The University of Nottingham

research paper series

Research Paper 2004/10

Intra-Industry Trade with Multinational Firms: Theory, Measurement and Determinants

by

Hartmut Egger, Peter Egger and David Greenaway

The Centre acknowledges financial support from The Leverhulme Trust under Programme Grant F114/BF

GEP Leverhulme Centre for Research on Globalisation and Economic Policy

The Authors

Hartmut Egger is Senior Assistant at the University of Zurich. Peter Egger is Professor of Economics at the University of Innsbruck. David Greenaway is Professor of Economics at the University of Nottingham and Director of the Leverhulme Centre for Research on Globalisation and Economic Policy (GEP).

Acknowledgements

David Greenaway acknowledges support from the Leverhulme Trust under programme grant F114/BF.

Intra-Industry Trade with Multinational Firms: Theory, Measurement and Determinants

by

Hartmut Egger, Peter Egger and David Greenaway

Abstract

A number of recent developments, including the analysis of firm level adjustment to falling trade costs, have contributed to a revival of interest in intra-industry trade. Most empirical work still relies on the standard Grubel-Lloyd measure. This however refers only to international trade, disregarding income flows stimulated by repatriated profits. Given the overwhelming importance of the latter, this is a major shortcoming**.** We provide a guide to measurement and estimation of the determinants of bilateral intra-industry trade shares from the perspective of new trade theory with multinational firms. We develop an analytically solvable general equilibrium model to investigate investment costs, multinational activities and income flows from repatriated profits. The robustness of our findings are investigated in five simulation analyses. We also discuss and quantify biases of different Grubel-Lloyd indices in an empirical assessment of intra-industry trade shares and identify repatriated profit flows of multinationals as a key determinant of biased measurement. To overcome this, we provide several alternative, bias-corrected versions of the Grubel-Lloyd index. Finally, we demonstrate that the determinants motivated by our theoretical analysis offer important insights into variations in the Grubel-Lloyd index. Our new specification outperforms any other previously estimated model as illustrated in regressions on numerically generated data.

JEL classification: F12, F23

Keywords: intra-industry trade, multinationals

Outline

Non-Technical Summary

The publication of Grubel and Lloyd (1975) stimulated enormous interest in intra-industry trade (IIT), for two reasons. First, the empirical phenomenon of high levels of trade in products from similar industries between countries with similar factor endowments seemed to be at odds with the standard Heckscher-Ohlin-Samuelson (HOS) workhorse model of international trade. Second, the observed increase in intra-industry trade coincided with what appeared to be relatively painless adjustment to economic integration in western Europe. The dislocation anticipated as inter-industry specialisation occurred did not materialise, giving rise to the so-called 'smooth adjustment hypothesis'.

In the decade that followed Grubel and Lloyd (1975) the literature exploded. Empirical analysis focused primarily on three things. First whether the phenomenon survived data disaggregation. Second, was IIT a peculiarity of trade in western Europe? Third, what were the drivers of the phenomenon?

Recent years have seen a revival of interest in intra-industry trade, stimulated by frontier work on trade costs, economic geography and a range of aspects of firm level adjustment to globalisation. One focus of this, from both a theoretical and measurement standpoint is intra-industry trade in a setting with multinational firms. This is a very important development from a theoretical standpoint because we have known for a long time that both phenomena co-exist, indeed are often co-terminous and we need good models for explaining this. But it is also important from a measurement perspective because of the importance of international production and intra-firm trade relative to armslength trade. FDI has grown about twice as fast as trade over the last decade. The principal sources and hosts are industrialised countries and two-way trade is closely associated with two-way FDI.

This paper contributes to this new literature in several ways. First, it generates a proof that the standard and still widely used Grubel-Lloyd index has to be adjusted to reflect more than the intra-industry trade share in a narrow sense. We build a general equilibrium model which shows that with multinational firms, both unbalanced profit repatriation and trade costs distort the index. We expose the biases resulting from these empirically relevant phenomena and construct several new versions of bias-corrected Grubel-Lloyd indices. Second, we develop a three-factor general equilibrium model of trade and multinationals to provide a detailed analysis of the role of investment cost differences between countries as a determinant of FDI and, hence, intra-industry trade. By introducing three factors, we emphasise the distinction between two important characteristics of headquarters: their provision of physical capital to set up plants, and the human-capital intensive generation of firm-specific assets through brand proliferation. Besides this more complete description of headquarter services, there is an advantage of analytical tractability since there are as many activities (homogeneous goods production, exporter and multinational production of manufactures) as there are factors (physical capital, skilled labour, unskilled labour). In this setting, we are able to evaluate not only the role of investment cost levels and differences in general but also their interaction with labour and capital endowments, depending on whether horizontal or vertical multinationals are active.

Third, a large number of numerical simulations of our model allow us to evaluate the robustness of our analytical findings with respect to simplifying assumptions as well as traditional determinants such as country size, capital-labour ratios and skilled-unskilled ratios.

Finally, we implement and report on an extensive empirical analysis, where uncorrected and bias-corrected versions of the Grubel-Lloyd index are used as regressors. This yields several conclusions. We find that biases not only affect the overall magnitude of the Grubel-Lloyd index but also systematically affect parameter estimates; cross-section estimates tend to be inconsistent if country-specific effects are excluded; the determinants generated by our theoretical model account for more than 50% of the variation in intra-industry trade-share data, implying that less than half of their variation is explained by traditionally used variables. Given the crucial importance of estimating accurately intra-relative to inter-industry trade, this is very significant.

1 Introduction

The publication of Grubel and Lloyd (1975) stimulated enormous interest in intra-industry trade (IIT), for two reasons. First, the empirical phenomenon of high levels of trade in products from similar industries between countries with relatively similar factor endowments seemed to be at odds with the standard Heckscher-Ohlin-Samuelson (HOS) workhorse model of international trade. Second, the observed increase in intra-industry trade coincided with what appeared to be relatively painless adjustment to economic integration in western Europe. The dislocation anticipated as inter-industry specialisation occurred did not materialise, giving rise to the so-called 'smooth adjustment hypothesis'.

In the decade that followed Grubel and Lloyd (1975) the literature exploded. Empirical analysis focused primarily on three things. First whether the phenomenon survived data disaggregation. Finger (1975) famously described IIT as a 'statistical artefact', a mirage created by the vagaries of statistical classification. Greenaway and Milner (1983) among others showed that although shares of IIT in total trade declined as trade data became more finely disaggregated, it did not disappear. In fact it remained prevalent. Second, was IIT a peculiarity of trade in western Europe. Studies in Tharakan (1983) demonstrated that it was not. Although average levels were lower in developing, and what are now referred to as transition economies, they were non-trivial. Third, what were the drivers of the phenomenon? Early cross-section work such as Loertscher and Wolter (1980) and Greenaway and Milner (1984) pointed to various aspects of industrial organisation but findings were not robust. Indeed, an application by Torstensson (1996) of extreme bounds analysis confirmed that the cross-industry determinants were very fragile.

This, and other work, progressed thinking on measurement and to a lesser extent explanation. Innovations on the theoretical front were much more dramatic, with the development and refinement of models of monopolistic competition and international trade (most notably Lancaster 1980, Krugman 1979 and 1980 and Helpman and Krugman 1985) as well as strategic interaction and intra-industry trade (eg Brander 1981 and Brander and Krugman 1982). These offered convincing explanations of the market structures under which we would expect IIT to be generated and have proved to be of lasting value. Many, and in particular Krugman (1981), focused on distributional consequences, emphasising the likelihood of greater symmetry between expanding and declining activities than in an HOS world and offering a theoretical underpinning to the potential for lower adjustment costs in an IIT setting as compared to HOS.

Recent years have seen a revival of interest in intra-industry trade, stimulated by frontier work on trade costs, economic geography and a range of aspects of firm level adjustment to globalization. One important focus of this, from both a theoretical and measurement standpoint is intra-industry trade in a setting with multinational firms. This is a very important development from a theoretical standpoint because we have known for a long time that both phenomena co-exist, indeed are often co-terminous and we need good models for explaining this. But it is also important from a measurement perspective because of the importance of international production and intra-firm trade relative to armslength trade. FDI has grown about twice as fast as trade over the last decade. The principal sources and hosts are industrialised countries and two-way trade is closely associated with two-way FDI.

An important development in understanding the relationship between IIT and intra-industry affiliate production is Markusen and Maskus (2001). From a specification based on numerical simulations of a two-factor knowledge capital model (associated with Carr et al., 2001 and Markusen, 2002), they find that intra-industry trade between the US and partner economies tends to **decrease** with greater similarity in size, which is at odds with the findings of Helpman (1987), Bergstrand (1990) or Hummels and Levinsohn (1995). They also found it decreased with the bilateral trade cost level, but increased with the bilateral level of investment costs. However, apart from these papers, this issue remains largely unexplored.

This paper contributes to this new literature in several ways. First, it generates a proof that the standard and still widely used Grubel-Lloyd index has to be adjusted to reflect more than the intra-industry trade share in a narrow sense. We build a general equilibrium model which shows that with multinational firms, both unbalanced profit repatriation and trade costs distort the index. We expose the biases resulting from these empirically relevant phenomena and construct several new versions of bias-corrected Grubel-Lloyd indices. Second, we develop a three-factor general equilibrium model of trade and multinationals to provide a detailed analysis of the role of investment cost differences between countries as a determinant of FDI and, hence, intra-industry trade. By introducing three factors, we emphasise the distinction between two important characteristics of headquarters: their provision of physical capital to set up plants, and the human-capital intensive generation of firm-specific assets through brand proliferation. Besides this more complete description of headquarter services, there is an advantage of analytical tractability since there are as many activities (homogeneous goods production, exporter and multinational production of manufactures) as there are factors (physical capital, skilled labour, unskilled labour). In this setting, we are able to evaluate not only the role of investment cost levels and differences in general, but also their interaction with labour and capital endowments, depending on whether horizontal or vertical multinationals are active.

Third, a large number of numerical simulations of our model allow us to evaluate the robustness of our analytical findings with respect to simplifying assumptions as well as traditional determinants such as country size, capital-labour ratios and skilled-unskilled ratios.

Finally, we implement and report on an extensive empirical analysis, where uncorrected and bias-corrected versions of the Grubel-Lloyd index are used as regressors. This yields several conclusions. We find that biases not only affect the overall magnitude of the Grubel-Lloyd index but also systematically affect parameter estimates; cross-section estimates tend to be inconsistent if country-specific effects are excluded; the determinants generated by our theoretical model account for more than 50% of the variation in intra-industry trade-share data, implying that less than half of their variation is explained by traditionally used variables. Given the crucial importance of estimating accurately intra-relative to inter-industry trade, this is very significant.

The remainder of the paper is organized as follows: Section 2 sets out our theoretical model of intra-industry trade with investment costs and introduces a corrected Grubel-Lloyd index. This is subjected to simulation analysis and a number of theoretical propositions are derived. Section 3 sets up our econometric analysis, reports our results and subjects them to sensitivity analysis. Section 4 concludes.

2 Theoretical background

2.1 The Grubel-Lloyd index

The Grubel and Lloyd (1971) index has become the standard measure for the intensity of intra-industry trade. In the two-country case, this is defined as[1]

$$GLI = \sum_k \frac{2 \times \min\left(EX_{ik}, IM_{ik}\right)}{\sum_k EX_{ik} + \sum_k IM_{ik}}, \tag{1}$$

where EX_{ik} is the value of country *i*'s exports of good *k*. IM_{ik} represents expenditures for country *i*'s imports of good *k*. Although this has been the index of choice for most researchers in this area for over 30 years, it is an inappropriate measure if there are *multinational activities* because GLI does not account for (unbalanced) repatriated profits of multinational firms and, therefore, underestimates the intra-industry trade share. For convenience, we use the term *trade imbalance bias* to refer to this measurement error.[2] To see this bias, consider the case of two economies with one sector of production and multinational activities of country *i* firms in country *j*. From payments balance it follows that $2 \times \min\left(EX_i, IM_i\right) < EX_i + IM_i$, if there are flows of repatriated profits due to multinational activities of country *i* firms in *j*. Thus, $GLI < 1$, according to (1). However, in a one-sector model there is by definition *only* intra-industry trade, so that the correct *GLI* must equal one.

To obtain an appropriate measure of the IIT share, we have to adjust the Grubel-Lloyd index for all income flows not due to goods trade, like repatriated profits.[3] More precisely, we correct the denominator of *GLI* for all output flows that are balanced by income flows not directly related to exports and imports. This gives a hypothetical measure of *balanced trade* in the denominator of *GLI*.[4] The *corrected Grubel-Lloyd index* for the two-country, multi-sector case is then:

---

[1] We do not distinguish between *c.i.f* and *f.o.b* data for the moment. For a rigorous discussion on different empirical specifications of the Grubel-Lloyd index see Subsection 3.1.

[2] Note that this has an entirely different motivation than the case made by Aquino (1978) for a correction for *aggregate* payments imbalance. As Greenaway and Milner (1981) showed this is neither defensible on theoretical grounds nor practicable.

[3] (See Subsection 3.1 and Appendix C for the quantification of this and other biases).

[4] This adjustment method was in fact first suggested by Grubel and Lloyd (1975). However, they did not develop it on the grounds that it lacked a clear theoretical motivation.

$$GLI^C = \sum_k \frac{2 \times \min\left(EX_{ik}, IM_{ik}\right)}{\sum_k EX_{ik} + \sum_k IM_{ik} - \left|\sum_k EX_{ik} - \sum_k IM_{ik}\right|}, \tag{2}$$

In our thought experiment with two one-sector economies and multinational activities of country *i* firms in country *j*, GLI^C gives a correct measure of the intra-industry trade share, i.e. $GLI^C = 1$.[5] According to (1) and (2), we obtain

$$SHI := \frac{GLI^C}{GLI} = 1 + \frac{\left|\sum_k EX_{ik} - \sum_k IM_{ik}\right|}{\sum_k EX_{ik} + \sum_k IM_{ik} - \left|\sum_k EX_{ik} - \sum_k IM_{ik}\right|} > 1 \tag{3}$$

as a measure of the trade imbalance bias in relative terms.

In what follows we are interested in the role of multinational activities and repatriated profits for income flows $\left|\sum_k EX_{ik} - \sum_k IM_{ik}\right|$. In particular we investigate how changes in the fixed costs of multinational activities as one key determinant of FDI-flows (see Amiti and Wakelin, 2003) affect the corrected Grubel-Lloyd index given in (2) and the ratio of the corrected and uncorrected indices as in (3). To identify the basic economic mechanisms, we start with two analytically solvable general equilibrium models, which account for horizontal and vertical multinational activities, then provide simulation analyses of five variants of new trade theory models with multinational firms.

2.2 Two analytically solvable models

Consider two countries with two sectors, which differ only with respect to factor endowments. In the industrial *X*-sector differentiated goods are produced, while output in agricultural *Y*-sector is homogeneous. Preferences of consumers are identical and represented by a Cobb-Douglas utility function:

$$U = X^{\alpha} Y^{1-\alpha}, \quad 0 < \alpha < 1 \tag{4}$$

where $X := \left[\sum_k x_k^{(\varepsilon-1)/\varepsilon}\right]^{\varepsilon/(\varepsilon-1)}$, $\varepsilon > 1$, is a CES-index, that accounts for home-produced and imported varieties of the industrial good.[6] Production technologies in the two sectors are given by $x = L$ and $Y = L$, respectively, where *L* is unskilled labour. In addition, production in the *X*-sector requires fixed set-up costs through the use of capital *K* and skilled non-

[5] Noteworthy, we can substitute $EX_{jk} = IM_{ik}$ in (2) if *f.o.b.* measures are used in the calculations of GLI^C. This will be important in our analytical investigation below.

[6] Country indices are neglected for the moment.

production labour S. We choose unskilled labour of country i as the numéraire and thus, set $w_{Li} = 1$. Exporting differentiated industrial output gives rise to iceberg transport costs of 1-$1/t$>0 (in real terms). Trade in the homogeneous good does not induce any trade frictions.

Horizontal multinational enterprises

In a symmetric equilibrium with identical unskilled wages in the two economies, demand in country i for a single variant of the differentiated good is given by

$$x_{ii} = \frac{\alpha E_i p_{ii}^{-\varepsilon}}{P_i} \quad \text{and} \quad x_{ji} = x_{ii}\tau\,, \tag{5}$$

where x_{ii} is a variety produced and consumed in country i, while x_{ji} is produced in j and exported to i.[7] $E_i := L_i + w_{Ki}K_i + w_{Si}S_i$ is total factor income (total expenditures) of country i and $P_i = p_{ii}^{1-\varepsilon}\left(h_i + h_j + n_i\right) + n_j p_{ji}^{1-\varepsilon}$ is a price index. n_i, n_j and h_i, h_j are exporters and horizontal multinationals of countries i and j, respectively. $\tau = t^{1-\varepsilon}$ is a measure of iceberg transport costs. It is well-known from the literature that profit maximization leads to a constant price-markup and, therefore to prices $p_{ii} = \varepsilon/\left(\varepsilon - 1\right)$ and $p_{ij} = t\varepsilon/\left(\varepsilon - 1\right)$.[8]

To set up an exporting firm (n) requires one unit of capital and one of skilled labour, whilst one unit of skilled labour and $g_i > 2$ units of capital are required to set up a horizontal multinational (h) in i with one plant in i and one in j. Thus, in equilibrium, zero-profit conditions of country i firms are given by[9]

$$\pi_{ni} = \frac{1}{\varepsilon - 1}\left[x_{ii} + \tau x_{jj}\right] - w_{Ki} - w_{Si} = 0\,, \tag{6}$$

$$\pi_{hi} = \frac{1}{\varepsilon - 1}\left[x_{ii} + x_{jj}\right] - g_i w_{Ki} - w_{Si} = 0\,, \tag{7}$$

due to $w_{Li} = w_{Lj} = 1$ in the case of diversification. Finally, the three factor market clearing conditions in country i are given by

---

[7] If x_{ji} units of the industiral good are produced in in country j, only $\left(1/t\right)x_{ji}$ units arrive in country i, due to the existence of iceberg transport costs.

