reduction in competition, but cannot afford to buy the by-now relatively large home firm. Finally,

¹³Strictly speaking, this is a myopic gain, since it does not take account of the effect of one takeover on the profitability of further takeovers between the remaining $n + n^* - 1$ firms. Neary (2004) shows that similar results can be obtained when the model is extended to allow firms to have forward-looking expectations of future takeovers.

¹⁴This figure is drawn on the assumption that the demand function is linear with intercept a . The home firm's cost competitiveness is measured by $a - c$.

from Salant et al. we know that the gain is strictly negative when the home firm's cost competitiveness rises to the level of the foreign firm, $a - c^*$.

All this implies that there is a range, denoted QR in Figure 5, within which home firms make positive profits but are vulnerable to acquisition by foreign firms. This corresponds to the region indicated in Figure 6 where, relative to Figure 4, the F region expands relative to the FH region. (Symmetric reasoning with the roles of the two countries reversed implies that there is a second region in Figure 6 where the H region expands relative to the FH region, as low-cost home firms acquire high-cost foreign ones.) Now consider the implications of this analysis for the effects of trade liberalisation. Starting in autarky, home firms face no competition, so they produce positive levels of output in that sub-region of F where their cost is less than the threshold level \tilde{c} . Trade liberalisation alone eliminates home firms in region F , and (in the case of symmetric countries) leads all foreign firms above the 45° line to export into the home market. If restrictions on cross-border M&As are also lifted, then as we have seen the region within which home firms can survive compresses even further. This gives the first prediction of the model: cross-border mergers take place in the same direction as trade flows, and they serve to move the pattern of international specialisation closer to what would prevail under perfect competition. In both senses cross-border mergers can be viewed as “instruments of comparative advantage”.

The second prediction of the model follows from the fact that G_{FH} is decreasing in the number of home firms n . A takeover of one home firm causes both curves in Figure 5 to pivot upwards around the point Q , but the gain to a takeover $\Delta\pi^*$ rises by more than the cost π . Intuitively, the potential acquiring firm is larger with $n - 1$ rivals than with n , so its gain in profits (which equal the increase in its price-cost margin times its total output) rises faster than the profits of the home firm (which are proportional to the square of its output). Hence the acquisition of one home firm increases the incentive for another to be acquired, so that mergers are likely to come in waves, with one bilateral acquisition prompting another until all the small and relatively inefficient firms in the sector have been acquired.

Finally, while the analysis so far has considered the incentives for cross-border M&As when trade is fully liberalized and the markets are integrated, it can be extended to allow for intermediate tariff levels, assuming that the markets are segmented.¹⁵ The effects of a small reduction in an existing tariff on the gain from a merger must now take account of the change in the acquiring firm's profits and the initial profits of the target firm in both markets. Trade liberalisation increases the home firm's profits on its initial exports, and this in itself makes it a more expensive takeover target, so making cross-border M&As less likely. On the other hand, trade liberalisation also increases the foreign firm's profits from exporting as well as reducing both firms' profits in their home markets, both of which make cross-border M&As more likely. In the neighbourhood

¹⁵I am very grateful to Philippe Martin for detailed suggestions on this case. See also van Long and Vousden (1995).

of autarky the first effect does not arise and so at high trade costs a small amount of trade liberalisation unambiguously raises the likelihood of cross-border M&As.

Turning finally to empirical evidence, Brakman, Garretsen and van Marrewijk (2005) is the only paper to date which explicitly tests the “instruments of comparative advantage” theory of cross-border M&As. They consider data on cross-border M&As between five OECD countries in twenty sectors over the period 1980-2004 and find strong evidence of a role for comparative advantage. Specifically, acquiring firms in cross-border mergers and acquisitions come disproportionately from sectors which have a revealed comparative advantage, as measured by the standard Balassa index. They also find evidence that mergers are positively autocorrelated within sectors, consistent with the hypothesis that mergers occur in waves. Of course, these results may be consistent with other theories too, so further work is needed to test their robustness.

