SCALE ECONOMIES, INTRA-INDUSTRY TRADE AND INDUSTRY LOCATION IN THE "NEW TRADE THEORY"

Trinity Economic Paper Series
Technical Paper No. 95/4
JEL Classification: F12, F15

Marius Brülhart
Department of Economics
Trinity College
Dublin 2.
email: mbrulhrt@mail.tcd.ie

Abstract

This paper exposes some common misinterpretations of the "new trade theory". First, the view that high scale economies give rise to high levels of intra-industry trade is challenged. It is shown that the monopolistic-competition trade model predicts a *negative* relationship between internal scale economies and intra-industry trade. Second, in spite of a common perception that the "new" theory explains ever growing levels of intra-industry trade in an integrating world economy, a scrutiny of the basic model indicates that reduced distance costs result in *lower* intra-industry trade. However, if temporary re-location rigidities are considered, integration entails an initial surge of intra-industry trade, which eventually withers away, when the centripetal forces towards inter-industry specialisation take over. This might contribute towards an explanation for observed reversals of intra-industry trade growth among industrial countries.

Acknowledgements

Special thanks to Kevin Hanrahan, Dermot McAleese and Paul Walsh for helpful suggestions and to Paddy Waldron for essential computing guidance. Earlier versions of this paper have been presented at the SPES Workshop, University of Nottingham, September 1994, and at the "Journées AFSE" in Nantes, June 1995.

Introduction

The "new trade theory" (NTT), conceived in the late 1970s, has fundamentally changed the way economists think about international trade flows. Models in this mould consider economies of scale and product differentiation, not different endowments of countries, as determinants of international specialisation and trade. The primary appeal of these models lies in the fact that they manage to explain certain features of modern trading patterns which do not accord with neo-classical theory. Trade among developed countries with relatively similar endowments has been the fastest growing component of global trade flows throughout the post-War period. This phenomenon does not conform with the neo-classical paradigm, whereby trading opportunities are greater the bigger are the differences between the productive endowments of countries. Furthermore, a large proportion of trade among industrialised nations consists of two-way flows of very similar products. Such intra-industry trade (IIT) is also difficult to accommodate within the neo-classical models, which predict that countries specialise in different sectors according to their comparative advantage and thus exchange different products.

The main ingredient of the "new" trade models was the assumption of increasing returns, which give rise to imperfectly competitive markets. Imperfect competition, in turn, made possible the theoretical explanation of IIT. Subsequently, numerous authors have reversed this sequence of reasoning and concluded that IIT must be a symptom of prevailing scale economies. It will be shown in this paper that such an interpretation of the mainstream "new" trade model (no longer a contradiction in terms) is wrong. In the standard model, the relationship between scale economies and IIT is discontinuous, since very *high* levels of scale economies are associated with *low* levels of IIT. This is not academic nit-picking, since via scale economies IIT has come to be interpreted as an indicator of imperfectly competitive market structure.[1] The departure from

[1]The link between IIT and imperfect competition has been questioned for two other reasons. First, some authors have dismissed IIT as a "statistical artefact", because of the arbitrary definition of industries in trade statistics (see Finger, 1975). Second, it has been shown that IIT can also be explained by the neo-classical model, if one allows for technological differences among countries (Davis, 1993; Bhagwati and Davis, 1994).

perfect competition modelling entails a significant change in welfare and trade policy implications.[2]

Although the link between IIT and scale economies is ambiguous, IIT is a meaningful concept. It matters in two ways. First, it can serve as an indicator of the similarity of industrial structures among countries. The higher the level of IIT, the more similar is the composition of industry in two trading nations, *ceteris paribus*. In other words, if certain sectors exhibit high IIT in a group of countries, then these sectors are likely to be relatively dispersed over the whole area of this group. Second, IIT, particularly when applied to *changes* in trade patterns, is likely to be associated with relatively smooth trade-induced industrial adjustment.[3]

The spatial interpretation of IIT is of considerable relevance for integrating country groups such as the European Union or the North American Free Trade Area. Much of the debate surrounding these projects concerns the question whether integration would promote the concentration of industrial sectors in particular areas, at the expense of other areas. The two commonly identified locational poles are, at one extreme, labour abundant, often peripheral areas and, at the other extreme, capital abundant, mainly central locations. Some commentators fear an exodus of industrial sectors to low-cost regions, while others predict a further clustering of manufacturing in economic core regions.

Neither of these scenarios is compatible with IIT. Yet, trade flows within the European Union, for example, have long been increasingly intra-industry in nature. Recently, some signs of stagnating IIT growth have been diagnosed for a number of European countries.[4] Does this mean that re-location of industries is only now starting to gather steam?

An *empirical* answer to this question is attempted in Brülhart (1995). This paper scrutinises the predictions of the "new" *theories* for IIT and industry location among integrating countries. There exists a common conception that industry location in the "new" models is indeterminate, since it is independent of

---

[2]see Richardson (1989), Helpman and Krugman (1989), and Krugman (1993).

[3]see Brülhart (1994).

[4]see Globerman and Dean (1990), Greenaway and Hine (1991) and Brülhart and McAleese (1995). A similar reversal was detected for the United States by Leamer (1995).

factor requirements and endowments. Therefore, the "new" models are often seen to sit very easily with the observation of ever growing levels of IIT. A fall in IIT, however, suggests the presence of agglomeration forces which do eventually push integrating countries towards sectoral specialisation. This paper takes a close look at some of the most influential models and concludes that - against the common stylised interpretation of the NTT - the twin forces of scale economies and distance costs tend to work towards a concentration of industrial activity, and hence declining IIT levels, as barriers to trade are reduced.

The paper is divided into three sections. First, the standard "new" trade model is examined in terms of the predicted link between internal scale economies and IIT. Second, the locational predictions of the same model as well as models based on external scale economies are explored. Third, some alternative scenarios, which introduce factors other than scale economies, are scrutinised. The main results are summarised in the Conclusion.

I. Scale Economies and Intra-Industry Trade

The seminal works of the NTT focus on decreasing costs at plant level.[5] These models assume cost functions characterised by a fixed set-up cost and a constant marginal cost, such as:

$$l_i = a + bx_i; \; i = 1,...,n; \quad a, b > 0. \qquad \textbf{(1)}$$

where l_i is the total factor input used to produce quantity x of the ith good. Due to the costless differentiability of products and consumers' taste for variety assumed in these models, no two firms have an incentive to produce identical goods in terms of consumption characteristics.[6] Hence, the production

[5]Dixit and Stiglitz (1977), Krugman (1979, 1980, 1981), Lancaster (1980) Helpman and Krugman (1985).

[6]This modelling structure implies monotonically rising firm numbers - and therefore falling concentration ratios - as market size increases. Empirical work in industrial economics has shown that this relationship does not hold where differentiation (e.g. through advertising) is not costless (see Sutton (1991, pp. 59ff.)). Since the relationship between market size and firm numbers is a crucial feature of the "new" trade models, such empirical findings could limit the explanatory power of these models significantly. This question, however, is beyond the scope of this paper.

function for a particular good is equivalent to the cost function of a single firm, and the scale economies are internal to the individual plants.