[8] Hence, the price index is given by $P_i := p_i^{1-\varepsilon}\left[h_i + h_j + n_i + n_j\tau\right]$ if $w_{Li} = w_{Lj} = 1$.

[9] Eqs. (6) and (7) build upon two simplifying assumptions, namely that (i) fixed costs of exporters and horizontal multinationals only differ with respect to the requirement of capital and that (ii) only factors of country i are used to set up country i firms (and their plants).

$$L_i = \left(h_i + h_j + n_i\right)x_{ii} + \tau n_i x_{jj} + Y_i\,, \tag{8}$$

$$S_i = n_i + h_i\,, \tag{9}$$

$$K_i = n_i + g_i h_i\,. \tag{10}$$

From (6)-(10), we obtain

$$w_{Ki} = \frac{1}{\varepsilon - 1}\frac{1-\tau}{g_i - 1}x_{jj}\,, \qquad w_{Si} = \frac{1}{\varepsilon - 1}\left[x_{ii} - x_{jj}\frac{1 - g_i\tau}{g_i - 1}\right] \tag{11}$$

for equilibrium wage rates in country *i* and

$$h_i = \frac{K_i - S_i}{g_i - 1}\,, \qquad n_i = \frac{g_i S_i - K_i}{g_i - 1} \tag{12}$$

for the equilibrium numbers of horizontal multinationals and exporters in country *i*. Equivalent expressions are obtained for wages and firm numbers in country *j*, if both sectors *X* and *Y* are active in both economies.

For the uncorrected and corrected Grubel-Lloyd indices we obtain, from (1) and (2),

$$GLI = \frac{2\varepsilon\tau \min\left[n_j x_{ii}, n_i x_{jj}\right]}{\varepsilon\tau\left(n_j x_{ii} + n_i x_{jj}\right) + \left|\left(\varepsilon\tau n_i + h_i\right)x_{jj} - \left(\varepsilon\tau n_j + h_j\right)x_{ii}\right|} \tag{1a}$$

and

$$GLI^C = \frac{2\varepsilon\tau \min\left[n_j x_{ii}, n_i x_{jj}\right]}{\varepsilon\tau\left(n_j x_{ii} + n_i x_{jj}\right) + \left|\left(\varepsilon\tau n_i + h_i\right)x_{jj} - \left(\varepsilon\tau n_j + h_j\right)x_{ii}\right| - \left|h_i x_{jj} - h_j x_{ii}\right|}, \tag{2a}$$

where $\left|\left(\varepsilon\tau n_i + h_i\right)x_{jj} - \left(\varepsilon\tau n_j + h_j\right)x_{ii}\right|$ is *Y*-trade[10], according to the balance of payment condition.[11] Moreover, $\left|h_i x_{jj} - h_j x_{ii}\right|$ is the balance of repatriated profits for which the denominator of GLI^C is adjusted. The respective share *SHI* is given by

$$SHI = 1 + \frac{\left|h_i x_{jj} - h_j x_{ii}\right|}{\varepsilon\tau\left(n_j x_{ii} + n_i x_{jj}\right) + \left|\left(\varepsilon\tau n_i + h_i\right)x_{jj} - \left(\varepsilon\tau n_j + h_j\right)x_{ii}\right| - \left|h_i x_{jj} - h_j x_{ii}\right|}, \tag{3a}$$

---

[10] By assumption, consumers prefer the home-supplied homogenous good in the case of identical prices. This implies a unique value of *Y*-trade in the absence of any trade friction for homogenous goods.

[11] Note that we consider f.o.b. trade flows (net of any iceberg transport costs) in eqs. (1a)-(3a) and throughout the rest of the theoretical analysis. This implies that $EX_{jk} = IM_{ik}$ (see Footnote 5). For a rigorous discussion on different concepts of the Grubel-Lloyd index, see Subsection 3.1.

For simplicity, we assume symmetry with respect to endowments[12] of *K* and *S* but allow for differences in endowments of unskilled labour *L*. Moreover, we assume the two economies are ex-ante equivalent with respect to cost parameter g, capturing physical capital related *FDI*-costs. Starting from this equilibrium we investigate how a marginal change in g_i (for given g_j) affects the IIT share GLI^C and assess the trade imbalance bias in relative terms by calculating the impact of g_i on *SHI*. Two scenarios can be distinguished:

Scenario I - $L_j < L_i$**:**[13]

Define $\tilde{x}_j := n_j x_{ii}$ and $\tilde{x}_i := n_i x_{jj}$. Then, using (11), (12) and E_i, P_i in (5) gives

$$\tilde{x}_j = \frac{\alpha}{\varepsilon} \frac{\left[\left(\frac{K-S}{g_i S - K} + \tau\right)\tilde{x}_i + (\varepsilon - 1) L_i\right](g_j S - K)}{\left(1 - \frac{\alpha}{\varepsilon}\right)(g_j - 1)S + (K - S) + \tau(g_j S - K)}, \tag{13}$$

and equivalently

$$\tilde{x}_i = \frac{\alpha}{\varepsilon} \frac{\left[\left(\frac{K-S}{g_j S - K} + \tau\right)\tilde{x}_j + (\varepsilon - 1) L_j\right](g_i S - K)}{\left(1 - \frac{\alpha}{\varepsilon}\right)(g_i - 1)S + (K - S) + \tau(g_i S - K)}. \tag{14}$$

From (13) and (14) it is obvious that $L_j < L_i$ implies $\tilde{x}_j\big|_{g_i = g_j} > \tilde{x}_i\big|_{g_i = g_j}$. Hence, we find[14]

$$GLI_{\text{SI}}^C = \frac{2\varepsilon\tau\tilde{x}_i}{(2\varepsilon\tau + h_j / n_j)\tilde{x}_j - (h_i / n_i)\tilde{x}_i - \left[(h_j / n_j)\tilde{x}_j - (h_i / n_i)\tilde{x}_i\right]} = \frac{\tilde{x}_i}{\tilde{x}_j}, \tag{15}$$

according to (2a), and

$$SHI_{\text{SI}} = 1 + \frac{h_j / n_j}{2\varepsilon\tau}\left(1 - \frac{h_i / n_i}{h_j / n_j}\frac{\tilde{x}_i}{\tilde{x}_j}\right), \tag{16}$$

according to (3a).

---

[12] These symmetry assumptions will be relaxed in the simulation analyses of Subsection 2.3.

[13] Remember our assumption that both sectors are active in the two countries. This requires that L_i and L_j are not too different.

[14] Index SI refers to **S**cenario **I**.

Result 1. *Consider* $L_j < L_i$ *and (ex ante)* $g_i = g_j$. *Then, a marginal increase of* g_i *(over* g_j*) raises the intra-industry trade share, i.e.* $dGLI_{\text{SI}}^{\text{C}} / dg_i > 0$, *and raises the trade imbalance bias in relative terms, i.e.* $dSHI_{\text{SI}} / dg_i > 0$.

Proof. See Appendix. ■

For $L_j < L_i$, an increase in g_i (for given g_j) makes the two economies "*more similar*", or in other words reduces country *i*'s home-market advantage due to its better endowment of *L*. It is well-known that the intra-industry trade share increases in the similarity of countries (see Helpman, 1987, Bergstrand, 1990, Hummels and Levinsohn, 1995), so that GLI^C increases in g_i. The aforementioned effect tends to reduce *SHI,* since the balance of repatriated profits, i.e. $\left(h_j / n_j\right)\tilde{x}_j - \left(h_i / n_i\right)\tilde{x}_i > 0$ becomes more equal, according to (15) and (16).[15] However, there is a second, counteracting effect. An increase in g_i reduces the number of country *i*'s horizontal multinationals (and increases its exporters). This lowers the flows of repatriated profits from *j* to *i* and, therefore, raises $\left(h_j / n_j\right)\tilde{x}_j - \left(h_i / n_i\right)\tilde{x}_i$ and stimulates the trade imbalance bias *SHI*. In sum, the firm number effect dominates and explains a negative impact of g_i on *SHI*. Or, put differently, if $L_j < L_i$ an increase of g_i, makes countries more similar in terms of their goods trade and therefore, raises GLI^C, but countries become more dissimilar in terms of their repatriated profits, which implies a higher *SHI*.

Scenario II - $L_j > L_i$**:**

From (13) and (14) it is clear that $L_j > L_i$ implies $\tilde{x}_j\big|_{g_i=g_j} < \tilde{x}_i\big|_{g_i=g_j}$. Hence, we find[16]

$$GLI_{\text{SII}}^{C} = \frac{2\varepsilon\tau\tilde{x}_j}{\left(2\varepsilon\tau + h_i / n_i\right)\tilde{x}_i - \left(h_j / n_j\right)\tilde{x}_j - \left[\left(h_i / n_i\right)\tilde{x}_i - \left(h_j / n_j\right)\tilde{x}_j\right]} = \frac{\tilde{x}_j}{\tilde{x}_i}, \tag{17}$$

according to (2a), and

$$SHI_{\text{SII}} = 1 + \frac{h_j / n_j}{2\varepsilon\tau}\left(\frac{h_i / n_i}{h_j / n_j} - \frac{\tilde{x}_j}{\tilde{x}_i}\right), \tag{18}$$

according to (3a).

---

[15] (One should keep in mind that repatriated profits are balanced if two economies are identical, implying $\sum_k EX_{ik} = \sum_k IM_{ik}$.)

[16] Index SII refers to **S**cenario **II**.

Result 2. *Consider* $L_j > L_i$ *and (ex ante)* $g_i = g_j$. *Then, a marginal increase of* g_i *(over* g_j*) reduces the intra-industry trade share, i.e.* $dGLI_{\text{SII}}^{\text{C}} / dg_i < 0$, *and lowers the trade imbalance bias in relative terms, i.e.* $dSHI_{\text{SII}} / dg_i < 0$.

Proof. See Appendix. ■

Under Scenario II, an increase in g_i reinforces *j*'s home-market advantage due to its better endowment of *L*. As a consequence, the dissimilarity between countries increases with g_i, which reduces the intra-industry trade share GLI^C. This stimulates *SHI,* since the balance of repatriated profits, i.e. $(h_i / n_i)\tilde{x}_i - (h_j / n_j)\tilde{x}_j > 0$ becomes less equal, according to (15) and (16). However, the induced decline in the number of country *i*'s horizontal multinational firms counteracts and dominates, so that $(h_i / n_i)\tilde{x}_i - (h_j / n_j)\tilde{x}_j$ declines, making countries more similar in terms of repatriated profits. This reduces *SHI*.

Vertical multinational enterprises

It is well-known from the literature that vertical multinationals (*v*) are more likely where countries differ sufficiently in their factor endowments or production technologies. In a two country model, vertical multinationals can only be active in one economy. We take the simplest possible framework that allows for vertical multinationals in country *i*, by assuming the following parameter constellation: $K_i > K_j = S_j = S_i$. Again, setting up an exporting firm requires one unit of capital and one of skilled labour; while one unit of skilled labour and $\gamma > 1$ units of capital are required for setting up a vertical multinational enterprise in country *i* with a single production plant in *j*.[17] In equilibrium, the zero profit conditions of exporters and vertical multinationals in *i* are given by[18]

$$\pi_{ni} = \frac{1}{\varepsilon - 1}\left[x_{ii} + \tau x_{jj} \right] - w_{Ki} - w_{Si} = 0, \qquad (19)$$

[17] We use γ instead of g to refer to the size of *FDI*-costs in the case of vertical multinational firms. The reason is that set-up costs of vertical multinationals fundamentally differ from set-up costs of horizontal multinationals, since in the former case only one production plant is required, while in the latter case two plants are operated.

[18] By assumption the endowments with unskilled labour are such that both the *X*-sector and the *Y*-sector are active in the two economies and that vertical multinationals as well as exporting firms survive in country *i*. Then, $w_{Li} = w_{Lj} = 1$, so that in this model vertical multinational activities are driven by a home-market effect (i.e. absolute size differences) and not by differences in unskilled wages.

$$\pi_{vi} = \frac{1}{\varepsilon - 1}\left[x_{jj} + \tau x_{ii}\right] - \gamma w_{K_i} - w_{Si} = 0, \tag{20}$$

respectively. (Note the similarity between (6) and (19).) In country *j* only exporting firms are active with profits

$$\pi_{nj} = \frac{1}{\varepsilon - 1}\left[x_{jj} + \tau x_{ii}\right] - w_{Kj} - w_{Sj} = 0. \tag{21}$$

The three factor market clearing conditions in country *i* are

$$L_i = n_i\left(x_{ii} + \tau x_{jj}\right) + Y_i, \tag{22}$$

$$S_i = n_i + v_i, \tag{23}$$

$$K_i = n_i + \gamma v_i. \tag{24}$$

And those in country *j* are

$$L_j = \left(n_j + v_i\right)\left(x_{jj} + \tau x_{ii}\right) + Y_j \tag{25}$$

$$K_j = S_j = n_j. \tag{26}$$

From (19), (20) and (22)-(24) we obtain[19]

$$w_{Ki} = \frac{1}{\varepsilon - 1}\frac{(1-\tau)\left(x_{jj} - x_{ii}\right)}{\gamma - 1}, \qquad w_{Si} = \frac{1}{\varepsilon - 1}\frac{(\tau\gamma - 1)x_{jj} + (\gamma - \tau)x_{ii}}{\gamma - 1} \tag{27}$$

and

$$n_i = \frac{\gamma K_j - K_i}{\gamma - 1}, \qquad v_i = \frac{K_i - K_j}{\gamma - 1} \tag{28}$$

for equilibrium wage rates and firm numbers in country *i*. Since only one firm type is active in *j*, we cannot distinguish between w_{Ki} and w_{Si}. Hence, equilibrium wages in *j* are given by

$$w_{Kj} + w_{Sj} = \frac{1}{\varepsilon - 1}\left[x_{jj} + \tau x_{ii}\right], \tag{29}$$

according to (21). The equilibrium firm number n_j is determined by (26).

Using $E_i = w_{Ki}K_i + w_{Si}K_j + L_i$, $P_i = p_{ii}^{1-\varepsilon}\left[n_i + \left(v_i + n_j\right)\tau\right]$ and $p_{ii} = \varepsilon/(\varepsilon - 1)$ in demand (5) as well as $E_j = \left(w_{Kj} + w_{Sj}\right)K_j + L_j$, $P_j = p_{jj}^{1-\varepsilon}\left[n_j + v_i + n_i\tau\right]$ and $p_{jj} = \varepsilon/(\varepsilon - 1)$ in the respective expression for country *j* gives after straightforward calculations explicit solutions

[19] $S_i = K_j$ is used in (27) and (28).

$$x_{ii} = \left(\frac{\alpha}{\varepsilon}\right)^2 (\gamma-1)(\varepsilon-1)\frac{\left[\left(K_i - K_j\right)+\tau\left(\gamma K_j - K_i\right)\right]L_j + \frac{\varepsilon}{\alpha}ML_i}{NM - \left(\frac{\alpha}{\varepsilon}\right)^2 (\gamma-1)\tau K_j\left[\left(K_i - K_j\right)+\tau\left(\gamma K_j - K_i\right)\right]}, \quad (30)$$

$$x_{jj} = x_{ii}\frac{(\gamma-1)\tau K_j L_i + \frac{\varepsilon}{\alpha}NL_j}{\left[\left(K_i - K_j\right)+\tau\left(\gamma K_j - K_i\right)\right]L_j + \frac{\varepsilon}{\alpha}ML_i}, \quad (31)$$

with $N := (1-\alpha/\varepsilon+\tau)(\gamma-1)K_j - (1-\alpha/\varepsilon)(1-\tau)\left(K_i - K_j\right)$, $M := (1-\alpha/\varepsilon+\tau)(\gamma-1)K_j$ $+(1-\tau)\left(K_i - K_j\right)$.

Fact 1. *Eqs. (30) and (31) are only consistent with positive wages* $w_{Ki} > 0$*, i.e. with* $x_{jj} > x_{ii}$*, according to* (27)*, if (i)* $\frac{\varepsilon}{\alpha}N > \left(K_i - K_j\right)+\tau\left(\gamma K_j - K_i\right)$ *and (ii)* $L_j > L_i$ *simultaneously hold.*

In the remainder of our analysis, we focus on positive wage equilibria with $w_{Ki} > 0$, i.e. sufficiently large[20] τ and L_j, according to Fact 1 and the definition of *N*. In addition $\tau\gamma > 1$ is sufficient for $w_{Si} > 0$.