6 Conclusion

This paper has reviewed the theory of foreign direct investment, using as an organizing principle an apparent conflict between received theory and recent trends in the globalized world. Conventional wisdom holds that the bulk of FDI is horizontal rather than vertical, aimed at replicating production facilities abroad to improve access to foreign markets rather than breaking up the production process to benefit from lower production costs. This view is reinforced by most, though not all, of the empirical evidence. Given this, we should expect falls in trade costs to discourage FDI, since this is one of the predictions of the standard model of horizontal FDI, which emphasizes a proximity-concentration trade-off. However, this seems to conflict with the experience of the 1990s, when trade liberalisation and technological change led to dramatic reductions in trade costs yet FDI grew much faster than trade. Two possible resolutions to this paradox have been explored. First, horizontal FDI in trading blocs is encouraged by intra-bloc trade liberalisation, because foreign firms establish plants in one country as export platforms to serve the bloc as a whole. Second, cross-border mergers, which are quantitatively more important than greenfield FDI, are encouraged rather than discouraged by falling trade costs.

One broad conclusion which follows from the literature reviewed here is that the distinction between horizontal and vertical FDI is useful for pedagogic purposes but otherwise not very helpful. In practice most multinational corporations pursue what Yeaple (2003a), following United Nations (1998), calls “complex integration strategies”, which do not fit neatly into either the horizontal or vertical categories. Export-platform FDI as discussed in Section 4 is one example of such a strategy: though modelled here as purely horizontal in the sense that no physical intermediate inputs are traded within the firm, it implies that the simple proximity-concentration trade-off does not apply easily to a world with more than two countries.

Cross-border mergers and acquisitions are another example, implying that trade and FDI can move in the same direction, even on the part of a single firm. Clearly more analytical and empirical work is required to disentangle the relative importance of these different aspects of FDI.

7 Appendix

7.1 Operating Profits and Trade Costs

Write the operating profits of the foreign firm serving the home market as a function of its output y and the trade cost t :

$$\tilde{\pi}(y, t) = [p(y) - (c + t)]y \quad (6)$$

where $p(y)$ is the inverse demand function and c is the unit cost of production. The function used in the text, $\pi(t)$, is the value of $\tilde{\pi}(y, t)$ when y is chosen at the profit-maximising level:

$$\pi(t) \equiv \underset{y}{Max} [\tilde{\pi}(y, t)] \quad (7)$$

This exhibits a variant of Hotelling's Lemma applied to a firm with monopoly power:

$$\pi' \equiv \frac{d\pi}{dt} = \tilde{\pi}_y \frac{dy}{dt} + \tilde{\pi}_t = \tilde{\pi}_t = -y < 0 \quad (8)$$

where the envelope property follows from the first-order condition: $\tilde{\pi}_y = p + yp' - (c + t) = 0$. Totally differentiating the first-order condition gives: $\tilde{\pi}_{yy}dy + \tilde{\pi}_{yt}dt = 0$, which implies that output is a decreasing function of the trade cost:

$$\frac{dy}{dt} = -\frac{\tilde{\pi}_{yt}}{\tilde{\pi}_{yy}} = \frac{1}{\tilde{\pi}_{yy}} < 0 \quad (9)$$

where the negative sign follows from the second-order condition: $\tilde{\pi}_{yy} = 2p' + yp'' < 0$. This also implies that the boundaries between the X and FDI regions in Figures 1 and 3 and between the D and FDI regions in Figure 2 are concave.

7.2 Merger Gains

To evaluate the merger gain G_{FH} , we make use of the always-useful result that a firm's operating profits in any market are proportional to the square of output: $\pi = by^2$, which follows from the first-order condition $p - c = by$, where b is the inverse demand slope. (Note that this holds under all market structures and demand systems, though when demand is non-linear b depends on the sales of all firms.)