The models of the NTT typically make a range of other restrictive assumptions. There is only one factor of production, usually called labour, and all goods are produced with the same (increasing returns) technology. Hence, there can be no comparative advantage among countries. All consumers share the same utility function, with symmetric preferences for all goods along a continuous spectrum of product characteristics. Monopolistic competition eliminates supernormal profits. These assumptions allow the models to reduce the array of dependent variables to one, namely the number of goods produced. While its restrictive hypotheses move the model far away from economic reality, such algebraic constructs manage to show formally that scale economies can give rise to international trade even where all countries share the same tastes, relative factor endowments and technologies.

Krugman's Two-Sector Model

The NTT suggests that all countries manufacture a number of product varieties which is proportional to the size of their total factor endowment (labour force), and that the international exchange of such similar goods shows up as IIT. However, in the crude form outlined above, these models provide no indication as to which industries locate where. Since they lump all manufacturing into a single differentiated sector, they cannot be used to analyse the relative importance of intra- and inter-industry trade.

Refining his basic model, Krugman (1980) has divided the spectrum of differentiated goods into two industries, whose products are demanded by different groups of consumers. He has assumed that relative demand for the goods of the two industries differs between the two countries in the model. Transport costs apply to cross-border transactions only. Two main results emerge from his formal analysis. First, each industry has an incentive to locate in the country with the bigger market for its products, and, second, if transport costs are not too low and scale economies not too pervasive, incomplete specialisation and the resulting IIT flows are the equilibrium outcome..

The mathematical elegance of the Krugman (1980) model makes it possible to evaluate the impact of various parameters on the level of IIT, even though this was not made explicit by the author. The model can be summarised quite simply.[7] The populations of both countries (country 2 is indicated by a star) consist of consumers purchasing only goods produced by firms in the *alpha* industry and of consumers desiring only products emanating from the *beta* industry (indicated by a tilde). Apart from this distinction, all consumers share the same type of utility functions, where the horizontally differentiated products of their respective industries enter symmetrically into demand, expressed as follows:

$$U = \left(\sum_i c_i^q \right)^{\frac{1}{q}}; \qquad \tilde{U} = \left(\sum_j \tilde{c}_j^q \right)^{\frac{1}{q}}; \qquad 0 < q < 1, \qquad \textbf{(2)}$$

where c_i is consumption of the *i*th good of the *alpha* industry. These utility functions result in a constant elasticity of substitution $1/(1-q)$.[8] The two countries are of equal size $\overline{L}$, with "mirror-image" markets for goods of the two industries:

$$L + \tilde{L} = L^* + \tilde{L}^* = \overline{L}; \qquad L = f\overline{L}; \qquad L^* = (1-f)\overline{L}; \qquad 0 < f < 1. \qquad \textbf{(3)}$$

All goods within an industry are produced with identical cost functions, of the type described by equation (1). Trade between the two countries is subject to "iceberg" transport costs, where only a fraction t of any good shipped reaches its destination. Therefore, consumer prices for imported goods (c.i.f.) are higher than prices for domestic products (f.o.b.). Due to the perfect symmetry of the model, the ratio of total demand by domestic residents for each foreign product to demand for each domestic product does not depend on relative prices and wages and is given by:

$$s = t^{\frac{q}{1-q}}; \qquad 0 < t < 1. \qquad \textbf{(4)}$$

Since both t and q take values between zero and one, $0 < s < 1$.

Total expenditure on goods of each industry can be expressed as the sum of domestic and foreign purchases:

---

[7]For a complete specification of the two-sector model, see Krugman (1980), p. 955 ff.

[8]For the formal derivation of this result, see Krugman (1981), p. 972.

$$npx = \frac{n}{n + sn^*} wL + \frac{sn}{sn + n^*} wL^*, \text{ and: } n^* px = \frac{sn^*}{n + sn^*} wL + \frac{n^*}{sn + n^*} wL^*. \qquad \textbf{(5)}$$

Since the price p and output x of each good are equal for all varieties in the two industries and countries, it is possible to derive the relative number of varieties produced in each country:

$$\frac{n}{n^*} = \frac{\frac{L}{L^*} - s}{1 - s\frac{L}{L^*}}; \qquad s < \frac{L}{L^*} < \frac{1}{s}, \qquad \textbf{(6)}$$

where n (n^*) is the number of domestically (foreign) produced goods within one industry. As a result of the "mirror image" assumption, the ratio (n/n^*) of one industry equals the inverse of the second industry's ratio. The inequality qualifying equation (6) defines the range of relative market sizes within which specialisation is incomplete, in the sense that neither n nor n^* is zero.

Manipulations of Krugman's Model

This is as far as Krugman (1980) has pursued the analysis. It is possible, however, to determine trade flows and the levels of IIT generated in this framework by simple manipulation of his model. As pointed out above, prices and output are identical for both industries and for all varieties, due to the perfect symmetry in demand and production. Hence, the ratio of country 1's total value of exports to the total value of its imports within, for instance, the *alpha* industry is determined as follows:

$$\begin{aligned} f &= \sum_{i=1}^{n} X_i \,/ \sum_{i=n+1}^{n+n^*} M_i\,. \\ &= \left(\frac{sn}{sn+n^*} L^*\right) / \left(\frac{sn^*}{n+sn^*} L\right). \\ &= \frac{L^*}{L}\frac{sn}{sn^*}\frac{n+sn^*}{sn+n^*} = \frac{n}{n^*}.^{9} \end{aligned} \qquad \textbf{(7)}$$

[9]For the last step see Krugman (1980, p. 957), who showed that $L/L^* = (n+sn^*)/(sn+n^*)$. If the model were extended so as to allow for variations in the scale of production as well as in the number of product varieties, the basic conclusions on the pattern of IIT would remain unaffected.

This ratio is sufficient for the calculation of IIT indices. Based on the formula developed by Finger and DeRosa (1979, p. 214), according to which IIT = (2*{min *X*, *M*}) / (*X* + *M*), we can define IIT as follows:

$$\text{IIT} = \begin{cases} \dfrac{2f}{1+f}, & \text{for } 0 < f < 1. \\ 1, & \text{for } f = 1. \\ \dfrac{2(1/f)}{1+(1/f)}, & \text{for } f > 1. \end{cases} \qquad \textbf{(8)}$$

The combination of equations (4) to (8) permits us to analyse equilibrium IIT levels accompanying variations in the exogenous parameters of the models. We shall examine the effect of changes in three parameters: relative market sizes (L/L^*), equilibrium scale economies (q) and transport costs ($1-t$). We illustrate the resulting IIT patterns through numerical simulations. Algebraic proofs of the detected relationships are given in the Appendix.

First, we plot IIT against the ratio of market sizes (L/L^*). If the market sizes of the two countries are not too different, namely if their ratio lies in the interval (s, $1/s$), incomplete specialisation and the resulting IIT occurs. This is illustrated in Figure 1. The diagram depicts IIT in the *alpha* industry, but, due to the symmetry of the model, the levels of IIT are identical in the *beta* industry. This figure displays the positive relationship between IIT and demand similarity predicted by the NTT and tested empirically by numerous researchers with considerable success.[10]

Second, since s is a function of transport costs ($1-t$) and of the demand parameter q (equation (4)), we can also depict IIT as a function of either of these variables, while holding everything else constant. Figure 2 shows the relationship between IIT and q, derived from equations (4) to (8), given various relative market sizes L/L^*. In equilibrium, q is related strictly negatively to (the potential for) scale economies, because q is an inverse indicator of consumers' taste for variety.[11] It can be seen in Figure 2 that a non-positive relationship exists between equilibrium scale economies and IIT. This holds for all scenarios

[10]Two recent examples are Lundberg (1992) and Hirschberg *et al*. (1993).