For the case of vertical multinationals in country *i* we can rewrite the Grubel-Lloyd indices in (1) and (2) as:

$$GLI = \frac{2\varepsilon\tau \min\left[\left(n_j + v_i\right)x_{ii}, n_i x_{jj}\right]}{\varepsilon\tau\left[\left(n_j + v_i\right)x_{ii} + n_i x_{jj}\right] + \left|\varepsilon\tau\left[\left(n_j + v_i\right)x_{ii} - n_i x_{jj}\right] - v_i\left(\tau x_{ii} + x_{jj}\right)\right|} \quad (1b)$$

and

[20] Using the definition of *N* allows us to rewrite condition (i) of Fact 1 as $(\varepsilon/\alpha)\left\{\left[(\tau-\alpha/\varepsilon)+\tau(1-\alpha/\varepsilon)\right](\gamma-1)K_j + (1-\tau)\left(\gamma K_j - K_i\right)\right\} > 0$, which implies that $\tau > \alpha$ is sufficient for condition (i). Moreover, if condition (i) is fulfilled, then

$$\frac{L_j}{L_i} > \frac{(\varepsilon/\alpha)M - \left[\tau\left(K_i - K_j\right)+\tau\left(\gamma K_j - K_i\right)\right]}{(\varepsilon/\alpha)N - \left[\left(K_i - K_j\right)+\tau\left(\gamma K_j - K_i\right)\right]} > 1$$

guarantees $x_{jj} > x_{ii}$ and thus, $w_{Ki} > 0$ in equilibrium.

$$GLI^C = \frac{2\varepsilon\tau \min\left[\left(n_j + v_i\right)x_{ii}, n_i x_{jj}\right]}{\varepsilon\tau\left[\left(n_j + v_i\right)x_{ii} + n_i x_{jj}\right] + \left|\varepsilon\tau\left[\left(n_j + v_i\right)x_{ii} - n_i x_{jj}\right] - v_i\left(\tau x_{ii} + x_{jj}\right)\right| - v_i\left(\tau x_{ii} + x_{jj}\right)}, \quad (2b)$$

where $\left|\varepsilon\tau\left[\left(n_j + v_i\right)x_{ii} - n_i x_{jj}\right] - v_i\left(\tau x_{ii} + x_{jj}\right)\right|$ is *Y*-trade[21], according to the balance of payments condition. Moreover, $v_i\left(\tau x_{ii} + x_{jj}\right)$ are income flows from country *j* to country *i*, due to vertical multinational activities. According to (1b) and (2b), *SHI* simplifies to

$$SHI = 1 + \frac{v_i\left(\tau x_{ii} + x_{jj}\right)}{\varepsilon\tau\left[\left(n_j + v_i\right)x_{ii} + n_i x_{jj}\right] + \left|\varepsilon\tau\left[\left(n_j + v_i\right)x_{ii} - n_i x_{jj}\right] - v_i\left(\tau x_{ii} + x_{jj}\right)\right| - v_i\left(\tau x_{ii} + x_{jj}\right)}. \quad (3b)$$

Three scenarios can be distinguished:

Scenario I - $\left(n_j + v_i\right)x_{ii} < n_i x_{jj}$, country *i* is a net exporter of *X*-goods[22]

In this case, we obtain

$$GLI_{\mathrm{SI}}^C = \left(1 + \frac{v_i}{n_j}\right)\frac{n_j x_{ii}}{n_i x_{jj}} = \left(1 + \frac{2\left(K_i - K_j\right)}{\gamma K_j - K_i}\right)\frac{x_{ii}}{x_{jj}}, \quad (32)$$

according to (2b). Since there are income flows from country *j* to country *i*, the balance of payments condition requires that *j* exports homogenous good *Y* if $\left(n_j + v_i\right)x_{ii} < n_i x_{jj}$ holds. Moreover, according to (3b), we obtain

$$SHI_{\mathrm{SI}} = 1 + \frac{1}{2}\frac{v_i\left(\tau x_{ii} + x_{jj}\right)}{\varepsilon\tau n_i x_{jj}} = 1 + \frac{1}{2\varepsilon\tau}\frac{K_i - K_j}{\gamma K_j - K_i}\left(\tau\frac{x_{ii}}{x_{jj}} + 1\right). \quad (33)$$

This implies Result 3.

Result 3. *Consider $\left(n_j + v_i\right)x_{ii} < n_i x_{jj}$. Then, an increase of investment cost parameter γ has a negative impact on intra-industry trade, i.e. $dGLI_{\mathrm{SI}}^C / d\gamma < 0$. Moreover, a higher γ leads to a lower trade imbalance bias, i.e. $dSHI_{\mathrm{SI}} / d\gamma < 0$.*

---

[21] Consider Footnote 10 on our assumptions regarding *Y*-trade.

[22] One can define $\Omega := \left(n_j + v_i\right)/n_i - x_{jj}/x_{ii}$ to find $\partial\Omega/\partial K_i > 0$ and $\partial\Omega/\partial L_i = -\left(\partial\Omega/\partial L_j\right)\times\left(L_i/L_j\right) > 0$, according to (26)-(31) and Fact 1. Roughly spoken, this implies that Scenario I is more likely if K_i and L_i are not too high and L_j is not too low, motivating interesting interaction effects that are accounted for in the econometric analysis below.

Proof. See Appendix. ■

An increase in γ tends to make vertical multinational activities less attractive and therefore reduces *X*-imports of country *i*. This implies a reduction of the IIT share since *i* was already a net exporter of differentiated goods. The intuition for the *SHI*-effect is as follows. Remember that the difference between GLI and GLI^C arises due to the existence of vertical multinationals in *i*. However, multinational activities become less attractive if γ increases. As a consequence, an increase of γ reduces flows of repatriated profits from *j* to *i* and reduces the downward bias of intra-industry trade flows if GLI instead of GLI^C is used. This gives rise to $dSHI_{\text{SI}} / d\gamma < 0$.

Scenario II - Country *i* is a net importer of both goods[23]

In this case, we obtain

$$GLI_{\text{SII}}^C = \frac{2\varepsilon\tau n_i x_{jj}}{2\varepsilon\tau n_i x_{jj}} = 1, \tag{34}$$

according to (2b). Indeed, $n_i x_{jj} < \left(n_j + v_i\right) x_{ii}$ implies that country *i* is a net importer of the differentiated *X*-good. If country *i* also imports the homogenous good, there is no inter-industry trade since net imports of *i* are equal to repatriated profits due to multinational activities of country *i* firms in country *j*. Moreover, $SHI_{\text{SII}} = SHI_{\text{SI}}$, given by (33).

Result 4. *Consider* $\left(n_j + v_i\right) x_{ii} > n_i x_{jj}$ *and* $\varepsilon\tau\left[\left(n_j + v_i\right) x_{ii} - n_i x_{jj}\right] - v_i\left(\tau x_{ii} + x_{jj}\right) < 0$. *Then,* $GLI_{\text{SII}}^C = 1$, *so that a marginal change of* γ *has no impact on the intra-industry trade share. The impact of* γ *on* SHI_{SII} *is negative.*

Proof. First, use (34) to see that γ has no impact on GLI_{SII}^C. Second, $dSHI_{\text{SII}} / d\gamma < 0$ follows from Result 3. ■

The intuition for the SHI_{SII}-effect of γ is analogous to the intuition of the SHI_{SI}-effect discussed below Result 3.

---

[23] Country *i* is a net importer of both types of goods if both $n_i x_{jj} < \left(n_j + v_i\right) x_{ii}$ and $\varepsilon\tau\left[\left(n_j + v_i\right) x_{ii} - n_i x_{jj}\right] - v_i\left(\tau x_{ii} + x_{jj}\right) < 0$ simultaneously hold. Such an outcome is only possible if there are profit flows from country *j* to country *i*.

Scenario III - Country *i* is a net importer of *X*-goods and exports the *Y*-good[24]

In this case, both $\left(n_j+v_i\right)x_{ii}>n_i x_{jj}$ and $\varepsilon\tau\left[\left(n_j+v_i\right)x_{ii}-n_i x_{jj}\right]-v_i\left(\tau x_{ii}+x_{jj}\right)>0$ must simultaneously hold. Thus, we obtain

$$GLI_{\mathrm{SIII}}^{C}=\frac{\varepsilon\tau n_i x_{jj}}{\varepsilon\tau n_j x_{ii}+\left(\varepsilon-1\right)\tau v_i x_{ii}-v_i x_{jj}}=\frac{\varepsilon\tau n_i/v_i}{\varepsilon\tau\left(n_j/v_i\right)x_{ii}/x_{jj}+\left(\varepsilon-1\right)\tau x_{ii}/x_{jj}-1}, \quad (35)$$

according to (2b), and

$$\begin{aligned}SHI_{\mathrm{SIII}}&=1+\frac{1}{2}\frac{v_i\tau x_{ii}+v_i x_{jj}}{\varepsilon\tau n_j x_{ii}+\left(\varepsilon-1\right)\tau v_i x_{ii}-v_i x_{jj}}\\&=1+\frac{1}{2}\frac{v_i\tau x_{ii}/x_{jj}+v_i}{\varepsilon\tau n_j x_{ii}/x_{jj}+\left(\varepsilon-1\right)\tau v_i x_{ii}/x_{jj}-v_i}\end{aligned}, \quad (36)$$

according to (3b). The requirement of balanced payments guarantees $GLI_{\mathrm{SIII}}^{C}>0$ and $SHI_{\mathrm{SIII}}>0$.

Result 5. *Consider* $\varepsilon\tau\left[\left(n_j+v_i\right)x_{ii}-n_i x_{jj}\right]-v_i\left(\tau x_{ii}+x_{jj}\right)>0$. *Then, an increase of investment cost parameter* γ *raises the intra-industry trade share, i.e.* $dGLI_{\mathrm{SIII}}^{C}/d\gamma>0$. *Moreover, the impact of* γ *on the trade imbalance bias in relative terms is negative, i.e.* $dSHI_{\mathrm{SIII}}/d\gamma<0$, *if* τ *is sufficiently large.*[25]

Proof. See Appendix. ■

Again, an increase in γ makes vertical multinational activities less attractive. This implies a reduction of low-skilled labour in the production of differentiated goods in *j*. The resulting expansion of the *Y*-sector in *j* reduces imports of homogenous goods, leading to less inter-industry trade. Moreover, there is an increase (a decline) of country *i*'s differentiated goods exports to (imports from) country *j*. Both effects raise GLI_{SII}^{C}. The intuition for the SHI_{SIII}-effect is not straightforward. An increase in γ makes multinational activities less attractive, thereby reducing the flows of repatriated profits. This tends to reduce the trade imbalance bias and thus, *SHI* (see the intuition of Result 3). However, a decline in overall trade flows induces a higher weight of repatriated profits and increases *SHI*, according to (36). It is difficult to

[24] Scenario III is more likely if K_i and L_i are not too low and L_j is not too high, see Footnote 22.

[25] Remember our discussion below Fact 1. A sufficient condition for $dSHI_{\mathrm{SIII}}/d\gamma<0$ is derived in the Appendix.

determine which effect dominates. However, we can show that a negative impact of investment costs γ on *SHI* is guaranteed if τ is sufficiently large (see Appendix).

2.3 Simulation analysis

As a complement to our analytical results, we assess the impact of investment costs and determinants of IIT based on numerically solved versions of the models of vertical and horizontal multinationals. Given the inherent non-linearities of Dixit - Stiglitz type models in general and possible nonmonotonicities due to complementary slackness of general equilibrium models of trade and MNEs in particular (Markusen, 2002), we implement numerical solutions to yield insights into appropriate specification choice and robustness.

We simulate various versions of our model. In so doing, we stick to the notion that both the model of vertical MNEs (Helpman, 1984, Helpman and Krugman, 1985) and that of horizontal MNEs (Markusen, 1984, Markusen and Venables, 1998, 2000) are restricted variants of the knowledge-capital model, where both types of firms may endogenously arise (Carr, Markusen and Maskus, 2001, Markusen, 2002). However, a pure horizontal model and a pure vertical one are also calibrated. Altogether, we set up five different models: a KK-model based on a Leontief technology in the X-sector; a KK-model based on a Cobb-Douglas technology in the X-sector; a KK-model based on a CES-technology in the X-sector assuming a more realistic technical rate of substitution of between 0 and 1 (see Sharma, 2002; we choose a relatively low value of 0.1); a horizontal Leontief-based model; and a vertical Leontief-based model.[26]

In sum, we compute the equilibrium Grubel-Lloyd index for 21×21×21=9261 cells of the factor cube and 5 different levels of country *i*'s fixed FDI-related investment costs (country *j*'s investment costs are always set at a fixed value). This gives 46305 equilibrium values for each model without trade cost differences. Additionally, we simulate a set of equilibria, where trade costs for exports from country *i* to *j* amount to 5%-25%, leaving those of exports from *j* to *i* always at 15%. Where countries differ in trade costs, there are a further 4×9261=37044 equilibrium values. Pooling the two sets of equilibria allows us to search for the preferred specification in the empirical analysis, accounting for the same variables. Altogether, there are 83349 observations for each model. Specifically, we estimate the following models:

---

[26] Table A.1 in the Appendix provides details on the calibration of the model.

$$
\begin{aligned}
LGLI_{ij}^{C} &= \alpha_0 + \alpha_1 \ln(GDP_i + GDP_j) \\
&+ \alpha_2 \ln\left(1 - \left(GDP_i / \left(GDP_i + GDP_j\right)\right)^2 - \left(GDP_j / \left(GDP_i + GDP_j\right)\right)^2\right) \\
&+ \alpha_3 \ln | K_i / L_i - K_j / L_j | \quad + \alpha_4 \ln | (S_i / L_i - S_j / L_j) | \\
&+ \alpha_5 \ln | (1 + INVC_i) - (1 + INVC_j) | + \alpha_6 \ln \left| TC_{ij} - TC_{ji} \right| + \zeta_{ij}
\end{aligned} \tag{M1}
$$

$$
\begin{aligned}
LGLI_{ij}^{C} &= \beta_0 + \beta_1 \ln(GDP_i + GDP_j) + \beta_2 (GDP_i - GDP_j)^2 + \beta_3 (K_i / L_i - K_j / L_j)^2 \\
&+ \beta_4 (S_i / L_i - S_j / L_j)^2 + \beta_5 (GDP_i - GDP_j)^2 \times (S_i / L_i - S_j / L_j)^2 \\
&+ \beta_6 \ln\left[0.5\left(1 + INVC_i\right) + 0.5\left(1 + INVC_j\right)\right] + \beta_7 \ln\left(0.5 TC_{ij} + 0.5 TC_{ji}\right) + \zeta_{ij}
\end{aligned} \tag{M2}
$$

$$
\begin{aligned}
LGLI_{ij}^{C} &= \chi_0 + \chi_1 \max\{\ln GDP_i, \ln GDP_j\} + \chi_2 \min\{\ln GDP_i, \ln GDP_j\} \\
&+ \chi_3 \ln | (K_i / L_i - K_j / L_j) | \quad + \chi_4 \ln | (S_i / L_i - S_j / L_j) | \\
&+ \chi_5 \ln | (1 + INVC_i) - (1 + INVC_j) | + \chi_6 \ln\left(0.5 TC_{ij} + 0.5 TC_{ji}\right) + \zeta_{ij}
\end{aligned} \tag{M3}
$$

$$
\begin{aligned}
LGLI_{ij}^{C} &= \delta_0 + \delta_1 \max\{\ln GDP_i, \ln GDP_j\} + \delta_2 \min\{\ln GDP_i, \ln GDP_j\} \\
&+ \delta_3 \max\{\ln(K_i / L_i), \ln(K_j / L_j)\} + \delta_4 \min\{\ln(K_i / L_i), \ln(K_j / L_j)\} \\
&+ \delta_5 \max\{\ln(S_i / L_i), \ln(S_j / L_j)\} + \delta_6 \min\{\ln(S_i / L_i), \ln(S_j / L_j)\} \\
&+ \delta_7 \max\{\ln(1 + INVC_i), \ln(1 + INVC_j)\} + \delta_8 \min\{\ln(1 + INVC_i), \ln(1 + INVC_j)\} \\
&+ \delta_9 \max\left\{\ln\left(TC_{ij}\right), \ln\left(TC_{ji}\right)\right\} + \delta_{10} \min\left\{\ln\left(TC_{ij}\right), \ln\left(TC_{ji}\right)\right\} + \zeta_{ij}
\end{aligned} \tag{M4}
$$

$LGLI_{ij}^{C}$ denotes the logistically transformed, corrected Grubel-Lloyd index, $INVC$ refers to investment costs g and γ of our theoretical analysis, respectively, and TC_{ji} (TC_{ij}) is a measure of transport costs for shipping differentiated goods from country j (i) to country i (j).[27]

M1 is closest to Helpman (1987) but with the addition of differences in investment costs; M2 is closest to Markusen and Maskus (2001), extended by the squared difference in capital-labour ratios; M3 is in the spirit of Hummels and Levinsohn (1995), with the addition of absolute difference in skilled-to-unskilled labour ratios; and M4 extends their idea of allowing for asymmetric influences between maximum and minimum levels of all variables.

[27] In terms of our analytical model, $TC_{ij} = \left(2 - 1/t_{ij}\right)$ gives the volume of production that is necessary if one unit of the differentiated good is consumed abroad.

Running those four specifications results in the following adjusted R^2 figures:

	Horizontal	Vertical	KK-Leontief	KK-CD	KK-CES
M1	0.0115	0.2345	0.1469	0.0082	0.0568
M2	0.0685	0.3507	0.1419	0.0569	0.0624
M3	0.0113	0.3157	0.1147	0.0113	0.0557
M4	0.1056	0.5085	0.1563	0.0933	0.0880

With the exception of the vertical model, the reported R^2 figures are relatively low, reflecting the high degree of non-linearity in these type of models. Of course, omitting the skills and friction variables in M1 or M3 would lead to specifications which are closer to Helpman (1987) and Hummels and Levisohn (1995), but inferior in terms of explanatory power. Similarly, omitting the capital terms in M2 would render the model closer to Markusen and Maskus (2001) but also inferior. On the other hand, using the maximum and minimum values of both trade and investment frictions in every model reduces the difference in adjusted R^2 figures, but without changing their ranking. Empirically, the repeated observation of each country in a bilateral setting and the use of country-specific effects improve the fit.

As can be seen, M4 consistently outperforms M1-M3. With regard to the estimated coefficients of M4, two weak hypotheses can be formulated. First, $\delta_1<0$ and $\delta_2<0$ are more likely if horizontal MNEs dominate[28] Note that horizontal MNE activity is market-seeking, i.e. growing with market size, and crowds out two-way trade in differentiated goods, explaining the expected sign of δ_1. The result $\delta_2<0$ is due to the non-linearities, caused by complementary slackness. Suppose that there is initially a very small country, so that it does not pay to set up horizontal MNEs. In such an economy, intra-industry trade accounts for a large share of trade. As we reallocate absolute factor endowments to this economy, at some point it is profitable to establish horizontal MNEs and intra-industry trade falls. As countries become more similar, the IIT share rises again. In our case, the effect induced by the complementary slackness dominates and δ_2 is negative. However, $\delta_1<0$ indicates that similarity in size is important and tends to increase the IIT share if δ_1 dominates δ_2. Vertical MNEs tend to foster intra-industry trade but they are stimulated by dissimilarities in country size, in line with our analytical investigation (see Fact 1).