The two components of the gain from a merger depend on the initial outputs of a typical home and foreign firm, which can be written as follows (using $A \equiv a - c$ and $A^* \equiv a - c^*$ to denote each firm's cost

competitiveness, and $\bar{n} \equiv n + n^*$ to denote the total number of firms):

$$y(c, c^*; n, n^*) = \frac{(n^* + 1)A - n^*A^*}{b(\bar{n} + 1)} \quad y^*(c, c^*; n, n^*) = \frac{(n + 1)A^* - nA}{b(\bar{n} + 1)} \quad (10)$$

It follows immediately that π is quadratic in A . As for the change in the acquiring foreign firm's profits, this can be factorized as follows:

$$\Delta\pi^*(c, c^*; n, n^*) = b[y^*(c, c^*; n - 1, n^*) + y^*(c, c^*; n, n^*)][y^*(c, c^*; n - 1, n^*) - y^*(c, c^*; n, n^*)] \quad (11)$$

Direct calculations (following Neary (2004, Appendix)) show that:

$$y^*(c, c^*; n - 1, n^*) - y^*(c, c^*; n, n^*) = \frac{1}{\bar{n}}y(c, c^*; n, n^*) \quad (12)$$

$$y^*(c, c^*; n - 1, n^*) + y^*(c, c^*; n, n^*) = 2y^*(c, c^*; n, n^*) + \frac{1}{\bar{n}}y(c, c^*; n, n^*) \quad (13)$$

Differentiating these with respect to c and n gives the properties of Figure 5 discussed in the text.

References

- Anderson, James E. and Eric van Wincoop (2004): "Trade Costs," *Journal of Economic Literature*, 42, September, 691-751.
- Anderson, James E. and J. Peter Neary (2005): *Measuring the Restrictiveness of International Trade Policy*, forthcoming, Cambridge, Mass.: MIT Press.
- Barba Navaretti, Giorgio, Anthony J. Venables et al. (2004): *Multinational Firms in the World Economy*, Princeton: Princeton University Press.
- Barry, Frank (1999) (ed.): *Understanding Ireland's Economic Growth*, London: Macmillan.
- Belderbos, René and Leo Sleuwaegen (1998): "Tariff jumping DFI and export substitution: Japanese electronics firms in Europe," *International Journal of Industrial Organization*, 16:5, September, 601-638.
- Belderbos, René (1997): *Japanese Electronics Multinationals and Strategic Trade Policies*, Oxford: Oxford University Press.
- Bergstrand, Jeffrey H. and Peter Egger (2005): "A knowledge-and-physical-capital model of international trade, foreign direct investment, and outsourcing: Part I, developed countries," mimeo., University of Notre Dame, March.
- Bertrand, Olivier and Hamid Zitouna (2003): "Trade liberalization and industrial restructuring: The role of cross-border mergers and acquisitions," mimeo., Université de Paris I.
- Blonigen, Bruce A., Ronald B. Davies, Glen R. Waddell, and Helen Naughton (2004): "FDI in space: Spatial autoregressive relationships in foreign direct investment," mimeo., University of Oregon.
- Brainard, S. Lael (1993): "An empirical assessment of the factor proportions explanation of multi-national sales," Working Paper No. W4583, Cambridge Mass.: NBER, December.
- Brainard, S. Lael (1997): "An empirical assessment of the proximity-concentration tradeoff between multinational sales and trade," *American Economic Review* 87, 520-544.
- Brakman, Steven, Harry Garretsen and Charles van Marrewijk (2005): "Cross-border mergers and acquisitions: On revealed comparative advantage and merger waves," mimeo., University of Groningen.
- Carr, David, James R. Markusen, and Keith E. Maskus (2001), "Estimating the knowledge-capital model of the multinational enterprise," *American Economic Review*, 91, 691-708.
- Dunning, John H. (1973): "The determinants of international production," *Oxford Economic Papers*, New Series, 25:3, November, 289-336.
- Ekholm, Karolina, Rikard Forslid and James R. Markusen (2003): "Export-platform foreign direct investment," Discussion Paper No. 3823, London: CEPR.
- Falvey, Rod (1998): "Mergers in open economies," *The World Economy*, 21, 1061-1076.