[11]The ratio $q/(1-q)$, in equilibrium, is equivalent to the ratio of variable to fixed costs bx/a (see Krugman, 1980, p. 957).

with the exception of perfectly equal market sizes, where IIT equals one throughout.

The intersections of the IIT paths with the horizontal axis in Figure 2 occur where $L/L^{*}=s$, below which all production of *alpha* goods is concentrated in country 2 (home to L^{*}), resulting in pure *inter*-industry trade flows. These 'break' points mark the frontier between the areas of complete and incomplete specialisation already detected in Figure 1.

It is interesting to note that, despite this result of the basic model underlying the NTT, most econometric analyses have worked with the null hypothesis of a strictly *positive* relationship between IIT and scale economies. Clark (1993, p. 341) commented that his empirical investigation resulted in a "failure to establish the theoretically expected positive relationship between intra-industry trade and scale economies", and Krugman (1994, p. 17) remarked that "the study of intra-industry trade has come to be regarded as one way of testing the importance of scale economies". Considering the predictions of the Krugman (1980) model, however, such inconclusive econometric results should come as no surprise.

The discontinuous relationship between IIT and scale economies has previously been noticed by several authors. Ethier (1982a, p. 390) stated that "although the existence of internal scale economies and product differentiation are essential to the theory, the degree of such phenomena need not be an essential determinant of the degree of intra-industry trade". Greenaway and Milner (1986, p. 137f.) have interpreted the empirical results in a way that is compatible with the NTT: "(The econometric) results give support to the view that scale economies and IIT are not continuously related; where the minimum efficient scale of production is large relative to market size, standardisation rather than differentiation is likely to be found." Tybout (1993, p. 441) pointed out that most empirical studies of IIT are "only loosely motivated by theory".

II. Scale Economies, Industry Location and Integration

We have shown that, given a certain level of trade barriers, the forces towards inter-industry specialisation grow with the importance of scale economies. In a cross-sectional study we would thus expect high locational concentration, and low IIT, in industries subject to high scale economies. It is now important to investigate how this configuration is affected by a change in trade barriers. What are the predictions of the NTT for a time-series study of IIT and industry location among integrating countries? We investigate this issue in two modelling frameworks, one of which assumes scale economies to be internal to individual firms, while the other considers external economies.

Internal Scale Economies

The third relationship we derive from the Krugman (1980) model determines the equilibrium IIT path accompanying economic integration. If q is held constant and t varies instead, a non-positive relationship emerges between t and IIT for all scenarios where market sizes differ between the two countries. Since $(1-t)$ measures 'iceberg' transport costs, Figure 3 indicates that IIT is non-negatively related to trade barriers. This model thus predicts that IIT will fall with progressing integration among countries. Again, the intersections of the IIT lines with the horizontal axis are determined by $L/L^{*}=s$.

It is only if we allow for zero transport costs ($t = 1$) - a totally unrealistic scenario - that we obtain the much discussed indeterminacy of industry location.[12] Harrigan (1994) has demonstrated that, in a world without any barriers to trade, the degree of scale economies only affects trade volumes but not the proportion of IIT.

Is it valid to conclude from Figure 3 that the model characterised by internal scale economies and monopolistic competition predicts IIT to fall as economic integration proceeds? Helpman and Krugman (1985, p. 38f.) have warned that "there are major potential pitfalls in mixing static and dynamic analysis", but asserted that their "static models can proxy for (...) dynamic effects". Figure 3 shows four *sets* of long-term equilibria. A gradual fall in

[12]This can be verified by setting $t = 1$ in Appendix equation (15), which yields a zero denominator.

international transport costs over time, *ceteris paribus*, would strictly speaking only be accompanied by one of the depicted IIT paths if relocation of production were instantaneous and costless. Yet, even if we allow for some transitory adjustment rigidities, the profit-seeking firms in the model would eventually increase concentration in the larger market area with each step of economic integration. This result contrasts with the well documented empirical finding of a negative relationship between IIT and proxy measures of transport costs or economic distance.[13]

The IIT paths depicted in Figure 3 do not reveal total trade volumes. It is apparent from equation (4) that falling transport costs increase demand for foreign products. Hence, the declining IIT levels, which accompany falling transport costs in a move rightwards along the horizontal axis of Figure 3, are based on ever higher absolute trade volumes. In other words, the expansion of trade resulting from trade liberalisation is strictly inter-industry in nature, and, assuming a smooth transition between equilibria, marginal IIT[14] is zero throughout. Trade liberalisation, in this model, always leads to an increase in *inter*-industry specialisation.

Such trade models based on internal scale economies provide an elegant formal framework for the explanation of 'non-classical' trade flows and produce some plausible predictions. However, they are based on many highly unrealistic assumptions. We limit our critical assessment to the treatment of scale economies such as described by equation (1), which is unrealistic for two reasons.

First, infinitely falling average costs are not compatible with the concept of a 'minimum efficient scale' (MES), used widely in industrial economics. In microeconomics, total costs are commonly modelled as a cubic function of production, which results in parabolic average costs schedules with one minimum. In line with the prevailing perception among industrial economists, Gomory and Baumol (1992, p. 7) have stated that "the empirical evidence indicates that the case of flat bottomed average cost curves is common". The models outlined above would only be appropriate where the domestic markets are smaller than the MES of a single plant and the combined world market is

[13] Recent examples are Balassa and Bauwens (1987), Lundberg (1992), Clark (1993).
[14] see Brülhart (1994).

smaller or equal to the MES. This is ultimately a subject for empirical investigation. It seems reasonable to expect, though, that such a configuration is the exception rather than the rule in real-world economies.

Second, the NTT generally neglects external economies. Whether increasing returns occur mainly at the firm level or within the industrial sector as a whole is also an empirical question. Yet, it seems radical to assume away one of the main forces analysed by economic geographers. It is indeed the re-emerging awareness of the importance of external scale economies which underlies the "new economic geography".[15]

External Scale Economies

According to equation (1), a firm's input requirements depend solely on the quantity of its own output. More generally, a cost function with no external effects can be expressed as follows:

$$l_i = f(x_i), \qquad \textbf{(9)}$$

where $f'(x_i) < \frac{f(x_i)}{x_i}$ indicates increasing returns, $f'(x_i) > \frac{f(x_i)}{x_i}$ represents decreasing returns, and $f'(x_i) = \frac{f(x_i)}{x_i}$ describes constant returns to scale.