[28] See Markusen and Maskus (2000) and Carr et al. (2001) for strong empirical support of horizontal MNEs.

Second, for similar reasons the maximum investment cost coefficient δ_7 is negative with two-way horizontal FDI (see Result 2). Hence, the share of IIT tends to rise with the similarity in investment costs. Finally, the difference between the maximum value and the corresponding minimum value of the skilled to unskilled labour coefficients ($\delta_5 - \delta_6$) tends by and large to be negative, which supports the common finding that IIT is higher between economies with more similar factor endowments (Helpman, 1987, Bergstrand, 1990, Hummels and Levinsohn, 1995). With regard to the impact of capital-labour ratios, remember that setting up horizontal MNEs is the most capital intensive activity. Horizontal MNEs seem empirically important and note that the share of intra-industry trade in total imbalance-corrected trade tends to rise if horizontal FDI increases. However, with co-existing horizontal and vertical or only vertical FDI the impact of capital-labour ratios gets less clear-cut.

2.4 Summary of the theoretical hypotheses for the Grubel-Lloyd index

From our analytical investigation we obtain the following hypotheses. With horizontal MNEs an increase in investment costs *g* tends to reduce GLI^C, if *g* increases in the *L*-abundant country. In contrast, if *g* rises in the country with scarce *L* supply, GLI^C increases. With vertical MNEs, an increase in *FDI*-costs tends to reduce the intra-industry trade share if in the country that hosts the vertical multinational firms factor *L* is relatively scarce (so that this economy is a net exporter of the differentiated good. In contrast, if the country that hosts the vertical multinationals is relatively *L*-abundant and, therefore, is a net importer of the differentiated good, GLI^C tends to be positively (non-negatively) affected by an increase in FDI-costs. The simulation exercise generates two additional hypotheses. First, the country size (δ_1,δ_2) coefficients are more likely negative, if horizontal MNEs dominate at a reasonable value of the elasticity of substitution (see Feenstra, 1994, for detailed empirical evidence). Second, intra-industry trade is by and large higher between economies with more similar skilled-to-unskilled labour endowments and investment costs with two-way horizontal FDI, but less likely the more important vertical FDI is.

3 Empirical analysis

3.1 The Grubel-Lloyd index in the empirical trade literature

Grubel and Lloyd (1971) had in mind a model with zero transport costs and no multinational firms. Both transport costs and MNE activity are now understood as essential characteristics of international exchange. However, their consequences for the measurement (and determinants) of intra-industry trade shares has to the best of our knowledge not yet been rigorously studied. Below, we provide several alternative versions of the Grubel-Lloyd index, which can cope with both transport costs and MNE activity. (We also explicitly discuss issues such as the interpretation of missing values in the disaggregated trade data, the index is based on.) Table 1 summarizes.[29]

> Table 1 <

It seems sensible to start with the original formulation of the index as also applied in Helpman (1987), Hummels and Levinsohn (1995), or Markusen and Maskus (2001). In the case of a two-country, new trade theory model with zero transport costs and no MNE activity, $GLI \equiv GLI_1^C \equiv GLI_2^C \equiv GLI_3^C \equiv GLI_4^C \equiv GLI_5^C$.

With multinational firms, trade is not necessarily (or even likely to be) balanced. To see the relevance of this, consider the simple thought experiment of two one-sector economies with MNEs. Not accounting for income flows due to repatriated profits leads to a downward bias of the Grubel-Lloyd index, i.e. $GLI - GLI_1^C < 0$, $GLI_2^C - GLI_3^C < 0$ and $GLI_4^C - GLI_5^C < 0$, which we refer to as the **trade imbalance bias** in absolute terms (in contrast to the relative measure of this bias, *SHI*, calculated above).[30] Hence, there remain three candidates for measuring the intra-industry trade share: GLI_1^C, GLI_3^C and GLI_5^C which differ if transport costs are positive.

Now consider the impact of transport costs, but stick for the moment to the usual assumption that $t_{ij} = t_{ji}$. In this case, $GLI_1^C \neq GLI_3^C \equiv GLI_5^C$. Note that $GLI_1^C \neq GLI_3^C$, because the denominator of GLI_1^C is higher than the denominator of GLI_3^C due to transport costs included

[29] The Grubel-Lloyd indices in Table 1 measure bilateral intra-industry trade in a multi-country world. Hence, EX_{ij} are country *i*'s exports to and IM_{ij} are country *i*'s imports from country *j*. Index *k* indicates different industries.

[30] The arguments in Greenaway et al. (2001) are related to our arguments. Bergstrand (1983) correctly points out that bilateral trade tends to be unbalanced also in a multilateral setting without MNEs. Our approach also covers this phenomenon.

in c.i.f. imports IM_{ij} but not in f.o.b. exports EX_{ij}. But also the numerator of GLI_1^C is higher unless $EX_{ijk} < IM_{ijk} \quad \forall k$. This **transport cost level bias** in absolute terms $\left|GLI_1^C - GLI_3^C\right|$ appears to be a non-linear function of *t*. Moreover, GLI_3^C and GLI_5^C share an advantage over GLI_1^C, since they lead to the same index values for the two economies (i.e., $GLI_{ij} \equiv GLI_{ji}$), while $GLI_{1,ij}^C \neq GLI_{1,ji}^C$ if transport costs are larger than zero.[31] In addition, if transport costs differ, i.e. if $t_{ij} \neq t_{ji}$, which is the empirically relevant case, also GLI_3^C differs from GLI_5^C. An approximation of this **transport cost difference bias** in absolute terms is $\left|GLI_3^C - GLI_5^C\right|$.

We focus on two versions of the corrected Grubel-Lloyd index in the empirical analysis: we use GLI_3^C as our preferred measure, since it is derived from our theoretical model and gives identical indices for the two economies. We also use GLI_1^C, since it is closest to the idea of the traditional Grubel-Lloyd index *GLI*, but avoiding the trade imbalance bias.

Furthermore, there is a **missing value interpretation bias** at the most disaggregated level (*k*). There are two opportunities to handle this. One could interpret them as "missing" in a narrow sense and skip all ℓ missing observations, before determining the minimum export and re-export (from partner statistics; or import values). Suppose that the data are sorted so that all ℓ missing observations come first and that $EX_{ji1},...,EX_{ji\ell}$ is missing with the true $0 \leq EX_{jik} < EX_{ijk} \quad \forall k \leq \ell$. If we ignore any trade imbalance for the moment and use $AGLI_3^C$ to assess the share of intra-industry trade, then[32] $\sum_{k=1}^{\ell} \frac{\left|EX_{ijk} - EX_{jik}\right|}{\sum_{k=1}^{\ell}\left(EX_{ij} + EX_{ji}\right)} > \sum_{k=\ell+1}^{K} \frac{\left|EX_{ijk} - EX_{jik}\right|}{\sum_{k=\ell+1}^{K}\left(EX_{ij} + EX_{ji}\right)}$ implies a missing value interpretation downward bias. Otherwise, $AGLI_3^C$ is upward biased.

[31] In our sample, the difference between $GLI_{1,ij}^C$ and $GLI_{1,ji}^C$ amounts to 43 percentage points for 5-digit-based data, which is about 312% of the corresponding mean.

[32] The same problems arise if $AGLI_1^C$ instead of $AGLI_3^C$ were used.

The alternative is to replace missing values by zeros. However, if $EX_{jik} = 0$ is used, although the true value were $0 < EX_{jik} < EX_{ijk} \quad \forall k \leq \ell$, there is a downward bias in GLI_3^C. In addition, the magnitude of trade imbalance may be biased under both correction methods. Since no information on the true values is available, we have to rely on assumptions to favor one approach over the other. Here, we stick to the working hypothesis that, on average, the true missing values are very small, if not zero.Summing up, we can label indices GLI_1^C and, especially, GLI_3^C and not $AGLI_1^C$ and $AGLI_3^C$ as the preferred measures.[33]

3.2 Econometric analysis

We estimate our M4 model on various concepts of the Grubel-Lloyd index, focusing on GLI as the traditional measure and GLI_1^C, GLI_3^C as our preferred measures. Moreover, we also account for $AGLI$, $AGLI_1^C$ and $AGLI_3^C$ as a robustness check on missing values.

> Table 2 <

Our data base comprises 422 observations of 1990-2000 bilateral average IIT share data of OECD countries for GLI_3^C and $AGLI_3^C$ after excluding missing values, while there are about twice as many observations for GLI, $AGLI$ and GLI_1^C, $AGLI_1^C$, due to their asymmetry between the *i*-to-*j* and the *j*-to-*i* trade flow definition. (A detailed data description is in Appendix D.) Table 2 summarizes our findings. First, those variables not usually considered in empirical analysis but motivated by our theoretical analysis (δ_5 - δ_{10}) account for 41%-69% of the regression models' explanatory power. This again emphasises the relevance of the MNE-related new trade theory literature for core empirical issues of international trade. Second, in line with the regressions on the data obtained from the numerically solved models with two-way horizontal FDI, coefficients for a variable's bilateral maximum and minimum value with the exception of country size (GDP) and capital-labour ratios tend to be different in terms of both their sign and absolute value. By and large, the evidence suggests that similarity

[33] For completeness, we should mention the so-called **Finger bias** which refers to the problem of potentially upward biased intra-industry trade figures due to the use of higher aggregated data than available and possible **statistical measurement bias** due to false reporting by the national statistical offices. A detailed empirical assessment of the size of these is presented in Appendix C.

(though in a non-linear way) in skilled-to-unskilled labour endowments and in both trade and investment impediments is in favor of more intra-industry trade in total trade.[34] Also the capital-labour ratio coefficients lends support to the horizontal model. Only the strong positive sign of the minimum GDP coefficient contradicts the simulated, purely horizontal model of FDI. Third, investment cost effects are mainly represented by the negative, significant coefficient of maximum bilateral investment costs, which is well in line with our theoretical findings. In the analytical model we find a negative impact of FDI-costs on the corrected Grubel-Lloyd index for some factor endowments (see Results 2 and 3).

It turns out that in the regressions based on the trade-imbalance uncorrected measures of the intra-industry trade share (*GLI*, *AGLI*) the role of transport costs seems to be over-estimated at the expense of relative factor endowments, as compared to their corrected counterparts. If missing observations at the disaggregated level really reflect very low values of trade rather than confidential information from a few, large firms' perspective, the results suggest that skipping missing values results in a downward bias (in absolute terms) of the effects of capital-labour ratios, maximum investment costs, and transport costs (to see this compare the coefficients for $AGLI$, $AGLI_1^C$, $AGLI_3^C$ with the corresponding ones to their left). Hence, our findings illustrate that measurement biases in IIT share indices do not only affect the mean (picked up by the constant), but there is some systematic bias, which is correlated with the most important explanatory variables. In sum, the results are well in line with the model of horizontal two-way FDI, but they lend less support to the existence of vertical FDI, irrespective of whether we use the preferred or the biased indices. This is not surprising given the composition of our country sample.

3.3 Sensitivity analysis

We check the sensitivity of our results with respect to the exclusion of extreme outliers and inclusion of exporter and importer fixed effects. With regard to outliers, we follow Belsley et al. (1980) and exclude all observations with absolute residuals exceeding two standard errors of the regression. On average, only 2% of observations have to be eliminated. Fixed country-specific effects are able to control for all other unobserved variables, especially, multilateral trade resistance terms in a multi-country setting (see Anderson and van Wincoop, 2003). Note that the parameters of the variables can still be estimated, since by definition there is enough

[34] This similarity aspect is also in line with our findings for the analytically solvable horizontal multinational

variation in maximum and minimum values. More precisely, in a sample such as ours it is impossible that each country exhibits the maximum or minimum value in all variables with respect to all included trading partners.

> Tables 3 and 4 <

Comparing the results across columns in the upper part of Table 3 confirms that the Finger-bias in our country sample tends not to systematically bias the parameter estimates. This conclusion is based on results that exclude extreme outliers but relies on the assumption that there are no omitted country-specific influences, which are correlated with the explanatory variables. The inclusion of country effects tends to reduce the collinearity among regressors and controls for all omitted country-specific influences, which may otherwise be picked up by the parameters of interest (see Baltagi, 2001). This has two important consequences. First, both size coefficients are now more supportive of the two-way horizontal MNE model than of its vertical counterpart (see the first two columns in the lower part of Tables 3 and 4, respectively). Second, the fixed effects model parameters tend to be much more sensitive to the impact of the Finger bias or the missing variable interpretation bias than their inconsistent counterparts. To see this, compare the columns in the lower part of Table 3 and note that the first comprises the preferred specification if our assumption about missing values holds. The Finger bias even changes the sign of δ_5 if we compute the GLI_1^C or GLI_3^C on the basis of 4-digit (3-digit) rather than of 5-digit data.[35]

3.4 Extensions

Here we provide insights into two additional issues: the role of differences in endowments with labour and physical capital, respectively, and the impact of investment costs on the ratio of the trade-imbalance corrected-to-uncorrected Grubel-Lloyd indices *SHI*.

> Table 5 <

---

firms model in Subsection 2.2. See our discussion below Results 1 and 2.

[35] Note that the reported F-tests on the parameters indicate that, by and large, using a simple measure of similarity or also the average of bilateral size, factor endowments, and trade and investment impediments is inferior to the chosen strategy of including each variable's bilateral maximum and minimum value separately.

Our analytical results suggest that the impact of an increase in investment costs on GLI^C is more likely to be positive the larger L_i is compared to L_j. To assess this, we construct an interaction term between the difference of maximum and minimum log investment costs and the corresponding log-difference in *L* (see Table 5). According to our theoretical results for two-way horizontal multinationals, we expect a negative sign of the maximum investment cost effect[36] (δ_7) but a positive one of the interaction term (δ_{11}).

As with the case of labour endowment differences, we formulate an interaction term between the impact of investment costs and differences in the endowment with physical capital.[37] Given all other endowments, our model suggests that maximum minus minimum investment costs are likely to have a positive impact on the intra-industry trade share, the larger the corresponding difference in capital endowments. Again, we expect a positive sign for the interaction term δ_{12}.

As the point estimates in Table 5 indicate, our empirical findings strongly support our theoretical hypotheses, irrespective of which of the preferred *GLI* concepts is used. However, one caveat applies. It is impossible to include simultaneously both interaction terms in the specifications. The reason is that capital stock levels are large in countries with large labour forces. Hence, size differences strongly dominate relative factor endowment differences in our sample, rendering the log difference in respective capital stocks and that in absolute labour endowments highly collinear.

> Table 6 <

Regarding *SHI*, we know that this ratio should fall with the difference between maximum and minimum foreign investment costs, in particular, if the country with the maximum investment

---

[36] Compare the findings of the simulation analyses in Subsection 2.3 and the summary of our theoretical hypotheses in Subsection 2.4.

[37] This interaction term is motivated by our analytical investigation in Subsection 0 for the case of vertical multinational firms, see Footnotes 22 and 24. Unfortunately, there is no comparable prediction for such an interaction term if horizontal multinationals are considered. This is due to our symmetry assumptions in Subsection 0.

costs is less well endowed with labour than its counterpart.[38] This is investigated in Table 6 for the two preferred concepts of the Grubel-Lloyd index. The results offer two insights. First, the point estimates of both effects exhibit the expected signs. Second, country-specific effects are important, indicating that bilateral trade-imbalances are a common phenomenon. Third, we have to concede that investment costs explain a relatively small though significant share of the deviation between the two indices as indicated by the R^2 figures. The other explanatory variables used in the previous tables only contribute insignificantly. Hence, other macro-economic variables, not accounted for in the above theoretical model and the empirical specifications are probably relevant in this regard. However, to study their impact is beyond the scope of this analysis.

4 Conclusions

In a review of the empirical analysis of international trade flows spanning the last 50 years, Leamer (1994) identifies "the extensive amount of intra-industry trade catalogued by Grubel and Lloyd (1975)….." as "….. one of the only two major empirical findings (which) seem to have had a major impact on the way (trade) economists think" (p.68). That conclusion would no doubt be revised in light of the growing influence of the firm level adjustment literature. Be that as it may, Leamer's conclusion articulates a widely accepted view that the apparent pervasiveness of intra-industry trade stimulated a revolution in the theoretical and empirical modelling of international trade.

From the standpoint of empirical investigation, it is obviously vital that the intra-industry trade share is measured as accurately as possible. Thirty years after the publication of Grubel and Lloyd (1975), their famous index remains the measure of choice for most investigators. Yet we know that it is grounded in the assumption of arms length trade but multinational activity is a feature of the landscape which should not be ignored. In this paper we have brought their presence to centre-stage. We have constructed a three factor general equilibrium model of trade with both horizontal and vertical multinationals, to identify precisely the impact of investment costs and multinational activity on intra-industry trade.

[38] For the case of vertical multinationals, Results 3, 4 and 5 predict a negative impact of investment costs on *SHI*. Moreover, as far as horizontal multinationals are considered, Result 2 shows that the *SHI*-effect is negative if $L_j > L_i$ so that the Grubel-Llyod index GLI^C declines in the investment cost parameterm.

The model and the measures of intra-industry trade derived from it have been subjected to extensive simulation analysis and rigorous econometric analysis. The latter focuses on the trade flows of 31 OECD countries.

Our analysis demonstrates clearly the role of investment costs and the biases inherent in the Grubel-Lloyd index when we fail to account for the presence of multinationals. Our econometric analysis confirms the superiority of our new corrected measures. It also shows that it is important to account for various new determinants of IIT alongside more traditional explanatory variables. Finally, our analysis lends further support to the relative importance of horizontal multinationals. We hope that the theoretical underpinning provided for our new measures and their robust empirical performance will commend their wider use.