Ferrett, Ben (2004): "Greenfield investment versus acquisition: Alternative modes of foreign expansion," mimeo, University of Nottingham.

Grossman, Gene, Elhanan Helpman and Adam Szeidl (2004): "Optimal integration strategies for the multinational firm," Working Paper No. 10189, Cambridge, Mass.: NBER.

Head, Keith and Thierry Mayer (2004): "Market potential and the location of Japanese investment in the European Union," *Review of Economics and Statistics*, 86:4, November, 959-972.

Head, Keith and John Ries (2005): "FDI as an outcome of the market for corporate control: Theory and evidence," mimeo., Sauder School of Business, University of British Columbia, May.

Helpman, Elhanan (1984): "A simple theory of international trade with multinational corporations," *Journal of Political Economy*, 92, 451-471.

Helpman, Elhanan, Marc Melitz and Stephen Yeaple (2004): "Exports versus FDI with heterogeneous firms," *American Economic Review*, 94, March, 300-316.

Horn, Henrik and Lars Persson (2001): "The equilibrium ownership of an international oligopoly," *Journal of International Economics*, 53, 307-333.

Jones, Ronald W. and J. Peter Neary (1984): "The positive theory of international trade," in R.W. Jones and P.B. Kenen (eds.): *Handbook of International Economics*, Volume 1, Amsterdam: North-Holland, 1-62.

Leahy, Dermot and J. Peter Neary (2005): "Strategic and efficiency motives for cross-border mergers," mimeo., University College Dublin (in preparation).

Long, Ngo Van and Neil Vousden (1995): "The effects of trade liberalization on cost-reducing horizontal mergers," *Review of International Economics*, 3, 141-155.

Markusen, James R. (1983): "Factor movements and commodity trade as complements," *Journal of International Economics*, 14, 341-356.

Markusen, James R. (2002): *Multinational Firms and the Theory of International Trade*, Cambridge, Mass.: MIT Press.

Motta, Massimo and George Norman (1996): "Does economic integration cause foreign direct investment?," *International Economic Review*, 37(4), 757-783.

Mundell, Robert (1957): "International trade and factor mobility," *American Economic Review*, 47:3, 321-335.

Neary, J. Peter (1995): "Factor mobility and international trade," *Canadian Journal of Economics*, 28, November 1995, S4-S23.

Neary, J. Peter (2002): "Foreign direct investment and the single market," *The Manchester School*, 70:3, June, 291-314.

Neary, J. Peter (2003): "Globalization and market structure," *Journal of the European Economic Association*, 1:2-3, April-May, 245-271.

Neary, J. Peter (2004): "Cross-border mergers as instruments of comparative advantage," Discussion Paper No. 4325, London: CEPR.

Neary, J. Peter and Cormac Ó Gráda (1991) "Protection, economic war and structural change: The 1930s in Ireland," *Irish Historical Studies*, 27, May, 250-266.

Nocke, Volker and Stephen Yeaple (2004): "Mergers and the composition of international commerce," mimeo., University of Pennsylvania.

Pavelin, Stephen and Frank Barry (2005): "The Single Market and the geographical diversification of leading firms in the EU," *The Economic and Social Review*, 36: 1, Spring, 1-17.

Salant, Stephen, S. Switzer and R. Reynolds (1983): "Losses due to merger: The effects of an exogenous change in industry structure on Cournot-Nash equilibrium," *Quarterly Journal of Economics*, 98, 185-199.

Smith, Alasdair (1987): "Strategic investment, multinational corporations and trade policy," *European Economic Review*, 31, 89-96.

UNCTAD (2000): *World Investment Report: Cross-Border Mergers and Acquisitions and Development*, New York and Geneva: United Nations.

United Nations (1998): *World Investment Report*, Centre for Transnational Corporations, New York: United Nations.

Yeaple, Stephen (2003a): "The complex integration strategies of multinational firms and cross-country dependencies in the structure of foreign direct investment," *Journal of International Economics*, 60:2, August, 293-314.