If, however, external conditions affect a firm's input requirements, cost functions would be of the form:

$$l_i = f(x_i, x), \qquad \textbf{(10)}$$

where $\frac{df}{dx} < \frac{f(x_i, x)}{x_i}$,

and where x can reflect any external influence. Two particular conceptions of x have been the main focus of past analyses. First, we can set $x = x$, where x stands for the output of products similar to x_i within one region or country. This scenario was first put forward by Alfred Marshall. It received some attention by trade theorists previous to the NTT and it has a long tradition in economic

[15]Krugman (1994, p. 23) has been at the forefront of this reappraisal: "During the early years of the new trade theory, I was strongly committed to the view that internal economies of scale, leading to intra-industry trade in differentiated products, were where the important action was; (...) I now tend to believe that external economies are at least equally important."

geography.[16] Second, if X represents world output of products similar to x_i, and if $x = X$, then we model what Ethier (1979) has termed "internationally decreasing costs", where external economies spill across regional and national borders and a sector's world-wide growth brings about a reduction in average costs for the sector as a whole, while individual firms experience constant or decreasing returns.[17]

In the past, it was often out of analytical expediency that trade theorists assumed scale economies to be external to individual firms, since this framework is compatible with perfect competition within industrial sectors.[18] However, theoretical as well as empirical investigations have led to a recent re-appraisal of the importance of these effects on more convincing grounds.

Most authors follow Marshall's initial conception by assuming that externalities remain contained within regional or national borders.[19] In their essence, *country-specific* externalities generate the same paths of specialisation and IIT as the ones described by the models based on internal scale economies and monopolistic competition[20]. The only difference is that localised external

---

[16]see Isard (1956), Chipman (1965), Ethier (1982b), Krugman (1991b), Gomory and Baumol (1992) and Krugman and Venables (1993).

[17]A good example for the importance of this distinction is GATT (1994). In a computable general equilibrium exercise, the effects on world income growth of the Uruguay Round package were estimated for the 1990-2005 period. Under the assumption of constant returns and perfect competition, the calculated annual gain in world income amounted to 184 billion U.S. dollars (1990 prices). When regionally contained external scale economies were introduced to the model, this figure rose to 218 billion dollars. Finally, assuming also that there are significant internal scale economies (and monopolistic competition), an annual income gain of 510 billion dollars was estimated.

[18]see Chipman (1965), p. 737.

[19]The importance of regionally contained external scale economies has recently been stressed by Porter (1990), Krugman (1991b) and Gomory (1994).

[20]From his theory based on internal scale economies, Krugman (1980, p. 957) concludes that "if the two countries have sufficiently dissimilar tastes, each will specialise in the industry for which it has the larger home market"; while his analysis of external economies applying to the manufacturing sector as a whole (Krugman, 1991a, p. 483) results in the strikingly similar finding that "in order to realise scale economies while minimising transport costs, manufacturing firms tend to locate in the region with larger demand". Panagariya (1981, p. 227), modelling an (external) increasing-returns as well as a decreasing-returns industry, found that "free trade will result in the larger country exporting the [internal-returns] commodity and the smaller country exporting the [decreasing-returns] commodity".

economies, unlike internal economies, induce countries to specialise between industrial sectors, even when market sizes are exactly equal in all countries. Where there are differences in market size, however, both forces push towards concentration of the industry near the larger market area and can thus reinforce each other.

The implications of *internationally* decreasing costs on specialisation and IIT are quite different. In a world producing horizontally differentiated goods with identical technologies in countries sharing equal factor endowments, two basic scenarios can be distinguished. First, consider the case where there are constant or decreasing returns at firm and industry level. In such a situation, global externalities provide no incentive for specialisation and trade, and transport costs ensure that all potential varieties are produced in every country. Unlike internal economies and localised externalities, internationally decreasing costs thus present no source of trade complementary to the classical determinants. Second, it is possible to think of international externalities arising in conjunction with internal scale economies or localised externalities. While international externalities in this scenario magnify the increase in global production and trade volumes resulting from trade liberalisation, they have no effect on the pattern of specialisation and IIT, since they affect the firms of each country symmetrically.

III. The Effects of Integration under Alternative Scenarios

The configuration generating the IIT patterns scrutinised in Sections I and II is highly abstract, because all products are subject to the same manufacturing technology, and both countries are endowed equally with production factors. So far, we have also assumed that adjustment to exogenous changes, namely trade liberalisation, is costless and instantaneous. This is a highly restrictive assumption, particularly when applied to the relocation of industrial production between countries.

By departing from the monopolistic-competition model of the NTT, we make a step towards economic reality. We now also consider the requirements and availability of various factor inputs. Meanwhile, we maintain the assumptions that internal or localised external scale economies apply and that consumers have a taste for variety, since both conditions are instrumental in the theoretical explanation of horizontal IIT.

Oligopoly and Adjustment

As a way of introducing adjustment rigidities to the analysis, we first consider the opposite extreme to the perfect locational mobility of production assumed up to now. We impose the constraint of a fixed number of firms (and product varieties) in each country. This constraint is incompatible with the highly stylised model discussed above, where all the effects of trade were embodied in the changing numbers of firms in each sector and country. With constant firm numbers, however, price and quantity changes have to be incorporated in the analysis. The assumption of oligopolistic markets also entails that general equilibrium modelling is no longer possible.

Krugman and Venables (1990, Section 3) have modelled this scenario.[21] The familiar features of a one-factor world divided into two countries, each producing varieties of a horizontally differentiated product with identical increasing-returns technologies are maintained. Two main differences apply. First, demand functions are linear. Second, the fixity of firm numbers rules out monopolistic competition. For simplicity, firms are assumed to interact as Cournot competitors in each national market, with every producer maximising profits while taking the output of the others as given. The fixed numbers of firms can be at any level below the equilibrium firm number resulting from monopolistic competition, hence super-normal profits are attained by the oligopolists.

Transport costs t are assumed to be borne by the producer. Cost function (1) thus has to be re-formulated separately for goods sold in their country of origin (1) or abroad (2):

$$l_{i1} = a + bx_{i1};\ l_{i2} = a + x_{i2}(b+t); \quad i = 1,...,n; \quad a,\ b,\ t > 0. \qquad \textbf{(11)}$$

[21]The most widely discussed model of IIT under oligopoly is that developed by Brander (1981) and Brander and Krugman (1983). Since all trade in that model is necessarily IIT, however, it is not very useful for our current purpose, which is to observe the variation in IIT levels, when the values of other parameters are changed.

The implications of this model are best illustrated graphically, based on a numerical example.[22] Figure 4 depicts equilibrium levels of industrial production, consumption and trade in the country with the smaller home market at various levels of transport costs. We have added the calculation and graphical representation of IIT to the analysis conducted by Krugman and Venables.[23] All quantities are represented as a proportion of domestic consumption at zero transport costs (vertical axis), and transport costs are expressed as a percentage of marginal costs (horizontal axis).

The IIT path associated in this model with various levels of trade liberalisation is in stark contrast to the pattern observed in the monopolistic competition case. While both models associate lower transport costs with higher trade volumes, only the scenario sketched in Figure 4 shows increased trade flows in *both* directions within the horizontally diversified sector. This equilibrium path thus establishes a strictly non-positive relationship between IIT and transport costs, or, if interpreted in a dynamic sense, strictly positive marginal IIT, provided that transport costs have fallen below the threshold above which no IIT arises.

The explanation of this pattern is straightforward. Through the very assumption of fixed firm numbers we rule out re-location and inter-industry specialisation. Trade liberalisation in this context increases firms' profits by offering greater export demand, while lowering profits by increasing competition from foreign rivals. Since there is no entry of firms and each firm produces a unique horizontally differentiated good, profit levels remain positive throughout even though the oligopolists raid each others' markets.