Appendix

A. Analytical appendix

Proof of Result 1

We define

$$\Gamma_j^h := \frac{\alpha}{\varepsilon}\frac{\left[\left(\frac{K-S}{g_i S-K}+\tau\right)\tilde{x}_i+(\varepsilon-1)L_i\right](g_j S-K)}{\left(1-\frac{\alpha}{\varepsilon}\right)(g_j-1)S+(K-S)+\tau(g_j S-K)}-\tilde{x}_j=0, \tag{A1}$$

according to (13), and

$$\Gamma_i^h := \frac{\alpha}{\varepsilon}\frac{\left[\left(\frac{K-S}{g_j S-K}+\tau\right)\tilde{x}_j+(\varepsilon-1)L_j\right](g_i S-K)}{\left(1-\frac{\alpha}{\varepsilon}\right)(g_i-1)S+(K-S)+\tau(g_i S-K)}-\tilde{x}_i=0, \tag{A2}$$

according to (14). Eqs. (A1) and (A2) imply system

$$\begin{aligned}\frac{\partial\Gamma_j^h}{\partial\tilde{x}_j}\frac{d\tilde{x}_j}{dg_i}+\frac{\partial\Gamma_j^h}{\partial\tilde{x}_i}\frac{d\tilde{x}_i}{dg_i}+\frac{\partial\Gamma_j^h}{\partial g_i}&=0\\ \frac{\partial\Gamma_i^h}{\partial\tilde{x}_j}\frac{d\tilde{x}_j}{dg_i}+\frac{\partial\Gamma_i^h}{\partial\tilde{x}_i}\frac{d\tilde{x}_i}{dg_i}+\frac{\partial\Gamma_i^h}{\partial g_i}&=0.\end{aligned} \tag{A3}$$

Straightforward calculations allow us to write $\left.\frac{\partial\Gamma_j^h}{\partial\tilde{x}_j}\right|_{g_i=g_j}=\left.\frac{\partial\Gamma_i^h}{\partial\tilde{x}_i}\right|_{g_i=g_j}=-1$, $\left.\frac{\partial\Gamma_j^h}{\partial\tilde{x}_i}\right|_{g_i=g_j}=\left.\frac{\partial\Gamma_i^h}{\partial\tilde{x}_j}\right|_{g_i=g_j}=\frac{\alpha}{\varepsilon}\frac{(K-S)+\tau(gS-K)}{B}<1$ and $\left.\frac{\partial\Gamma_j^h}{\partial g_i}\right|_{g_i=g_j}=-\frac{\alpha}{\varepsilon}\frac{K-S}{gS-K}\frac{S\tilde{x}_i}{B}<0$, $\left.\frac{\partial\Gamma_i^h}{\partial g_i}\right|_{g_i=g_j}=-\frac{2-\alpha/\varepsilon}{\alpha/\varepsilon}\times\left.\frac{\partial\Gamma_j^h}{\partial g_i}\right|_{g_i=g_j}>0$, where $B:=\left(1-\frac{\alpha}{\varepsilon}\right)(g-1)S+(K-S)+\tau(gS-K)$ and $g\equiv g_i=g_j$ have been used. Applying Cramer's rule to system (A3), we therefore obtain

$$\left.\frac{d\tilde{x}_j}{dg_i}\right|_{g_i=g_j}=\frac{\frac{\partial\Gamma_i^h}{\partial g_i}\frac{\partial\Gamma_j^h}{\partial\tilde{x}_i}-\frac{\partial\Gamma_j^h}{\partial g_i}\frac{\partial\Gamma_i^h}{\partial\tilde{x}_i}}{\left(1+\frac{\partial\Gamma_j^h}{\partial\tilde{x}_i}\right)\left(1-\frac{\partial\Gamma_j^h}{\partial\tilde{x}_i}\right)}=\frac{\frac{\partial\Gamma_j^h}{\partial g_i}+\frac{\partial\Gamma_i^h}{\partial g_i}\frac{\partial\Gamma_j^h}{\partial\tilde{x}_i}}{\left(1+\frac{\partial\Gamma_j^h}{\partial\tilde{x}_i}\right)\left(1-\frac{\partial\Gamma_j^h}{\partial\tilde{x}_i}\right)}<0, \tag{A4}$$

$$\left.\frac{d\tilde{x}_i}{dg_i}\right|_{g_i=g_j} = \frac{\frac{\partial\Gamma_j^h}{\partial g_i}\frac{\partial\Gamma_j^h}{\partial\tilde{x}_i} - \frac{\partial\Gamma_j^h}{\partial\tilde{x}_j}\frac{\partial\Gamma_i^h}{\partial g_i}}{\left(1+\frac{\partial\Gamma_j^h}{\partial\tilde{x}_i}\right)\left(1-\frac{\partial\Gamma_j^h}{\partial\tilde{x}_i}\right)} = \frac{\frac{\partial\Gamma_i^h}{\partial g_i} + \frac{\partial\Gamma_j^h}{\partial g_i}\frac{\partial\Gamma_j^h}{\partial\tilde{x}_i}}{\left(1+\frac{\partial\Gamma_j^h}{\partial\tilde{x}_i}\right)\left(1-\frac{\partial\Gamma_j^h}{\partial\tilde{x}_i}\right)} > 0\,, \tag{A5}$$

according to (A3).

Next, we differentiate $GLI_{\mathrm{SI}}^C = \tilde{x}_i / \tilde{x}_j$, according to (15), with respect to g_i and obtain

$$\left.\frac{dGLI_{\mathrm{SI}}^C}{dg_i}\right|_{g_i=g_j} = \frac{1}{\tilde{x}_j^2}\left[\frac{d\tilde{x}_i}{dg_i}\tilde{x}_j - \frac{d\tilde{x}_j}{dg_i}\tilde{x}_i\right] > 0\,, \tag{A6}$$

which is positive, according to (A4) and (A5).

To determine the impact of g_i on *SHI*, we differentiate (16) with respect to g_i. This gives

$$\left.\frac{dSHI_{\mathrm{SI}}}{dg_i}\right|_{g_i=g_j} = \frac{h_j/n_j}{2\varepsilon\tau}\frac{\tilde{x}_i}{\tilde{x}_j}\left[\frac{S}{gS-K} - \frac{1}{\tilde{x}_i}\left(\frac{d\tilde{x}_i}{dg_i} - \frac{d\tilde{x}_j}{dg_i}\frac{\tilde{x}_i}{\tilde{x}_j}\right)\right]. \tag{A7}$$

Substituting $\left.\frac{\partial\Gamma_j^h}{\partial\tilde{x}_i}\right|_{g_i=g_j} = \frac{\alpha}{\varepsilon}\frac{(K-S)+\tau(gS-K)}{B}$, $\left.\frac{\partial\Gamma_j^h}{\partial g_i}\right|_{g_i=g_j} = -\frac{\alpha}{\varepsilon}\frac{K-S}{gS-K}\frac{S\tilde{x}_i}{B}$ and $\left.\frac{\partial\Gamma_i^h}{\partial g_i}\right|_{g_i=g_j} = -\frac{2-\alpha/\varepsilon}{\alpha/\varepsilon}\times\left.\frac{\partial\Gamma_j^h}{\partial g_i}\right|_{g_i=g_j}$ in (A4) and (A5) it follows that the bracket expression on the right-hand side of (A7) is strictly decreasing in $\tilde{x}_i/\tilde{x}_j$, according to (A4) and (A5). Thus, $dSHI_{\mathrm{SI}}/dg_i|_{g_i=g_j}$ is positive for all $\tilde{x}_j|_{g_i=g_j} > \tilde{x}_i|_{g_i=g_j}$, if it is positive for $\tilde{x}_j = \tilde{x}_i$. We therefore, calculate

$$\left.\frac{1}{\tilde{x}_i}\left(\frac{d\tilde{x}_i}{dg_i} - \frac{d\tilde{x}_j}{dg_i}\right)\right|_{g_i=g_j} = \frac{1}{\tilde{x}_i}\frac{\partial\Gamma_i^h/\partial g_i - \partial\Gamma_j^h/\partial g_i}{1+\partial\Gamma_j^h/\partial\tilde{x}_i} = 2\left(1-\frac{\alpha}{\varepsilon}\right)\frac{K-S}{B}\frac{S}{gS-K}\,, \tag{A8}$$

according to (A4), (A5) and our considerations above. Since $2(1-\alpha/\varepsilon)(K-S)/B<1$, it follows that $dSHI_{\mathrm{SI}}/dg_i|_{g_i=g_j} > 0$ for all possible $\tilde{x}_j|_{g_i=g_j} > \tilde{x}_i|_{g_i=g_j}$, since $dSHI_{\mathrm{SI}}/dg_i|_{g_i=g_j} > 0$ for $\tilde{x}_j|_{g_i=g_j} = \tilde{x}_i|_{g_i=g_j}$. This completes the proof of Result 1. ■

Proof of Result 2

First, note that $dSHI_{\text{SII}} / dg_i\big|_{g_i=g_j} < 0$ directly follows from (17), (A4) and (A5). Second, regarding the impact of g_i on SHI_{SII} we calculate

$$\left.\frac{dSHI_{\text{SII}}}{dg_i}\right|_{g_i=g_j} = -\frac{h_j / n_j}{2\varepsilon\tau}\left[\frac{S}{gS-K} - \frac{1}{\tilde{x}_i}\left(\frac{d\tilde{x}_i}{dg_i}\frac{\tilde{x}_j}{\tilde{x}_i} - \frac{d\tilde{x}_j}{dg_i}\right)\right], \tag{A9}$$

according to (18). The right hand side of (A9) is strictly increasing in $\tilde{x}_j / \tilde{x}_i$, according to (A4) and (A5). (For details see the proof of Result 1.) Hence, $dSHI_{\text{SII}} / dg_i\big|_{g_i=g_j}$ is negative for all $\tilde{x}_j\big|_{g_i=g_j} < \tilde{x}_i\big|_{g_i=g_j}$, if it is negative for $\tilde{x}_j = \tilde{x}_i$. This follows immediately from the proof of Result 1 and completes the proof of Result 2. ■

Proof of Result 3

We use the definitions of *M* and *N* and differentiate

$$\frac{x_{ii}}{x_{jj}} = \frac{\left[\left(K_i - K_j\right) + \tau\left(\gamma K_j - K_i\right)\right]L_j + \left(\varepsilon / \alpha\right)ML_i}{\left(\gamma - 1\right)\tau K_j L_i + \left(\varepsilon / \alpha\right)NL_j}, \tag{A10}$$

according to (31), with respect to γ. This gives after straightforward calculations

$$\frac{d\left(x_{ii} / x_{jj}\right)}{d\gamma} = K_j \frac{x_{ii}}{x_{jj}}\left[\frac{\tau L_j + \left(\varepsilon / \alpha\right)\left(1 - \alpha / \varepsilon + \tau\right)L_i}{\left[\left(K_i - K_j\right) + \tau\left(\gamma K_j - K_i\right)\right]L_j + \left(\varepsilon / \alpha\right)ML_i} - \frac{\tau L_i + \left(\varepsilon / \alpha\right)\left(1 - \alpha / \varepsilon + \tau\right)L_j}{\left(\gamma - 1\right)\tau K_j L_i + \left(\varepsilon / \alpha\right)NL_j}\right] \tag{A11}$$

and thus,

$$\frac{d\left(x_{ii} / x_{jj}\right)}{d\gamma} = -K_j \frac{x_{ii}}{x_{jj}}\frac{\left(K_i - K_j\right)\left(1-\tau\right)}{\phi \times \psi}\left\{L_i^2\frac{\varepsilon}{\alpha}\tau + L_j^2\frac{\varepsilon}{\alpha}\left[\left(1+\tau\right)\left(1 - \frac{\alpha}{\varepsilon}\right) + \tau\right] + L_i L_j\left[\left(\frac{\varepsilon}{\alpha}\right)^2\left(1 - \frac{\alpha}{\varepsilon} + \tau\right)\left(2 - \frac{\alpha}{\varepsilon}\right) + \tau\right]\right\} < 0, \tag{A12}$$

where

$$\phi := \left[\tau L_j + \frac{\varepsilon}{\alpha}\left(1 - \frac{\alpha}{\varepsilon} + \tau\right)L_i\right]\left(\gamma K_j - K_i\right) + \left[L_j + \frac{\varepsilon}{\alpha}\left(2 - \frac{\alpha}{\varepsilon}\right)L_i\right]\left(K_i - K_j\right), \tag{A13}$$

$$\psi := \left[\tau L_i + \frac{\varepsilon}{\alpha}\left(1 - \frac{\alpha}{\varepsilon} + \tau\right)L_j\right]\left(\gamma K_j - K_i\right) + \left[\tau L_i + \frac{\varepsilon}{\alpha}\left(2 - \frac{\alpha}{\varepsilon}\right)\tau L_j\right]\left(K_i - K_j\right) \tag{A14}$$

have been considered.

Finally, using $d\left(x_{ii}/x_{jj}\right)/d\gamma<0$ in the first derivatives of (32) and (33) with respect to γ gives Result 3. ■

Proof of Result 5

Differentiating GLI_{SIII}^{C}, according to (35), with respect to γ gives

$$\begin{aligned}\frac{dGLI_{\text{SIII}}^{C}}{d\gamma}=&\frac{\varepsilon\tau\left[d\left(n_i/v_i\right)/d\gamma-GLI_{SIII}^{C}\times\left(x_{ii}/x_{jj}\right)\times d\left(n_j/v_i\right)/d\gamma\right]}{\varepsilon\tau\left(n_j/v_i\right)x_{ii}/x_{jj}+\left(\varepsilon-1\right)\tau x_{ii}/x_{jj}-1}\\&-GLI_{\text{SIII}}^{C}\frac{\varepsilon\tau\left(n_j/v_i\right)+\left(\varepsilon-1\right)\tau}{\varepsilon\tau\left(n_j/v_i\right)x_{ii}/x_{jj}+\left(\varepsilon-1\right)\tau x_{ii}/x_{jj}-1}\frac{d\left(x_{ii}/x_{jj}\right)}{d\gamma}.\end{aligned}\tag{A15}$$

Using $n_i/v_i=\left(\gamma K_j-K_i\right)/\left(K_i-K_j\right)$ and $n_j/v_i=n_i/v_i+1$ (which implies $d\left(n_i/v_i\right)/d\gamma=d\left(n_j/v_i\right)/d\gamma>0$), according to (28), and noting that $GLI_{\text{SIII}}^{C}<1$ and $x_{ii}/x_{jj}<1$ must hold in a positive wage equilibrium, it is straightforward that the first term on the right hand side of (A15) must be positive. Together with $d\left(x_{ii}/x_{jj}\right)/d\gamma<0$, according to (A12), this implies $dGLI_{\text{SIII}}^{C}/d\gamma>0$.

Next, we calculate the first derivative of SHI_{SIII} with respect to v_i and obtain

$$\frac{\partial SHI_{\text{SIII}}}{\partial v_i}=\frac{\left(SHI_{\text{SIII}}-1\right)}{v_i}\frac{\varepsilon\tau n_j x_{ii}/x_{jj}}{\varepsilon\tau n_j x_{ii}/x_{jj}+\left(\varepsilon-1\right)\tau v_i x_{ii}/x_{jj}-v_i}>0,\tag{A16}$$

according to (36). Moreover, differentiating SHI_{SIII} with respect to x_{ii}/x_{jj} gives

$$\frac{\partial SHI_{SIII}}{\partial\left(x_{ii}/x_{jj}\right)}=-\frac{\left(SHI_{SIII}-1\right)\varepsilon\tau\left(n_j+v_i\right)}{\left[1+\tau x_{ii}/x_{jj}\right]\left\{\varepsilon\tau n_j x_{ii}/x_{jj}+\left(\varepsilon-1\right)\tau v_i x_{ii}/x_{jj}-v_i\right\}}.\tag{A17}$$