Yeaple, Stephen (2003b): "The role of skill endowments in the structure of U.S. outward foreign direct investment," *Review of Economics and Statistics*, 85:3, August, 726-734.

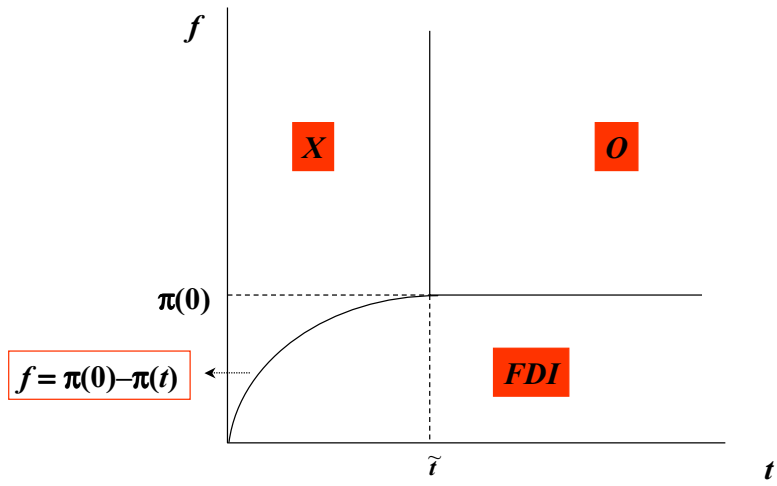


Fig. 1: The Proximity-Concentration Trade-Off I: The Trade-Cost-Jumping Motive

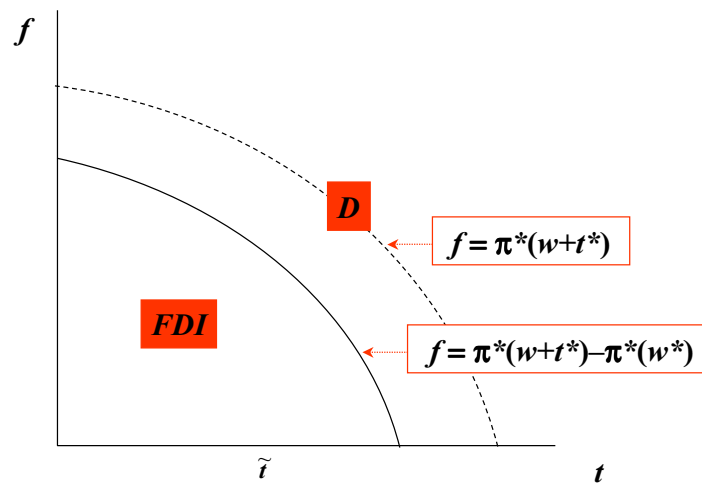
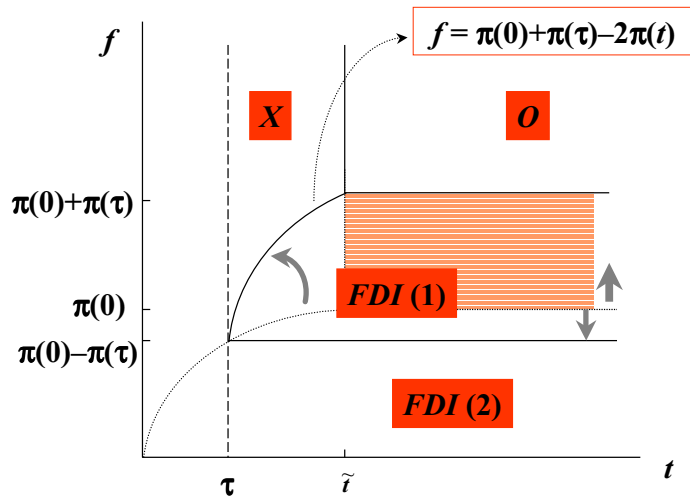
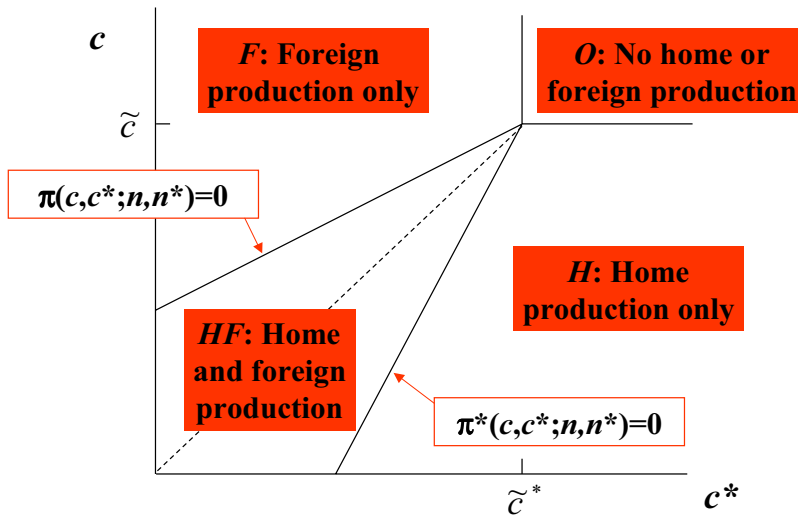