This analysis is of interest since it might well reflect the short-term patterns of adjustment after a step of trade liberalisation.[24] It seems plausible to think that

---

[22]For the precise functions and parameters underlying the numerical simulations see Krugman and Venables (1990), pp. 61f. and p. 75.

[23]IIT was calculated according to equations (8), setting $f = (n_A * x_{iAB}) / (n_B * x_{iBA})$. We thus represent IIT in quantity terms. Consideration of price effects and the calculation of IIT in value terms does not alter the intersections of the IIT line with the axes of Fig. 4, but it flattens the positive segment of the curve.

[24]Indeed, if product differentiation is costly, and incumbent firms manage to erect lasting barriers to entry, then this pattern could persist indefinitely (see Sutton, 1981). This is, of course, the scenario underlying the models of Brander (1981) and Brander and Krugman (1983).

production adjusts first within established firm structures, only to give way to new locational patterns after a certain period of transition.[25] Such rigidities can arise for a number of reasons such as temporarily limited producer horizons due to sunk costs. Profit-maximising firms may thus delay re-location until existing fixed costs are written off, beyond the date when total marginal costs at a new location fall below existing ones.

A combination of the fixed-firm oligopoly model with the free-entry NTT configuration could thus yield a plausible explanation of the recent indications that IIT growth among industrial countries has stalled. If there is a significant time-lag between the liberalisation of trade and the re-location of production, IIT shoots up initially, but withers away later, when the forces for industrial agglomeration come into full effect.[26]

Upward Sloping Factor Supply

A perfectly elastic supply of factor inputs to the differentiated industry has so far been assumed implicitly in the partial equilibrium analysis of this Section. Concentration of increasing-returns activities in one country had no effect on wages. This assumption might be realistic for industries that are small relative to the total economy and which use non-specific inputs. However, the bigger an industry and the more specific its factor requirements, the more important is the consideration of general equilibrium effects in the form of factor-price equalisation. Such effects result in a negative externality from industrial concentration, working in the opposite direction to the internal or external economies of increased production scale.

To model imperfectly elastic input supply we retain one non-specific factor input, l, with price w, for the diversified industry, while the perfectly competitive "rest of the economy" also uses a specific input, k, with price r.[27]

[25] This could be thought of in terms of entry and exit of firms, or in terms of re-locating existing firms. In other words, our analysis only considers *where* the production of certain goods takes place, irrespective of factor ownership.

[26] The same finding is made by Torstensson (1995, p. 11).

[27] It might seem paradoxical to assume the factor used by the scrutinised industry to be non-specific, when factor specificity is one of the reasons for an upward-sloping sectoral input supply curve. This choice of modelling features is motivated by analytical convenience, and it implicitly assumes the diversified industry to be large enough, relative to the rest of the economy, to affect the

Following Krugman and Venables (1990), we assume that the two countries have equal factor endowment ratios l/k, but different labour endowments $l_{1,2}$. Factor market equilibrium in country 1 is thus defined as follows:

$$l_1 = n_1(\{x_{11} + x_{12}\}c + f) + y_1 b_w(w_1, r_1), \text{ and}$$

$$k_1 = y_1 b_r(w_1, r_1), \qquad \textbf{(12)}$$

where y is the output of the "rest of the economy" and b its average cost function. If $b_r > 0$, there are diminishing returns to l and the wage rate increases with the size of the diversified industry, which uses more of l per unit of output than the perfectly competitive sectors. An expansion of the increasing-returns industry thus results both in cost savings due to the reduced amount of fixed costs f per unit of output x, and in cost increases due to a higher factor price ratio w/r. The interaction of these two effects can give rise to a non-monotonous specialisation path following trade liberalisation.

Numerical simulations based on their algebraic model revealed to Krugman and Venables (1990, p. 71) a U-shaped production curve in the smaller country along the path of trade liberalisation, where, at the initial stages, "the smaller economy experiences falling manufacturing production and downward pressure on wages as barriers fall. As the process continues, manufacturing production increases, and there is a convergence of peripheral wages to those of the centre." The process of industrial concentration towards the larger market, induced by the exploitation of scale economies, occurs at some intermediate stages of liberalisation, while in the vicinity of the complete elimination of trade barriers input prices replace market proximity as the main determinant of industrial location.

In spite of the non-monotonous evolution of peripheral production, the IIT pattern generated by this model resembles the IIT path in a world with fixed firm numbers: the lower the trade barriers, the higher are the levels of IIT. In other words, the "U-curve" of production does not yield a "U-curve" of IIT. In the interval of trade barriers contained between the point where all international trade is choked off and somewhere around the bottom of the "U-curve", industrial activity concentrates in the gravitational core and unidirectional trade in

general wage level. Different causes for the imperfect elasticity of factor supply do not affect the induced patterns of specialisation and trade.

industrial goods from the centre to the periphery occurs. Once, however, trade barriers have fallen far enough, production re-locates to the low-wage periphery and, as the size of the peripheral industry grows relative to the established activities in the core, the proportion of two-way trade increases steadily. The assumption of inelastic factor supply eliminates the possibility of total re-location of industrial activities after complete elimination of trade barriers, so that IIT remains positive throughout the process of trade liberalisation.

Comparative Advantage

In the situation discussed above, countries only differed in terms of market size. It is a small step to introduce different factor endowments into the analysis of industry location and trade. If the differentiated industry is assumed to use both factors, l and k, and the relative factor endowments are not equal in the two countries ($l_1/k_1 \neq l_2/k_2$), neo-classical determinants of trade can be analysed in conjunction with the effects of scale economies and product differentiation.

The input requirements of the increasing-returns industry in a two-factor world have been modelled in two different ways. In Krugman and Venables (1990), both fixed and marginal costs have the same capital-labour ratio, yielding the following cost function for a firm i in country 1:

$$TC_i = f_i * b_i(w_1, r_1) + (x_{i11} + x_{i12}) * b_i(w_1, r_1). \quad \textbf{(13)}$$

Under this configuration, trade liberalisation attracts the sector to the country whose factor endowment ratio matches the sector's more closely. If the production technology of the diversified industry is relatively labour intensive and the smaller country is relatively labour abundant, then comparative advantage countervails the centripetal forces due to scale economies and market proximity, and the locus of production in the small country as a function of falling trade barriers takes the U shape outlined above, generating a tendency for rising IIT. However, if the large country also happens to enjoy a comparative advantage in the increasing-returns industry, then the centripetal forces are reinforced and IIT vanishes rapidly as liberalisation proceeds.

Lawrence and Spiller (1983) have applied the alternative hypothesis that fixed costs consist entirely of capital, and that labour is the only source of variable costs. Hence, the total cost function for firm i in country 1 becomes

$$TC_i = f_i r_1 + (x_{i11} + x_{i12}) w_1. \qquad \textbf{(14)}$$

In this instance, the relative importance of fixed and variable cost determines factor requirements. The greater the scale of operation, the smaller is the weight of fixed costs in total outlays, and the stronger is the locational attraction of the labour-abundant country. Conversely, where capital-intensive fixed costs are important relative to labour-intensive variable costs, the capital-abundant country enjoys a comparative advantage.