In view of (26), (28), (A16) and (A17), we therefore obtain

$$\begin{aligned}\frac{dSHI_{\text{SIII}}}{d\gamma}&=\frac{\partial SHI_{\text{SIII}}}{\partial v_i}\frac{dv_i}{d\gamma}+\frac{\partial SHI_{\text{SIII}}}{\partial\left(x_{ii}/x_{jj}\right)}\frac{d\left(x_{ii}/x_{jj}\right)}{d\gamma}\\&=-\frac{\left(SHI_{\text{SIII}}-1\right)\varepsilon\tau}{\left[1+\tau x_{ii}/x_{jj}\right]Q}\left\{\frac{x_{ii}}{x_{jj}}\left(1+\tau\frac{x_{ii}}{x_{jj}}\right)\frac{K_j}{\gamma-1}+\left(n_j+v_i\right)\frac{d\left(x_{ii}/x_{jj}\right)}{d\gamma}\right\}\end{aligned}\tag{A18}$$

with $Q := \varepsilon\tau n_j x_{ii}/x_{jj} + (\varepsilon-1)\tau v_i x_{ii}/x_{jj} - v_i$. We substitute $d(x_{ii}/x_{jj})/d\gamma$, according to (A12), in (A18). Thereby, we consider ϕ and ψ, according to (A13) and (A14), and note that $x_{ii}/x_{jj} = \phi/\psi$, according to (31). Moreover, we use $n_j = K_j(\gamma-1)/(\gamma-1)$ and $v_i = (K_i - K_j)/(\gamma-1)$, according to (26) and (28). After tedious calculations we then obtain

$$\frac{dSHI_{\text{SIII}}}{d\gamma} = -\frac{(SHI_{\text{SIII}}-1)\varepsilon\tau}{\phi\psi\left[1+\tau x_{ii}/x_{jj}\right]Q}\frac{K_j}{\gamma-1}\frac{x_{ii}}{x_{jj}}D, \tag{A19}$$

with $D := \left\{T_1(K_i-K_j)^2 + T_2(K_i-K_j)(\gamma K_j - K_i) + T_3(\gamma K_j - K_i)^2 + \tau\phi^2\right\}$ and

$$\frac{(SHI_{\text{SIII}}-1)\varepsilon\tau}{\phi\psi\left[1+\tau x_{ii}/x_{jj}\right]Q}\frac{K_j}{\gamma-1}\frac{x_{ii}}{x_{jj}} > 0,$$

according to (36). Thereby,

$$\begin{aligned} T_1 := {} & \tau\frac{\varepsilon}{\alpha}\left(2\tau-\frac{\alpha}{\varepsilon}\right)L_i^2 - \frac{\varepsilon}{\alpha}\left[2\left(1-\tau^2\right)\left(1-\frac{\alpha}{\varepsilon}\right)-\tau\left(2\tau-\frac{\alpha}{\varepsilon}\right)\right]L_j^2 \\ & -\left\{\left(\frac{\varepsilon}{\alpha}\right)^2\left(2-\frac{\alpha}{\varepsilon}\right)\left[2(1-\tau)\left(1-\frac{\alpha}{\varepsilon}\right)-\tau\left(2\tau-\frac{\alpha}{\varepsilon}\right)\right]-\tau(2\tau-1)\right\}L_iL_j \end{aligned}, \tag{A20}$$

$$T_2 := 2\tau\left\{\frac{\varepsilon}{\alpha}\left(1-\frac{\alpha}{\varepsilon}+\tau\right)L_i^2 + \frac{\varepsilon}{\alpha}\tau\left(2-\frac{\alpha}{\varepsilon}\right)L_j^2 + \left[\left(\frac{\varepsilon}{\alpha}\right)^2\left(2-\frac{\alpha}{\varepsilon}\right)\left(1-\frac{\alpha}{\varepsilon}+\tau\right)+\tau\right]L_iL_j\right\}, \tag{A21}$$

$$T_3 := \left[\tau L_j + \frac{\varepsilon}{\alpha}\left(1-\frac{\alpha}{\varepsilon}+\tau\right)L_i\right]\left[\tau L_i + \frac{\varepsilon}{\alpha}\left(1-\frac{\alpha}{\varepsilon}+\tau\right)L_j\right] \tag{A22}$$

have been considered. ($T_2 > 0$ and $T_3 > 0$ hold for all $\tau > 0$.) Function *D* has the following properties: $D|_{\tau=1} > 0$,

$$D|_{\tau=0} = -\frac{\varepsilon}{\alpha}\left(1-\frac{\alpha}{\varepsilon}\right)\left\{2L_j^2(K_i-K_j)^2 + \frac{\varepsilon}{\alpha}L_iL_j\left[2\left(2-\frac{\alpha}{\varepsilon}\right)(K_i-K_j)^2 - \left(1-\frac{\alpha}{\varepsilon}\right)(\gamma K_j-K_i)^2\right]\right\}$$

and $dD/d\tau > D|_{\tau=0}$ for all $\tau > 0$. Thus, we can distinguish two cases: If $D|_{\tau=0} > 0$, then $D > 0$ and thus, $dSHI_{\text{SIII}}/d\gamma < 0$, according to (A19), for all $\tau \in (0,1)$. However, if $D|_{\tau=0} < 0$, then there exists a unique $\underline{\tau} \in (0,1)$ such that $dSHI_{\text{SIII}}/d\gamma < 0$ ($D > 0$) $\forall \tau > \underline{\tau}$ and $(dSHI_{SIII}/d\gamma)|_{\tau=\underline{\tau}} = 0$ ($D = 0$). This follows immediately from (A19)-(A22) and the fact that $D\big|_{\tau=1} > 0$. Hence, in the case of $D\big|_{\tau=0} < 0$, $\tau > \underline{\tau}$ guarantees $dSHI_{\text{SIII}} / d\gamma < 0$.[39] This completes the proof of Result 5. ■

B. Simulation appendix

Table A.2 provides details on the assumptions about the chosen parameter values in the numerical simulation exercise.

> Table A.2 <

Our choice of the parameter related to the technical rate of substitution points to a complementary relationship between factors of production, which is in line with recent evidence (see Sharma, 2002). The choice of the elasticity of substitution parameter between varieties is well in line with the findings in Feenstra (1994), and that one of the factor shares broadly reflects the findings in Mankiw et al. (1992). The assumption that iceberg trade costs vary around 15% is well in line with the stylized facts (see Baier and Bergstrand, 2001).

C. Descriptive statistics on different measurement biases

To provide a complete picture of the size of both intra-industry trade shares and the various biases discussed, we report descriptive statistics of bilateral Grubel-Lloyd indices according to each concept, computed on the basis of three different levels of aggregation (5-digit, 4-digit, and 3-digit) as published by the OECD using the Standard International Trade Classification.

> Table A.2-A.3 <

The figures in Table A.2 illustrate that the average uncorrected intra-industry trade share amounts to about 14-21% for the average bilateral OECD relationship between 1990 and 2000, depending of which level of aggregation is used. Trade imbalance corrected figures, of course, tend to be considerably higher. In almost all cases, irrespective of which concept or aggregation level is chosen, the standard error in the share is about as large as the mean. As

[39] A more detailed proof is relegated to a supplement, made available in the GEP working paper version of the paper.

the last column in the table indicates, the major part of this variation is due to the cross-section rather than the time dimension. However, all concepts where missing values at the disaggregated level are interpreted as reflecting zero trade, tend to exhibit much more time variation than the others. For this reason, cross-sectional rather than time series (or panel data) analysis seems better suited for intra-industry trade share measurement, since measurement errors in the time dimension are likely to cancel out. Table A.3 displays the correlation matrix between all measurement concepts. Obviously, the various corrections are strong enough showing up in correlation coefficients as small as 0.14 between *GLI* and the (not preferred) $AGLI_3^C$, but also that one between the preferred GLI_3^C (GLI_1^C) and the usually used *GLI* amounts only to 0.36 (0.59). Although Tables A.2 and A.3 provide first insights into the relative size of the various biases discussed, Table A.4 focuses more directly on this issue and summarizes average bias figures.

> Table A.4 <

The reported biases are computed in the following way. To quantify the **trade imbalance bias**, we calculate $GLI - GLI_1^C$, $GLI_2^C - GLI_3^C$ and $GLI_4^C - GLI_5^C$, as indicated in Subsection 3.1. It is obvious that this contributes more than any other bias. At the average aggregation level, the uncorrected intra-industry trade share is downward biased by about 14 percentage points, which is about 51%-81% of the mean. Of course, the trade imbalance bias is related to the level of MNE activity. For instance, when regressing the (logit-transformed) absolute value of the bias on the log absolute difference between two partner countries world outward FDI stocks, we obtain a coefficient of about 0.10, which is significant at 10%.

For the **transport cost level bias**, we subtract the respective export based intra-industry trade share indices from their *uncorrected* counterparts. (In Table A.4, we treat GLI_3^C as the preferred measure of the intra-industry trade share.) In particular, $GLI - GLI_2^C$, $AGL - AGL_2^C$ and $GLI_1^C - GLI_3^C$, $AGLI_1^C - AGLI_3^C$ have been computed. The bias is always displayed with the intra-industry trade share concept it is affecting. From Table A.4, we see that the transport cost bias is relatively small, amounting to 0.6 percentage points on average. Transport costs tend to upward bias (by about 7%-10%) the uncorrected intra-industry trade share, whereas their impact on trade-imbalance corrected intra-industry trade is – on average – almost negligible (between -0.9% and -1.7% of the corresponding intra-industry trade share).

However, we would expect that this bias to be much larger in a sample of non-OECD economies.

As mentioned above, the **transport cost difference bias** drives a wedge between the import-based concepts and the export-based concepts of intra-industry trade share measurement. Accordingly, only $GLI_4^C - GLI_2^C$, $AGLI_4^C - AGLI_2^C$ and $GLI_5^C - GLI_3^C$, $AGLI_5^C - AGLI_3^C$ are computed to estimate this. In our sample, this bias is even smaller than the transport cost level bias, indicating that the asymmetry between two trading partners' transport costs is relatively small. (Again, a much larger bias of this type might be present if we considered non-OECD countries.)

The **missing value interpretation bias** has to be interpreted with care, since no information is available on whether *GLI* is closer to the true value or *AGLI* (and similarly for the corrected figures). As our working hypothesis, we take the extreme position and assume that all missing values indicate zero trade flows. In this case, the missing value interpretation bias amounts to 8.2 percentage points (about 47%) of intra-industry trade shares on average.

For the **Finger bias**, we subtract for each concept the intra-industry trade share data of the respective higher level of aggregation from its next lower counterpart (i.e., SITC 4-digit based shares minus 5-digit based ones and 3-digit based ones minus their 4-digit based counterparts). Then, we average the resulting differences over the two aggregation levels and all country pairs and years. According to our results, using 4-digit instead of 5-digit data exerts an upward bias of about 3.4 percentage points on the average intra-industry trade share (i.e., about 25%). Of course, using 3-digit data instead causes an upward bias by about twice as much.

In a final step, we aggregate the aforementioned biases, taking GLI_3^C at the SITC 5-digit level as the preferred measure of the intra-industry trade share. Of course, the discussed biases do not simply add up, since they exhibit a non-zero covariance. The overall biases are reported in the last two columns of Table A.4, Importantly, the last column of Table A.4 is independent of the Finger bias, since all bias figures are with respect to 5-digit based intra-industry trade share measures. We see that the traditionally used Grubel-Lloyd index is downward biased by about 10 percentage points, which is about 43% of the corrected value.

D. Data appendix

Data sources and definition

We use bilateral export and import flow data at the Standard International Trade Classification 5-digit, 4-digit and 3-digit level as published by the OECD (International Trade by Commodity Statistics, 1990-2000). Bilateral transport costs are based on trade-weighted averages of c.i.f./f.o.b. figures from this source.

Real GDP figures are from the World Bank's World Development Indicators and measured in constant US dollars of 1995.

Capital stock data had to be computed by the perpetual inventory method as discussed in Leamer (1984, pp.232-234). Since no data on depreciation rates are available for our countries, the same value as in Leamer (i.e., 13.3%) is assumed. Data on human capital measure the average years of schooling of participants in the active labour force (see Baier, Dwyer and Tamura, 2002, for more details). Endowment data were kindly provided by Scott Baier.

Investment cost data are based on score variables published in the World Economic Forum's Global Competitiveness Report. Amiti and Wakelin (2003) provide a detailed description. The data were kindly provided by Keith Maskus.

Table A.5 provides the correlation matrix and summary statistics for the explanatory variables.

> Table A.5 <

Country sample

The country sample the regression results are based on consists of bilateral trade flows between the following 31 countries:

Australia, Austria, Belgium, Canada, China, Czech Republic, Denmark, Finland, France, Germany, Greece, Hong Kong, Hungary, Iceland, Ireland, Italy, Japan, Republic of Korea, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Spain, Sweden, Switzerland, Turkey, United Kingdom, USA.

References

Amiti, Mary and Katharine Wakelin, Investment Liberalization and International Trade, *Journal of International Economics* 61 (1), 101-126.

Anderson. James E. and Eric van Wincoop (2003), Gravity with Gravitas: A Solution to the Border Puzzle, *American Economic Review* 93 (1), 170-192.

Anderson, James E. and Eric van Wincoop (2004), Trade Costs, preliminary draft, in preparation for the *Journal of Economic Literature*.

Aquino, Antonio (1978), Inta-Industry Trade and Inter-Industry Specialization as Concurrent Sources in International Trade in Manufactures, *Weltwirtschaftliches Archiv* 114 (2), 275-296.

Baier, Scott L. and Jeffrey H. Bergstrand (2001), The Growth of World Trade: Tariffs, Transport Costs, and Income Similarity, *Journal of International Economics* 53 (1), 1-27.

Baier, Scott L., Gerald P. Dwyer and Robert Tamura (2002), How Important Are Capital and Total Factor Productivity for Economic Growth?, unpublished manuscript.

Baltagi, Badi H. (2001), *Econometric Analysis of Panel Data*, second edition, Wiley, Chichester.

Belsley, D. A., Edwin Kuh, and Roy E. Welsch (1980), *Regression Diagnostics*, Wiley, New York.

Bergstrand, Jeffrey H. (1983), Measurement and Determinants of Intra-Industry International Trade, in P. K. Matthew Tharakan (ed.), Intra-industry trade: *Empirical and Methodological Aspects*, Elsevier Science & Technology Books.

Bergstrand, Jeffrey H. (1990), The Heckscher-Ohlin-Samuelson Model, the Linder Hypothesis, and the Determinants of Bilateral Intra-industry trade, *Economic Journal*, 100 (403), 1216-1229.

Bhagwati, Jagdish and Donald R. Davis (1999), Intraindustry Trade: Issues and Theory, in: Melvin, James R., James C. Moore, and Raymond Riezman, eds., Trade Theory and Econometrics: Essays in Honor of John S. Chipman, Studies in the Modern World Economy, vol. 15., Routledge, London and New York, 1-16.

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

Brander, James A. and Paul R. Krugman (1983), A 'Reciprocal dumping' Model of International Trade, *Journal of International Economics* 15 (3-4), 313-321.

Carr, David, Keith Maskus and James R. Markusen (2001), Estimating the Knowledge-Capital Model of the Multinational Enterprise, *American Economic Review* 91 (2001), 693-708.

Davis, Donald R. (1995), Intra-industry Trade: A Heckscher-Ohlin-Ricardo Approach, *Journal of International Economics* 39 (3-4), 201-226.

Davis, Donald R. (1997), Critical Evidence on Comparative Advantage: North-North Trade in a Multilateral World, *Journal of Political Economy* 105 (4), 1051-1060.

Dixit, Avinash and Joseph Stiglitz (1977), Monopolistic Competition and Optimum Product Diversity; *American Economic Review* 67 (3), 297-308.

Durkin, John T. and Markus Krygier (1998), Comparative Advantage and the Pattern of Trade within Industries, *Review of International Economics*, 6 (2), 292-306.

Durkin, John T. and Markus Krygier (2000), Differences in GDP Per Captia and the Share of Intraindustry Trade: The Role of Vertically Differentiated Trade, *Review of International Economics*, 8 (4), 760-774.

Feenstra Robert C. (1994), New Product Varieties and the Measurement of International Prices, *American Economic Review* 84 (1), 157-177.

Finger, J. Michael (1975), Trade Overlap and Intra-industry trade, *Economic Inquiry* 13 (4), 581-589.

Greenaway, David and Chris Milner (1981), Trade Imbalance Effects in the Measurement of Intra-Industry Trade, *Weltwirtschaftliches Archiv* 117 (4), 756-762.

Greenaway, David and Chris Milner (1983), On the Measurement of Intra-industry trade, *Economic Journal* 93 (372), 900-908.

Greenaway, David and Chris Milner (1984), A Cross-Section Analysis of Intra-Industry Trade in the UK, *European Economic Review*, 25, 319-344.

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

Greenaway, David, Peter Lloyd and Chris Milner (2001), New Concepts and Measures of the Globalisation of Production, *Economics Letters* 73 (1), 57-63.

Grubel, Herbert G. and Peter J. Lloyd, (1971), The Empirical Measurement of Intra- Industry Trade, *Economic Record* 47 (120), 494-517.

Helpman, Elhanan (1984), A Simple Theory of International Trade with Multinational Corporations, *Journal of Political Economy* 32 (3), 451-71.

Helpman, Elhanan (1987), Imperfect Competition and International Trade: Evidence from Fourteen Industrial Countries, *Journal of the Japanese and the International Economies* 1 (1), 62-81.

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

Hummels, David and James Levinsohn (1995), Monopolistic Competition and International Trade: Reconsidering the Evidence, *Quarterly Journal of Economics* 110 (3), 799-836.

Krugman, Paul R. (1979), Increasing Returns, Monopolistic Competition, and International Trade, *Journal of International Economics* 9 (4), 469-479.

Krugman, Paul R. (1980), Scale Economies, Product differentiation, and the Pattern of Trade, *American Economic Review* 70 (5), 950-959.

Krugman, Paul R. (1981), Intra-industry Specialisation and the Gains from Trade, *Journal of Political Economy*, 89, 959-973.

Lancaster, Kelvin (1980), Intra-Industry Trade under Perfect Monopolistic Competition, *Journal of International Economics* 10 (2), 151-175.

Leamer, Edward E. (1984), *Sources of International Comparative Advantage*, MIT Press, Cambridge, MA.

Leamer, Edward E. (1994), Testing Trade Theory, in David Greenaway and L. Alan Winters *Surveys in International Trade*, Blackwell, Oxford.

Loertscher, Rudolf and Frank Wolter (1980), Determinants of Intra-Industry Trade: Among Countries and across Industries, *Weltwirtschaftliches Archiv* 116 (2), 280-293.

Mankiw, Gregory N., David Romer, and David N. Weil (1992), A Contribution to the Empirics of Economic Growth, *Quarterly Journal of Economics* 107 (2), 407-437.

Markusen, James R. (1995), The Boundaries of Multinational Enterprises and the Theory of International Trade, *Journal of Economic Perspectives* 9 (2), 169-189.

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

Markusen, James R. and Keith Maskus (2001), A Unified Approach to Intra-industry trade and Direct Foreign Investment, NBER Working Paper No. 8335.

Markusen, James R. and Anthony Venabales (1998), Multinational Firms and the New Trade Theory, *Journal of International Economics* 46 (2), 183-204.

Markusen, James R. and Anthony Venables (2000), The Theory of Endowment, Intra-Industry, and Multinational Trade, *Journal of International Economics* 52 (2), 209-234.

Torstensson, Johan (1996), How Robust are Models of Intra-Industry Trade?, *Oxford Bulletin of Economics and Statistics*, 58, 507-524.

Venables, Anthony J., Patricia G. Rice and Martin Stewart (2003), The Geography of Intra-Industry Trade: Empirics, Topics in Economic Analysis & Policy 3 (11), http://www.bepress.com/bejeap/topics/vol3/iss1/art11.