Fig. 2: Vertical FDI



**Fig. 3: The Proximity-Concentration Trade-Off II:
External Trade-Cost-Jumping + Export-Platform Motives**



**Fig. 4: Equilibrium Production Patterns
in Free Trade without FDI**

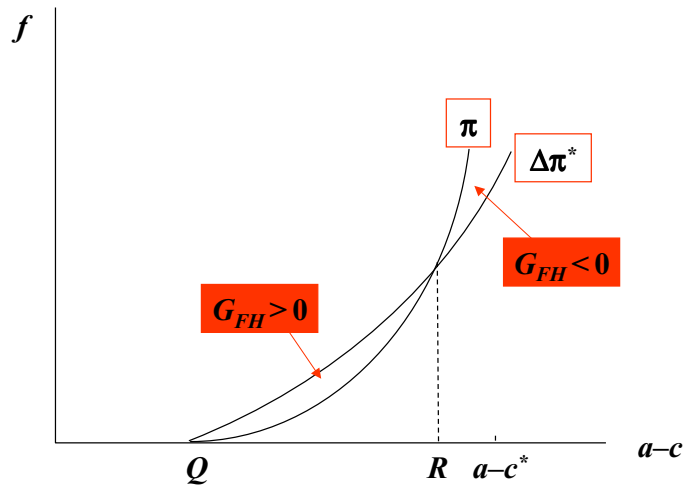


Fig. 5: The Components of Gain from a Cross-Border Merger

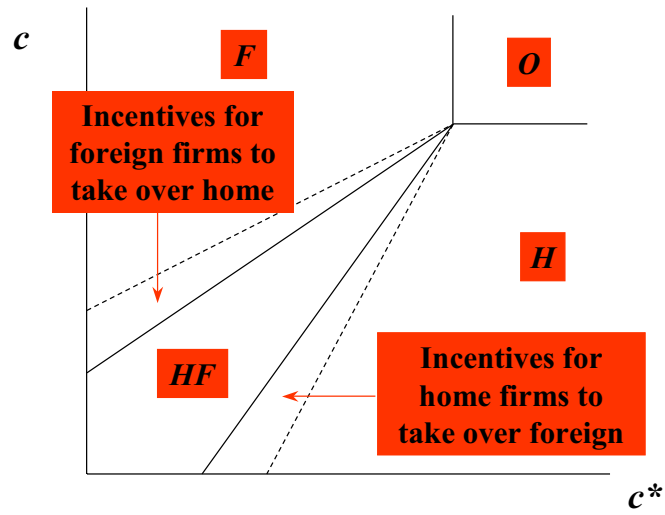


Fig. 6: Cross-Border Merger Incentives