The Lawrence and Spiller specification inspires an interesting hypothesis, since it can serve as a link to the theory of industrial innovation and the product cycle. It does not seem unrealistic to assume that new, technologically advanced products involve high fixed costs relative to variable costs, since considerable research and development costs have to be imputed on relatively small sales due to (as yet) small market demand. The manufacture of standardised goods, on the other hand, is likely to have relatively small fixed costs, so that the availability of the variable-cost factor - presumably low-skilled labour - becomes more important. It can thus be hypothesised that with progressing integration the production of standardised goods concentrates in the (low-skill) labour abundant country, while innovative products will emerge from the capital (or high-skill labour) abundant country. Such a process would result in an increase in *vertical* IIT following trade liberalisation.

It should be noted that economic integration facilitates not only the international exchange of goods but also the transfer of production factors. Krugman and Venables (1990) have pointed out two effects of factor mobility both of which can be conjectured to reduce IIT. First, as factor prices are equalised through factor mobility, the relative importance of market access increases. Second, if labour migrates from the small to the large country, the market-size differential rises and the centripetal forces are strengthened further.

Conclusion

In the "new trade theory" (NTT), scale economies, even though a necessary condition for the appearance of horizontal intra-industry trade (IIT), ultimately also lead to its elimination. This result stems from the way in which scale economies are modelled. The standard NTT models assume marginal costs to lie below average costs for any level of output. Hence, there is no such thing as a minimum efficient scale, and a larger firm is always more competitive than a smaller one. This leads to the intuitive result that the reduction of distance costs sharpens the competitive advantage of the bigger incumbents. Since the bigger firms are the ones with the larger home market, the NTT suggests that integration leads to a concentration of industrial activity in those countries which had previously offered the largest markets for a particular range of products.[28]

If, however, there are rigidities which slow down the re-location of industrial activity in an integrating area, then integration leads to an initial surge of IIT, followed by a decline once the move towards a specialised long-term equilibrium sets in. This pattern seems to conform quite well with observed reversals in IIT growth among industrialised countries.

The effects on IIT of a combination of factor market considerations with the modelling of scale economies can be summarised as follows. Where factor market conditions favour location in the larger country, the locational pull of scale economies is reinforced and the decline of IIT is accelerated. Where relative factor endowments favour production in the smaller country, however, the centripetal forces are attenuated, and IIT is likely to rise, or at least to remain high, as economic integration proceeds.

The NTT provides a powerful and compact explanation of trade flows among identically endowed countries. However, it is based on two highly abstract concepts: indefinitely falling average costs and costless differentiability of products. Further research might strive to incorporate more realistic modelling of scale economies and temporary as well as permanent limitations to firm numbers and product specifications. This would lead to game-theoretical

[28]In Krugman's (1980) model, large-country advantage arises not from different firm size but from lower average c.i.f. prices for all firms of equal size.

analysis of oligopolistic markets, which is inevitably messy but potentially much closer to the real world of international trade.

Figure 1
Intra-Industry Trade and Relative Market Sizes

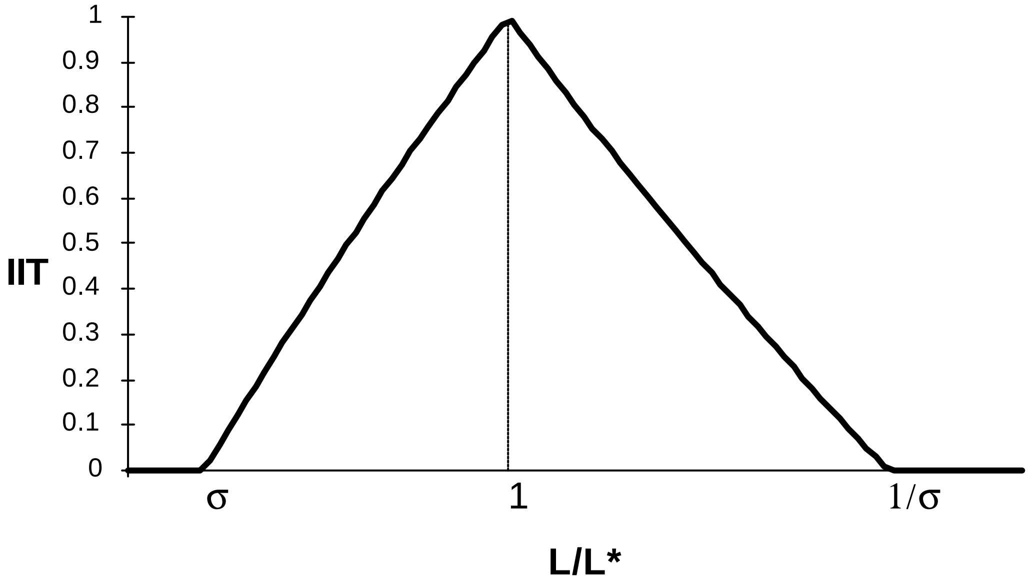


Figure 2
Intra-Industry Trade and Scale Economies
(t = 0.8)

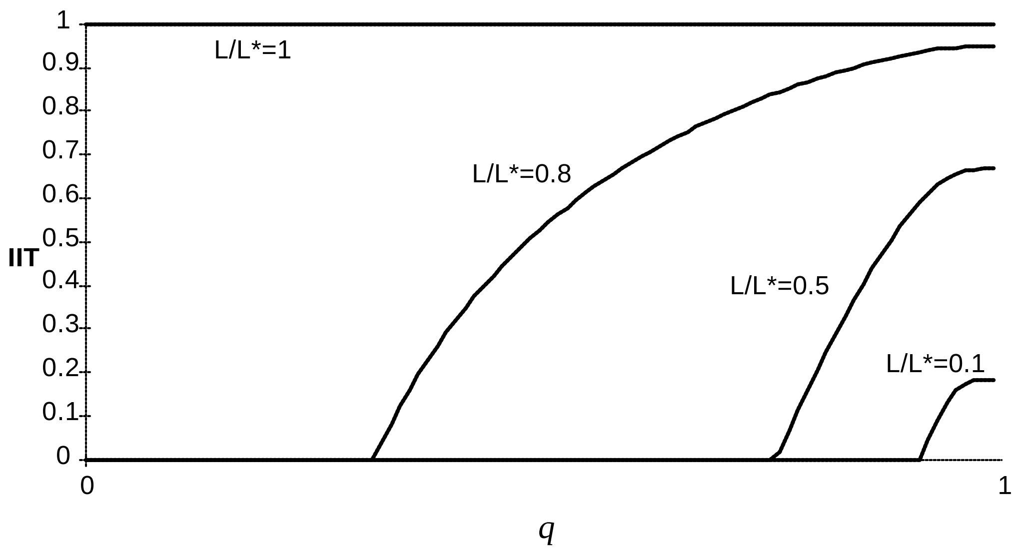


Figure 3
Intra-Industry Trade and Transport Costs
(q = 0.5)