Table 1 – Alternative Definitions of the Grubel-Lloyd Index

Label	Definition	Interpretation
GLI	$2 \cdot \sum_k \min(EX_{ijk}, M_{ijk}) \big/ \left(EX_{ij} + M_{ij}\right)$	Where $EX_{ij} = \sum_k EX_{ijk}$ are aggregate f.o.b exports of and $M_{ij} = \sum_k M_{ijk}$ are the corresponding c.i.f imports of country *i*. Missing values at the disaggregated level are treated as 0.
$AGLI$		As *GLI*, but missing values at the disaggregated level are skipped and not interpreted as 0.
GLI_1^C	$\sum_k \min(EX_{ijk}, M_{ijk}) \big/ \min(EX_{ij}, M_{ij})$	As *GLI*, but taking into account that part of the trade volume serves to balance imbalanced trade in invisibles as induced by the presence of MNEs.
$AGLI_1^C$		As GLI_1^C, but missing values at the disaggregated level are skipped and not interpreted as 0.
GLI_2^C	$2 \cdot \sum_k \min(EX_{ijk}, EX_{jik}) \big/ (EX_{ij} + EX_{ji})$	As *GLI*, but only considering trade flows at f.o.b. With positive transport costs, $M_{ijk} \neq EX_{jik}$ and $M_{ij} \neq EX_{ji}$.
$AGLI_2^C$		As GLI_2^C, but missing values at the disaggregated level are skipped and not interpreted as 0.
GLI_3^C	$\sum_k \min(EX_{ijk}, EX_{jik}) \big/ \min(EX_{ij}, EX_{ji})$	As GLI_2^C, but taking into account that part of the trade volume serves to balance imbalanced trade in invisibles as induced by the presence of MNEs.
$AGLI_3^C$		As GLI_3^C, but missing values at the disaggregated level are skipped and not interpreted as 0.
GLI_4^C	$2 \cdot \sum_k \min(M_{ijk}, M_{jik}) \big/ (M_{ij} + M_{ji})$	As GLI_2^C, but based on trade flows at c.i.f. instead of f.o.b. GLI_4^C differs from GLI_2^C if $t_{ij} \neq t_{ji}$.
$AGLI_4^C$		As GLI_4^C, but missing values at the disaggregated level are skipped and not interpreted as 0.
GLI_5^C	$\sum_k \min(M_{ijk}, M_{jik}) \big/ \min(M_{ij}, M_{ji})$	As GLI_4^C, but taking into account that part of the trade volume serves to balance imbalanced trade in invisibles as induced by the presence of MNEs. GLI_5^C differs from GLI_3^C if $t_{ij} \neq t_{ji}$.
$AGLI_5^C$		As GLI_5^C, but missing values at the disaggregated level are skipped and not interpreted as 0.

Table 2 - The Determinants of Intra-Industry Trade Shares (Between Regression Results; 1990-2000 Data; Left-Hand-Side Variable is Logit-Transformed)
(All Left-Hand-Side Variables are Based on 5-digit SITC Figures)

		GLI	AGLI	GLI^C_1	$AGLI^C_1$	GLI^C_3	$AGLI^C_3$
Maximum GDP: max{ln(GDP_i),ln(GDP_j)}	δ_1	0.066 *	-0.021	0.035	-0.047 *	-0.181 ***	-0.242 ***
		1.89	0.87	1.00	1.79	2.78	5.20
Minimum GDP: min{ln(GDP_i),ln(GDP_j)}	δ_2	0.498 ***	0.136 ***	0.470 ***	0.144 ***	0.483 ***	0.147 ***
		13.18	5.16	12.35	5.02	6.74	2.88
Maximum Capital-Labor Ratio: max{ln(K_i/L_i),ln(Kj/L_j)}	δ_3	0.255 ***	-0.019	0.304 ***	-0.001	0.898 ***	0.366 ***
		2.64	0.29	3.12	0.01	4.89	2.78
Minimum Capital-Labor Ratio: min{ln(K_i/L_i),ln(Kj/L_j)}	δ_4	0.283 ***	0.071	0.346 ***	0.165 ***	0.199	-0.038
		4.35	1.57	5.23	3.35	1.60	0.43
Maximum Endowment with Skilled Labor: max{S_i/L_i,S_j/L_j}	δ_5	-1.047 *	-1.397 ***	-2.070 ***	-2.141 ***	-3.918 ***	-2.824 ***
		1.72	3.28	3.34	4.60	3.43	3.37
Minimum Endowment with Skilled Labor: min{S_i/L_i,S_j/L_j}	δ_6	0.112	1.210 ***	-0.278	0.590 *	0.428	1.781 ***
		0.24	3.72	0.59	1.67	0.48	2.83
Maximum Investment Costs: max{ln($INVC_i$),ln($INVC_j$)}	δ_7	-1.039 ***	-0.510 ***	-0.878 ***	-0.396 ***	-1.155 ***	-0.892 ***
		5.74	4.03	4.79	2.90	3.36	3.65
Minimum Investment Costs: min{ln($INVC_i$),ln($INVC_j$)}	δ_8	0.043	0.171	0.166	0.345 **	0.474	0.619 **
		0.24	1.33	0.90	2.49	1.38	2.52
Maximum Transport Costs: max{ln(TC_{ij}),ln(TC_{ji})}	δ_9	-0.735 ***	-0.338 ***	-0.544 ***	-0.147 ***	-0.654 ***	-0.162 *
		11.11	7.31	8.15	3.05	5.46	1.93
Minimum Transport Costs: min{ln(TC_{ij}),ln(TC_{ji})}	δ_{10}	0.187 **	0.178 ***	0.142 *	0.064	0.246 *	0.179 *
		2.47	3.38	1.87	1.14	1.82	1.87
Constant	δ_0	-20.650 ***	-4.840 ***	-17.838 ***	-2.369	-14.549 ***	1.583
		10.64	3.56	9.10	1.64	3.94	0.62
Observations		866	866	866	866	422	422
R^2		0.52	0.24	0.43	0.14	0.35	0.21
Share of R^2, accounted for by variables 5-10 in %		48.96	68.70	41.39	50.64	47.88	49.17
F-tests (p-values):							
$\delta_1=-\delta_2$		0.000 ***	0.000 ***	0.000 ***	0.003 ***	0.000 ***	0.096 *
$\delta_3=-\delta_4$		0.000 ***	0.487	0.000 ***	0.044 **	0.000 ***	0.025 **
$\delta_5=-\delta_6$		0.153	0.683	0.000 ***	0.002 ***	0.006 ***	0.255
$\delta_7=-\delta_8$		0.000 ***	0.004 ***	0.000 ***	0.687	0.036 **	0.240
$\delta_9=-\delta_{10}$		0.000 ***	0.000 ***	0.000 ***	0.007 ***	0.000 ***	0.756

Absolute t-statistics below coefficients. *** significant at 1%; ** significant at 5%; * significant at 10%.

Table 3 - Sensitivity Analysis of Preferred Models (Traditional, Trade-Imbalance-Adjusted Indices; Assuming that Trade Costs Generate Exporter Income)

		Estimates are Based on Between Models and Exclude Extreme Outliers					
		Based on 5-digit data		Based on 4-digit data		Based on 3-digit data	
		GLI^C_1	$AGLI^C_1$	GLI^C_1	$AGLI^C_1$	GLI^C_1	$AGLI^C_1$
Maximum GDP: max{ln(GDP_i),ln(GDP_j)}	δ_1	0.064 ***	-0.004	-0.075 ***	-0.119 ***	-0.174 ***	-0.109 ***
		3.70	0.25	2.76	8.08	6.75	6.78
Minimum GDP: min{ln(GDP_i),ln(GDP_j)}	δ_2	0.441 ***	0.160 ***	0.476 ***	0.233 ***	0.436 ***	0.277 ***
		23.91	9.24	16.95	14.64	16.40	16.34
Maximum Capital-Labor Ratio: max{ln(K_i/L_i),ln(Kj/L_j)}	δ_3	0.218 ***	-0.064	0.678 ***	0.207 ***	0.674 ***	0.197 ***
		4.57	1.43	9.18	5.09	9.74	4.51
Minimum Capital-Labor Ratio: min{ln(K_i/L_i),ln(Kj/L_j)}	δ_4	0.357 ***	0.170 ***	0.138 ***	0.058 **	0.072	0.095 ***
		11.48	5.69	3.03	2.23	1.68	3.41
Maximum Endowment with Skilled Labor: max{S_i/L_i,S_j/L_j}	δ_5	-1.337 ***	-1.900 ***	-1.777 ***	-1.925 ***	-1.790 ***	-1.894 ***
		4.38	6.60	4.00	7.70	4.31	7.09
Minimum Endowment with Skilled Labor: min{S_i/L_i,S_j/L_j}	δ_6	-0.274	0.740 ***	0.403	1.329 ***	0.736 **	0.870 ***
		1.20	3.43	1.17	6.96	2.28	4.24
Maximum Investment Costs: max{ln($INVC_i$),ln($INVC_j$)}	δ_7	-0.809 ***	-0.600 ***	-1.023 ***	-0.771 ***	-0.893 ***	-1.020 ***
		9.22	7.28	7.78	10.37	7.20	12.68
Minimum Investment Costs: min{ln($INVC_i$),ln($INVC_j$)}	δ_8	0.152 *	0.432 ***	0.175	0.297 ***	0.191	0.397 ***
		1.69	5.15	1.34	3.96	1.54	4.89
Maximum Transport Costs: max{ln(TC_{ij}),ln(TC_{ji})}	δ_9	-0.479 ***	-0.149 ***	-0.649 ***	-0.365 ***	-0.642 ***	-0.437 ***
		15.23	5.25	14.45	14.08	15.20	15.60
Minimum Transport Costs: min{ln(TC_{ij}),ln(TC_{ji})}	δ_{10}	0.116 ***	0.053 *	0.091 *	0.100 ***	0.156 ***	0.165 ***
		3.22	1.65	1.81	3.45	3.29	5.19
Constant	δ_0	-20.102 ***	-3.923 ***	-19.250 ***	-7.010 ***	-14.958 ***	-6.399 ***
		22.91	4.67	14.80	9.40	12.09	8.16
Observations		859	849	834	840	834	840
R^2		0.82	0.47	0.76	0.73	0.77	0.73
Share of R^2, accounted for by variables 5-10 in %		46.10	54.34	53.52	61.09	54.04	59.52
F-tests (p-values):							
δ_1=-δ_2		0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***
δ_3=-δ_4		0.000 ***	0.030 **	0.000 ***	0.000 ***	0.000 ***	0.000 ***
δ_5=-δ_6		0.000 ***	0.000 ***	0.004 ***	0.026 **	0.019 **	0.000 ***
δ_7=-δ_8		0.000 ***	0.029 **	0.000 ***	0.000 ***	0.000 ***	0.000 ***
δ_9=-δ_{10}		0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***
		Estimates Include Fixed Exporter and Importer Effects and Exclude Extreme Outliers					
		Based on 5-digit data		Based on 4-digit data		Based on 3-digit data	
		GLI^C_1	$AGLI^C_1$	GLI^C_1	$AGLI^C_1$	GLI^C_1	$AGLI^C_1$
Maximum GDP: max{GDP_i,GDP_j}	δ_1	-0.254 *	-0.656 **	-3.470 ***	-3.250 ***	-4.254 ***	-3.346 ***
		1.75	1.98	8.19	11.44	10.99	12.07
Minimum GDP: min{GDP_i,GDP_j}	δ_2	-0.204 *	-0.598 *	-3.665 ***	-3.096 ***	-4.332 ***	-3.055 ***
		1.81	1.82	8.67	10.91	11.30	11.07
Maximum Capital-Labor Ratio: max{K_i/L_i,Kj/L_j}	δ_3	1.786 ***	-0.384	3.636 ***	3.969 ***	1.525 **	3.702 ***
		2.81	0.68	4.67	7.42	2.08	7.42
Minimum Capital-Labor Ratio: min{K_i/L_i,Kj/L_j}	δ_4	1.796 ***	-0.027	3.245 ***	4.047 ***	1.339 *	3.697 ***
		2.83	0.05	4.20	7.61	1.84	7.45
Maximum Endowment with Skilled Labor: max{S_i/L_i,S_j/L_j}	δ_5	-0.559	-1.356 **	3.216 ***	-3.820 ***	1.870 ***	-3.735 ***
		1.00	2.35	4.54	8.15	2.75	8.48
Minimum Endowment with Skilled Labor: min{S_i/L_i,S_j/L_j}	δ_6	0.636	0.173	6.063 ***	-2.189 **	4.423 ***	-2.246 ***
		1.03	0.28	7.70	4.18	5.80	4.62
Maximum Investment Costs: max{ln($INVC_i$),ln($INVC_j$)}	δ_7	-0.482 ***	-0.163 *	-0.639 ***	-0.077 *	-0.771 ***	-0.195 **
		4.13	1.76	4.43	1.81	5.62	2.09
Minimum Investment Costs: min{ln($INVC_i$),ln($INVC_j$)}	δ_8	0.585 ***	0.565 ***	0.520 ***	0.533 ***	0.526 ***	0.770 ***
		4.71	4.91	3.44	5.21	3.63	7.80
Maximum Transport Costs: max{ln(TC_{ij}),ln(TC_{ji})}	δ_9	-0.219 ***	0.004	-0.177 ***	-0.127	-0.281 ***	-0.148 ***
		7.37	0.13	4.78	5.19	8.04	6.28
Minimum Transport Costs: min{ln(TC_{ij}),ln(TC_{ji})}	δ_{10}	0.036	0.018	-0.157 ***	-0.034	-0.092 ***	-0.032
		1.16	0.66	4.25	1.37	2.65	1.34
Constant	δ_0	-32.393 **	37.157 ***	91.088 ***	85.284 ***	180.820 ***	94.208 ***
		2.09	2.69	4.41	6.27	9.55	7.15
Observations		857	848	838	834	838	836
R^2		0.90	0.70	0.91	0.85	0.92	0.89
Share of R^2, accounted for by variables 5-10 in %		51.02	50.40	51.42	50.63	51.34	50.79
F-tests (p-values):							
δ_1=-δ_2		0.496	0.058 *	0.000 ***	0.000 ***	0.000 ***	0.000 ***
δ_3=-δ_4		0.005 ***	0.714	0.000 ***	0.000 ***	0.050 **	0.000 ***
δ_5=-δ_6		0.943	0.288	0.000 ***	0.000 ***	0.000 ***	0.000 ***
δ_7=-δ_8		0.556	0.015 **	0.574	0.001 ***	0.226	0.000 ***
δ_9=-δ_{10}		0.000 ***	0.268	0.000 ***	0.000 ***	0.000 ***	0.000 ***
Fixed exporter effects		0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***
Fixed importer effects		0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***

Absolute t-statistics below coefficients. *** significant at 1%; ** significant at 5%; * significant at 10%.

Table 4 - Sensitivity Analysis of Preferred Models (Export-Based, Trade-Imbalance-Adjusted Indices; Assuming that Trade Costs Do Not Generate Income)

		Estimates are Based on Between Models and Exclude Extreme Outliers					
		Based on 5-digit data		Based on 4-digit data		Based on 3-digit data	
		GLI^C_3	$AGLI^C_3$	GLI^C_3	$AGLI^C_3$	GLI^C_3	$AGLI^C_3$
Maximum GDP: $\max\{GDP_i,GDP_j\}$	δ_1	-0.106 ***	-0.147 ***	-0.173 ***	-0.191 ***	-0.253 ***	-0.194 ***
		3.09	5.49	4.25	7.27	6.69	6.96
Minimum GDP: $\min\{GDP_i,GDP_j\}$	δ_2	0.503 ***	0.144 ***	0.528 ***	0.242 ***	0.447 ***	0.340 ***
		14.01	5.15	12.82	8.85	11.75	11.93
Maximum Capital-Labor Ratio: $\max\{K_i/L_i,Kj/L_j\}$	δ_3	0.760 ***	0.269 ***	0.690 ***	0.292 ***	0.668 ***	0.205 ***
		8.08	3.70	6.47	4.13	6.80	2.76
Minimum Capital-Labor Ratio: $\min\{K_i/L_i,Kj/L_j\}$	δ_4	0.231 ***	-0.030	0.256 ***	0.154 ***	0.201 ***	0.222 ***
		3.84	0.60	3.69	3.28	3.19	4.47
Maximum Endowment with Skilled Labor: $\max\{S_i/L_i,S_j/L_j\}$	δ_5	-2.968 ***	-2.713 ***	-2.300 ***	-3.010 ***	-2.589 ***	-2.759 ***
		4.85	5.74	3.44	6.64	4.25	5.84
Minimum Endowment with Skilled Labor: $\min\{S_i/L_i,S_j/L_j\}$	δ_6	0.104	1.717 ***	-0.106	1.014 ***	0.280	0.459
		0.24	4.88	0.21	2.97	0.60	1.28
Maximum Investment Costs: $\max\{\ln(INVC_i),\ln(INVC_j)\}$	δ_7	-0.951 ***	-0.704 ***	-0.861 ***	-0.593 ***	-0.912 ***	-0.924 ***
		5.62	5.31	4.42	4.43	5.12	6.62
Minimum Investment Costs: $\min\{\ln(INVC_i),\ln(INVC_j)\}$	δ_8	0.465 ***	0.478 ***	0.600 ***	0.637 ***	0.634 ***	0.792 ***
		2.78	3.53	3.11	4.79	3.58	5.68
Maximum Transport Costs: $\max\{\ln(TC_{ij}),\ln(TC_{ji})\}$	δ_9	-0.603 ***	-0.192 ***	-0.480 ***	-0.173 ***	-0.396 ***	-0.232 ***
		10.16	4.21	7.01	3.86	6.55	4.78
Minimum Transport Costs: $\min\{\ln(TC_{ij}),\ln(TC_{ji})\}$	δ_{10}	0.266 ***	0.183 ***	0.069	0.024	0.013	0.104 *
		4.03	3.54	0.90	0.47	0.20	1.91
Constant	δ_0	-15.706 ***	-0.147	-16.643 ***	-5.373 ***	-11.550 ***	-3.805 ***
		9.29	0.11	8.83	4.08	6.59	2.86
Observations		413	413	415	415	415	415
R^2		0.74	0.47	0.69	0.57	0.75	0.28
Share of R^2, accounted for by variables 5-10 in %		51.45	52.06	50.66	49.55	50.56	50.29
F-tests (p-values):							
$\delta_1=-\delta_2$		0.000 ***	0.938	0.000 ***	0.113	0.000 ***	0.000 ***
$\delta_3=-\delta_4$		0.000 ***	0.003 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***
$\delta_5=-\delta_6$		0.000 ***	0.055	0.001 ***	0.000 ***	0.000 ***	0.000 ***
$\delta_7=-\delta_8$		0.003 ***	0.080 *	0.160	0.733	0.100 *	0.319
$\delta_9=-\delta_{10}$		0.000 ***	0.768	0.000 ***	0.000 ***	0.000 ***	0.000 ***