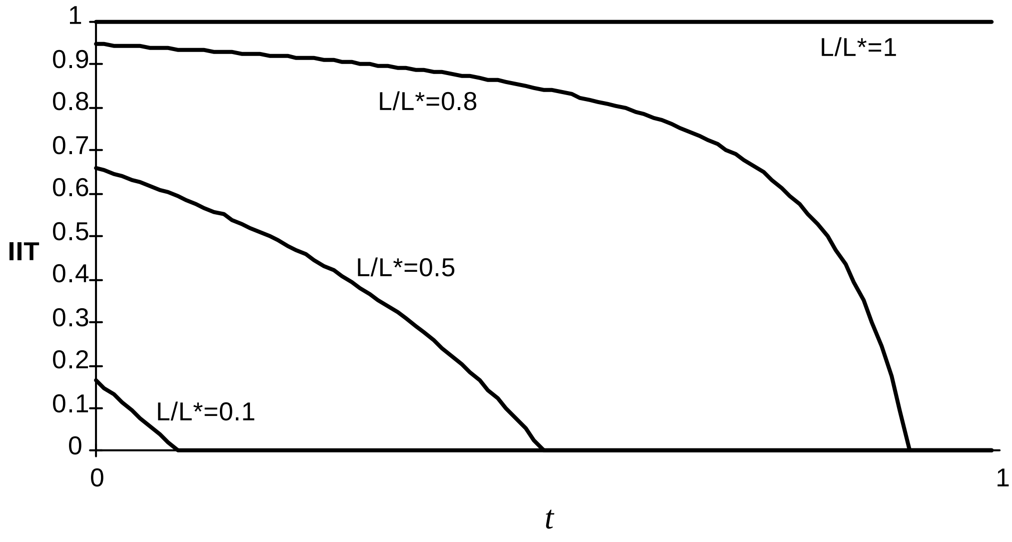
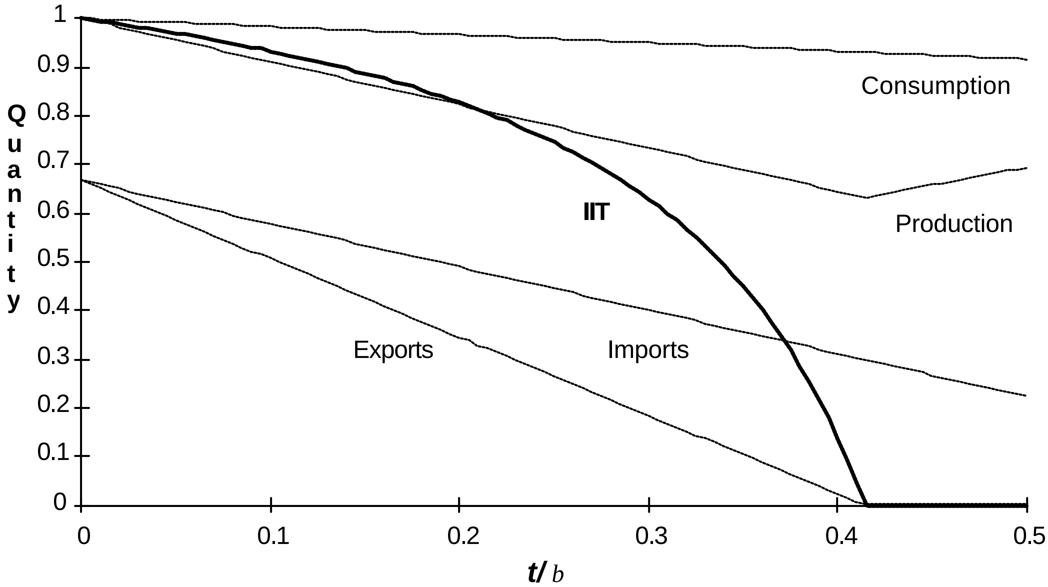


Figure 4
Fixed Number of Firms: Output and Trade



Appendix: Mathematical Proofs

The relationships between equilibrium IIT and various country characteristics, derived in Section I from the monopolistic competition model by Krugman (1980), were illustrated graphically on the basis of numerical simulations. More rigorous algebraic proofs of the signs of these relationships are provided here.

Because the IIT formula differs between a situation of net imports and one of net exports (equation (8)), we limit this analysis to a situation of net imports. This does not curtail the generality of the proofs, since the net-import and net-export intervals are perfectly symmetrical (see Figure 1).

We thus assume that the home market for the particular industry is smaller than the foreign market:

$L < L^*$.

Consequently, the amount of home-produced varieties within this industry is smaller than the number of varieties produced abroad (see equation (5)):

$n < n^*$.

The following formula thus applies for the calculation of IIT (see equations (4) to (7)):

$$IIT = \frac{2f}{1+f} = \left(2\frac{\frac{L}{L^*} - t^{\frac{q}{1-q}}}{1 - t^{\frac{q}{1-q}}\frac{L}{L^*}}\right) \Big/ \left(1 + \frac{\frac{L}{L^*} - t^{\frac{q}{1-q}}}{1 - t^{\frac{q}{1-q}}\frac{L}{L^*}}\right). \qquad \textbf{(15)}$$

Since $L < L^*$ and $s < 1$, $\frac{L}{L^*} < \frac{1}{s}$, and hence equation (15) is always defined.

The relationships depicted in Sections I and II can be formally validated by determining the signs of the first partial derivatives of equation (15).

In Figure 1, we have plotted a non-positive relationship between IIT and the size difference in the two countries' markets for the varieties produced by a certain industry. We thus differentiate equation (15) with respect to (L/L^*), which yields the following:

$$\frac{dIIT}{d\frac{L}{L^*}} = \frac{2(1+s)}{(1+\frac{L}{L^*})^2(1-s)}, \tag{16}$$

which is positive, and thus suggests that IIT rises as the size of the domestic market approaches the size of the larger foreign market.

Figure 2 shows a non-negative relationship between IIT and q, an inverse measure of equilibrium scale economies. Differentiating equation (15) with respect to q, we obtain:

$$\frac{dIIT}{dq} = \frac{2s\ln(t)(-1+\frac{L}{L^*})}{(-1+q)^2(-1+s)^2(1+\frac{L}{L^*})}, \tag{17}$$

where all elements are positive, except for the two last factors of the numerator, $\ln(t)$ and $(-1+\frac{L}{L^*})$. Hence, the ratio (17) is positive, which confirms the negative relationship between IIT and equilibrium scale economies.

Finally, the non-positive link between IIT and transport costs, depicted in Figure 3, can be proven algebraically by differentiating equation (15) with respect to t:

$$\frac{dIIT}{dt} = \frac{2(\frac{qs}{t})(-1+\frac{L}{L^*})}{(1-q)(-1+s)^2(1+\frac{L}{L^*})}, \tag{18}$$

where all elements are positive, except for the last factor of the numerator, $(-1+\frac{L}{L^*})$. Consequently, (18) is negative. This confirms our finding that the model predicts equilibrium IIT to fall during a process of economic integration.

References

Balassa, Bela and Bauwens, Luc (1987) "Intra-Industry Specialisation in a Multi-Country and Multi-Industry Framework". *Economic Journal*, vol. 97, pp. 923-939.

Bhagwati, Jagdish and Davis, Donald R. (1994) "Intraindustry Trade Issues and Theory". Harvard Institute of Economic Research, Discussion Paper No. 1695.

Brander, James (1981) "Intra-Industry Trade in Identical Commodities". *Journal of International Economics*, vol 11, pp. 1-14.

Brander, James and Krugman, Paul (1983) "A 'Reciprocal Dumping' Model of International Trade". *Journal of International Economics*, vol. 15, pp. 313-321.

Brülhart, Marius (1994) "Marginal Intra-Industry Trade: Measurement and Relevance for the Pattern of Industrial Adjustment". *Weltwirtschaftliches Archiv*, vol. 130, pp. 600-613.

Brülhart, Marius (1995) "A Test of the "New Trade Theory": Industrial Specialisation in the European Union". *Trinity Economic Papers*, No. 95/5.

Brülhart, Marius and McAleese, Dermot (1995) "Intra-Industry Trade and Industrial Adjustment: The Irish Experience". *The Social and Economic Review*, vol. 26, pp. 107-129.

Chipman, John S. (1965) "A Survey of the Theory of International Trade: Part 2, The Neo-Classical Theory". *Econometrica*, vol. 33, pp. 685-760.

Clark, Don P. (1993) "Recent Evidence on Determinants of Intra-Industry Trade". *Weltwirtschaftliches Archiv*, vol. 129, pp. 332-343.

Davis, Donald R. (1993) "Intraindustry Trade: A Heckscher-Ohlin-Ricardo Approach". *mimeo*, Harvard University, United States.

Dixit, Avinash K. and Stiglitz, Joseph E. (1977) "Monopolistic Competition and Optimum Product Diversity". *The American Economic Review*, vol. 67, pp. 297-308.

Ethier, Wilfred (1979) "Internationally Decreasing Costs and World Trade". *Journal of International Economics*, vol. 9, pp. 1-24.

Ethier, Wilfred (1982a) "National and International Returns to Scale in Modern Theory of International Trade". *American Economic Review*, vol. 72, pp. 389-405.

Ethier, Wilfred (1982b) "Decreasing Costs in International Trade and Frank Graham's Argument for Protection". *Econometrica*, vol. 50, pp. 1243-1268.

Finger, J.M. (1975) "Trade Overlap and Intra-Industry Trade". *Economic Inquiry*, vol. 13, pp. 581-589.

Finger, J.M. and DeRosa, Dean A. (1979) "Trade Overlap, Comparative Advantage and Protection". In: Giersch, H. (ed.) *On the Economics of Intra-Industry Trade.* J.C.B. Mohr, Tübingen.

General Agreement on Tariffs and Trade (GATT) (1994) "Market Access for Goods and Services: Overview of the Results". GATT Secretariat, Geneva, Switzerland.

Globerman, Steven and Dean, James W. (1990) "Recent Trends in Intra-Industry Trade and Their Implications for Future Trade Liberalization". *Weltwirtschaftliches Archiv*, vol. 126, pp.25-49.

Gomory, Ralph E. (1994) "A Ricardo Model with Economies of Scale". *Journal of Economic Theory*, vol. 62, pp. 394-419.

Gomory, Ralph E. and Baumol, William J. (1992a) "Scale Economies, the Regions of Multiple Trade Equilibria, and Gains from the Acquisition of Industries". C.V. Starr Center for Applied Economics, New York University, Economic Research Report No. 92-10.

Greenaway, David and Hine, Robert C. (1991) "Intra-Industry Specialization, Trade Expansion and Adjustment in the European Economic Space". *Journal of Common Market Studies*, vol. 24, pp. 603-622.

Greenaway, David and Milner, Chris (1986) *The Economics of Intra-Industry Trade*. Oxford, Basil Blackwell.

Harrigan, James (1994) "Scale Economies and the Volume of Trade". *Review of Economics and Statistics*, vol. 76, pp. 321-328.

Helpman, Elhanan and Krugman, Paul (1985) *Market Structure and Foreign Trade*. Cambridge (Mass.), MIT Press.

Helpman, Elhanan and Krugman, Paul (1989) *Trade Policy and Market Structure*. Cambridge (Mass.), MIT Press.

Hirschberg, Joseph G.; Sheldon, Ian, M. and Dayton, James R. (1994) "An Analysis of Bilateral Intra-Industry Trade in the Food Processing Sector". *Applied Economics*, vol. 26, pp. 159-167.

Isard, Walter (1956) *Location and Space-Economy*. New York, John Wiley and Sons.

Krugman, Paul (1979) "Increasing Returns, Monopolistic Competition, and International Trade". *Journal of International Economics*, vol. 9, pp. 469-479.

Krugman, Paul (1980) "Scale Economies, Product Differentiation, and the Pattern of Trade". *American Economic Review*, vol. 70, pp. 950-959.

Krugman, Paul (1981) "Intraindustry Specialization and the Gains from Trade". *Journal of Political Economy*, vol. 89, pp. 959-973.

Krugman, Paul (1991a) "Increasing Returns and Economic Geography". *Journal of Political Economy*, vol. 99, pp. 483-499.

Krugman, Paul (1991b) *Geography and Trade.* Cambridge (Mass.), MIT Press.

Krugman, Paul (1993) "The Narrow and Broad Arguments for Free Trade". *American Economic Review*, vol. 83 (AEA Papers and Proceedings), pp. 362-366.

Krugman, Paul (1994) "Empirical Evidence on the New Trade Theories: The Current State of Play". In: CEPR Conference Report, *New Trade Theories*, London.

Krugman, Paul and Venables, Anthony J. (1990) "Integration and the Competitiveness of Peripheral Industry". In: Bliss, C. and de Macedo, J. (ed.) *Unity with Diversity in the European Community*, Cambridge University Press, pp. 56-75.

Krugman, Paul and Venables, Anthony J. (1993) "Integration, Specialization and Adjustment". *NBER Working Paper*, No. 4559, Cambridge, Mass.

Lancaster, Kelvin (1980) "Intra-Industry Trade under Perfect Monopolistic Competition". *Journal of International Economics*, vol. 10, pp. 151-175.

Leamer, Edward (1995) "A Trade Economist's View of U.S. Wages and 'Globalization'". *mimeo*, University of California, Los Angeles.

Lundberg, Lars (1992) "The Structure of Swedish International Trade and Specialization: 'Old' and 'New' Explanations". *Weltwirtschaftliches Archiv*, vol. 128, pp. 266-287.

Lundberg, Lars and Hansson, Pär (1986) "Intra-Industry Trade and its Consequences for Adjustment". In: Greenaway and Tharakan, *Imperfect Competition and International Trade.* London, Wheatsheaf, pp. 129-147.

Panagariya, Arvind (1981) "Variable Returns to Scale in Production and Patterns of Specialization". *American Economic Review*, vol. 71, pp. 221-230.

Porter, Michael (1990) *The Competitive Advantage of Nations.* New York, Free Press.

Richardson, J. David (1989) "Empirical Research on Trade Liberalisation with Imperfect Competition: A Survey". *OECD Economic Studies*, vol. 12, pp. 7-50.

Sutton, John (1991) *Sunk Costs and Market Structure*. Cambridge (Mass.), MIT Press.

Torstensson, Johan (1995) "Economic Integration, Market Size and the Location of Industries". *CREDIT Research Paper*, No. 95/2, University of Nottingham.

Tybout, James R. (1993) "Internal Returns to Scale as a Source of Comparative Advantage: The Evidence". *American Economic Review*, vol. 83 (AEA Papers and Proceedings), pp. 440-444.