		Estimates Include Fixed Exporter and Importer Effects and Exclude Extreme Outliers					
		Based on 5-digit data		Based on 4-digit data		Based on 3-digit data	
		GLI^C_3	$AGLI^C_3$	GLI^C_3	$AGLI^C_3$	GLI^C_3	$AGLI^C_3$
Maximum GDP: $\max\{\ln(GDP_i),\ln(GDP_j)\}$	δ_1	-2.181 ***	-1.579 **	-4.195 ***	-3.442 ***	-4.455 ***	-2.987 ***
		3.68	2.43	5.59	6.07	5.38	5.28
Minimum GDP: $\min\{\ln(GDP_i),\ln(GDP_j)\}$	δ_2	-2.265 ***	-1.387 **	-4.253 ***	-3.276 ***	-4.411 ***	-2.685 ***
		3.85	2.14	5.70	5.80	5.34	4.77
Maximum Capital-Labor Ratio: $\max\{\ln(K_i/L_i),\ln(Kj/L_j)\}$	δ_3	3.386 ***	1.512	5.855 ***	5.413 ***	5.062 ***	3.248 ***
		3.08	1.40	4.57	5.33	3.53	3.23
Minimum Capital-Labor Ratio: $\min\{\ln(K_i/L_i),\ln(Kj/L_j)\}$	δ_4	2.698 **	1.509	5.371 ***	5.400 ***	4.896 ***	3.290 ***
		2.47	1.40	4.20	5.33	3.42	3.27
Maximum Endowment with Skilled Labor: $\max\{S_i/L_i,S_j/L_j\}$	δ_5	-0.996	-0.690	2.700 **	-0.441	1.500	-1.399
		0.93	0.49	2.43	0.47	1.30	1.58
Minimum Endowment with Skilled Labor: $\min\{S_i/L_i,S_j/L_j\}$	δ_6	3.759 ***	2.406 *	5.466 ***	0.566	3.277 **	0.123
		3.29	1.73	4.51	0.56	2.56	0.13
Maximum Investment Costs: $\max\{\ln(INVC_i),\ln(INVC_j)\}$	δ_7	-1.080 ***	-0.481 **	-0.806 ***	-0.549 ***	-0.822 ***	-0.866 ***
		4.78	2.23	3.42	2.86	3.25	4.40
Minimum Investment Costs: $\min\{\ln(INVC_i),\ln(INVC_j)\}$	δ_8	0.906 ***	0.774 ***	1.192 ***	1.017 ***	0.992 ***	0.999 ***
		3.94	3.40	4.66	4.89	3.58	4.61
Maximum Transport Costs: $\max\{\ln(TC_{ij}),\ln(TC_{ji})\}$	δ_9	-0.147 ***	-0.043	-0.100	0.004	-0.051	-0.010
		2.64	0.81	1.67	0.08	0.80	0.20
Minimum Transport Costs: $\min\{\ln(TC_{ij}),\ln(TC_{ji})\}$	δ_{10}	-0.073	0.074	-0.081	0.001	-0.156 **	-0.060
		1.32	1.39	1.37	0.02	2.54	1.17
Constant	δ_0	55.571 **	47.727	103.984 ***	73.757 ***	132.360 ***	90.169 ***
		2.02	1.66	3.05	3.00	3.59	3.49
Observations		413	411	413	413	411	413
R^2		0.89	0.62	0.89	0.78	0.88	0.81
Share of R^2, accounted for by variables 5-10 in %		51.45	52.06	50.66	49.55	50.56	50.29
F-tests (p-values):							
$\delta_1=-\delta_2$		0.000 ***	0.023 **	0.000 ***	0.000 ***	0.000 ***	0.000 ***
$\delta_3=-\delta_4$		0.006 ***	0.161	0.000 ***	0.000 ***	0.001 ***	0.001 ***
$\delta_5=-\delta_6$		0.174	0.519	0.000 ***	0.945	0.032 **	0.450
$\delta_7=-\delta_8$		0.601	0.372	0.276	0.105	0.658	0.657
$\delta_9=-\delta_{10}$		0.000 ***	0.371	0.000 ***	0.883	0.000 ***	0.032 **
Fixed exporter effects		0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***
Fixed importer effects		0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***	0.000 ***

Absolute t-statistics below coefficients. *** significant at 1%; ** significant at 5%; * significant at 10%.

Table 5 - The Role of Labor and Capital Endowments for the Impact of Investment Costs (5-Digit Data Based; All Regressions Include Country Effects and Exclude Outliers)

		The Role of Labor Endowments				The Role of Physical Capital Endowments			
		GLI^C_1	$AGLI^C_1$	GLI^C_3	$AGLI^C_3$	GLI^C_1	$AGLI^C_1$	GLI^C_3	$AGLI^C_3$
Maximum GDP: max{ln(GDP_i),ln(GDP_j)}	δ_1	-0.249 ***	-0.691 **	-2.042 ***	-1.343 **	-0.312 ***	-0.679 **	-2.151 ***	-1.191 *
		10.76	2.13	3.43	2.18	10.95	2.09	3.56	1.92
Minimum GDP: min{ln(GDP_i),ln(GDP_j)}	δ_2	-0.185 ***	-0.633 *	-2.125 ***	-1.129 *	-0.250 ***	-0.623 *	-2.224 ***	-0.976
		10.57	1.96	3.59	1.83	10.77	1.92	3.70	1.58
Maximum Capital-Labor Ratio: max{ln(K_i/L_i),ln(Kj/L_j)}	δ_3	2.024 ***	-0.326	3.185 ***	1.683	2.158 ***	-0.335	3.300 ***	1.552
		3.27	0.59	2.89	1.58	3.49	0.60	2.97	1.45
Minimum Capital-Labor Ratio: min{ln(K_i/L_i),ln(Kj/L_j)}	δ_4	2.047 ***	0.003	2.540 **	1.619	2.181 ***	-0.004	2.611 **	1.479
		3.31	0.01	2.31	1.52	3.53	0.01	2.36	1.38
Maximum Endowment with Skilled Labor: max{S_i/L_i,S_j/L_j}	δ_5	-0.575	-1.466 **	-1.281	-2.029	-0.558	-1.471 **	-1.290	-2.130
		1.04	2.56	1.18	1.46	1.01	2.56	1.18	1.52
Minimum Endowment with Skilled Labor: min{S_i/L_i,S_j/L_j}	δ_6	0.556	0.184	3.213 ***	1.784	0.543	0.175	3.091 ***	1.708
		0.91	0.30	2.79	1.30	0.89	0.29	2.66	1.24
Maximum Investment Costs: max{ln($INVC_i$),ln($INVC_j$)}	δ_7	-0.682 ***	-0.317 **	-1.401 ***	-1.613 ***	-0.683 ***	-0.311 **	-1.432 ***	-1.620 ***
		4.80	2.34	4.52	5.55	4.80	2.29	4.57	5.54
Minimum Investment Costs: min{ln($INVC_i$),ln($INVC_j$)}	δ_8	0.930 ***	0.700 ***	1.292 ***	1.690 ***	0.937 ***	0.687 ***	1.351 ***	1.688 ***
		6.11	4.91	4.36	5.91	6.17	4.82	4.52	5.88
Maximum Transport Costs: max{ln(TC_{ij}),ln(TC_{ji})}	δ_9	-0.22 ***	0.01	-0.17 ***	-0.10 *	-0.22 ***	0.01	-0.18 ***	-0.10 *
		7.36	0.27	3.07	1.95	7.46	0.35	3.13	1.84
Minimum Transport Costs: min{ln(TC_{ij}),ln(TC_{ji})}	δ_{10}	0.025	0.010	-0.038	0.122 **	0.026	0.010	-0.026	0.118 **
		0.83	0.37	0.68	2.27	0.87	0.35	0.47	2.19
Interaction: Δ\{ln($INVC_i$),ln($INVC_j$)\}·ln(L_i) if $INVC_i$>$INVC_j$, else Δ\{ln($INVC_i$),ln($INVC_j$)\}·ln(L_j)	δ_{11}	0.033 ***	0.019 **	0.037 *	0.103 ***	-	-	-	-
		3.24	1.99	1.78	5.23	-	-	-	-
Interaction: Δ\{ln($INVC_i$),ln($INVC_j$)\}·ln(K_i) if $INVC_i$>$INVC_j$, else Δ\{ln($INVC_i$),ln($INVC_j$)\}·ln(K_j)	δ_{12}	-	-	-	-	0.02 ***	0.01 *	0.03 **	0.07 ***
		-	-	-	-	3.25	1.91	2.03	5.28
Constant	δ_0	-38.86 **	38.31 ***	52.93 *	35.78	-38.38 **	37.94 ***	56.89 **	30.37
		2.55	2.80	1.90	1.29	2.52	2.77	2.03	1.09
Observations		857	848	413	411	857	848	413	411
R^2		0.90	0.71	0.89	0.65	0.90	0.70	0.88	0.64

Δ\{ln($INVC_i$),ln($INVC_j$)\} is defined as max\{ln($INVC_i$),ln($INVC_j$)\} - min\{ln($INVC_i$),ln($INVC_j$)\}. Absolute t-statistics below coefficients. *** significant at 1%; ** significant at 5%; * significant at 10%.

Table 6 - Explaining GLI^C/GLI

	GLI^C_1/GLI_1	GLI^C_3/GLI_3
$\Delta\{\ln(INVC_i),\ln(INVC_j)\}$	-2.191	-89.663 **
	1.52	2.12
Interaction: $\Delta\{\ln(INVC_i),\ln(INVC_j)\}\cdot\ln(L_i)$ if $INVC_i > INVC_j$, else $\Delta\{\ln(INVC_i),\ln(INVC_j)\}\cdot\ln(L_j)$	1.551 *	7.719 **
	1.80	2.06
Constant	-2.452	-9.184
	0.20	0.18
Observations	857	413
R^2	0.11	0.03
F-tests (p-values):		
Joint significance of all other explanatory variables (see Footnote)	0.199	0.433
Fixed exporter effects	0.000 ***	0.000 ***
Fixed importer effects	0.011 **	0.000 ***

$\Delta\{\ln(INVC_i),\ln(INVC_j)\}$ is defined as $\max\{\ln(INVC_i),\ln(INVC_j)\} - \min\{\ln(INVC_i),\ln(INVC_j)\}$. Coefficients of $\max\{\ln(GDP_i),\ln(GDP_j)\}$, $\min\{\ln(GDP_i),\ln(GDP_j)\}$, $\max\{\ln(K_i/L_i),\ln(K_j/L_j)\}$, $\min\{\ln(K_i/L_i),\ln(K_j/L_j)\}$, $\max\{\ln(S_i/L_i),\ln(S_j/L_j)\}$, $\min\{\ln(S_i/L_i),\ln(S_j/L_j)\}$, $\max\{\ln(TC_{ij}),\ln(TC_{ji})\}$, $\min\{\ln(TC_{ij}),\ln(TC_{ji})\}$ not reported due to their insignificance (see the F-statistics).

Absolute t-statistics below coefficients. *** significant at 1%; ** significant at 5%; * significant at 10%.

Table A.1 - Simulation Set-up

	Vertical Model	Horizontal Model	Knowledge-Capital Model		
			Leontief	Cobb-Douglas	CES
	V	H	KK1	KK2	KK3
Endowments of i:					
Share of K	[0.62,0.77]	[0.45,0.55]	[0.40,0.60]	[0.40,0.60]	[0.40,0.60]
Share of S	[0.48,0.52]	[0.45,0.55]	[0.40,0.60]	[0.40,0.60]	[0.40,0.60]
Share of L	[0.15,0.25]	[0.45,0.55]	[0.40,0.60]	[0.40,0.60]	[0.40,0.60]
Input coefficients:					
a_{KX}	0	0	0.3	see footnote	
a_{SX}	0	0	0.2	see footnote	
a_{LX}	0.5	0.5	0.5	see footnote	
Investment costs:					
Additional foreign investment costs of i	[1.10,1.30]	[0.10,0.30]	[0.10,0.30]	[0.10,0.30]	[0.10,0.30]
Additional foreign investment costs of j	1.2	0.2	0.2	0.2	0.2
Trade costs of differentiated goods:					
Iceberg parameter TC_{ij}	[1.05,1.25]	[1.05,1.25]	[1.05,1.25]	[1.05,1.25]	[1.05,1.25]
Iceberg parameter of TC_{ji}	1.15	1.15	1.15	1.15	1.15

In all experiments, we set ε=6 (see Feenstra, 1994) and α=0.8 (according to UN Comtrade data for 1990-2000). The stepwidth between minimum and maximum additional foreign investment costs of country i is always 0.05. We assume the following values for world endowments: K=60; S=40; L=100. In Model KK2 S=80, and in models KK1 and KK3 S=200 and K=300 are asuumed to ensure that exporters and horizontal multinationals co-exist in the center of the factor cube. The factor box is always split into 21 segments of equal size in any of the two dimensions, so that there are 21×21×21 equilibria to be solved for each level of investment costs. For the Cobb-Douglas case, we assume the production technology $z_i=K_i^{0.3}S_i^{0.2}L_i^{0.5}$ with i=1,2. For the more general case of a Constant Elasticity of Substitution Technology, we assume $z_i=[0.3K_i^{\rho}+0.2S_i^{\rho}+0.5L_i^{\rho}]^{1/\rho}$ with ρ=-10 and i=1,2.

Table A.2 - Summary Statistics for Different Concepts of the Grubel-Lloyd Index

	Observations	Mean	Std. Dev.	Minimum	Maximum	Time Invar.[a)]
5-digit SITC data						
GLI (usual definition; missings=0)	8429	0.14	0.13	0.00	0.64	0.92
AGLI (missings≠ 0)	8429	0.21	0.14	0.00	1.00	0.69
GLI^C_1 (GLI balance-adjusted)	8429	0.23	0.23	0.00	1.00	0.91
$AGLI^C_1$ (AGLI balance-adjusted)	8429	0.35	0.24	0.00	1.00	0.72
GLI^C_2 (GLI export-based)	7259	0.13	0.14	0.00	1.00	0.88
$AGLI^C_2$ (AGLI export-based)	7259	0.20	0.14	0.00	1.00	0.69
GLI^C_3 (GLI^C_2 balance adjusted)	7259	0.24	0.27	0.00	1.00	0.94
$AGLI^C_3$ ($AGLI^C_2$ balance adjusted)	7259	0.36	0.28	0.00	1.00	0.77
GLI^C_4 (GLI import-based)	7429	0.14	0.15	0.00	1.00	0.95
$AGLI^C_4$ (GLI import-based)	7429	0.21	0.15	0.00	1.00	0.81
GLI^C_5 (GLI^C_4 balance adjusted)	7429	0.25	0.27	0.00	1.00	0.96
$AGLI^C_5$ ($AGLI^C_4$ balance adjusted)	7429	0.36	0.27	0.00	1.00	0.83
4-digit SITC data						
GLI (usual definition; missings=0)	8495	0.17	0.15	0.00	0.71	0.91
AGLI (missings≠ 0)	8495	0.23	0.15	0.00	1.00	0.75
GLI^C_1 (GLI balance-adjusted)	8495	0.27	0.24	0.00	1.00	0.90
$AGLI^C_1$ (AGLI balance-adjusted)	8495	0.38	0.24	0.00	1.00	0.73
GLI^C_2 (GLI export-based)	7345	0.15	0.15	0.00	1.00	0.87
$AGLI^C_2$ (AGLI export-based)	7345	0.21	0.15	0.00	1.00	0.73
GLI^C_3 (GLI^C_2 balance adjusted)	6878	0.13	0.19	0.00	1.00	0.92
$AGLI^C_3$ ($AGLI^C_2$ balance adjusted)	6878	0.24	0.17	0.00	1.00	0.78
GLI^C_4 (GLI import-based)	7345	0.28	0.28	0.00	1.00	0.95
$AGLI^C_4$ (GLI import-based)	7345	0.38	0.28	0.00	1.00	0.86
GLI^C_5 (GLI^C_4 balance adjusted)	6878	0.25	0.31	0.00	1.00	0.96
$AGLI^C_5$ ($AGLI^C_4$ balance adjusted)	6878	0.41	0.27	0.00	1.00	0.84
3-digit SITC data						
GLI (usual definition; missings=0)	8491	0.21	0.17	0.00	0.78	0.91
AGLI (missings≠ 0)	8491	0.26	0.17	0.00	1.00	0.80
GLI^C_1 (GLI balance-adjusted)	8491	0.33	0.25	0.00	1.00	0.88
$AGLI^C_1$ (AGLI balance-adjusted)	8491	0.41	0.25	0.00	1.00	0.75
GLI^C_2 (GLI export-based)	7337	0.19	0.17	0.00	1.00	0.86
$AGLI^C_2$ (AGLI export-based)	7337	0.24	0.17	0.00	1.00	0.77
GLI^C_3 (GLI^C_2 balance adjusted)	7472	0.21	0.18	0.00	1.00	0.89
$AGLI^C_3$ ($AGLI^C_2$ balance adjusted)	7472	0.25	0.17	0.00	1.00	0.78
GLI^C_4 (GLI import-based)	7337	0.34	0.29	0.00	1.00	0.92
$AGLI^C_4$ (GLI import-based)	7337	0.41	0.28	0.00	1.00	0.86
GLI^C_5 (GLI^C_4 balance adjusted)	7472	0.35	0.27	0.00	1.00	0.92
$AGLI^C_5$ ($AGLI^C_4$ balance adjusted)	7472	0.42	0.27	0.00	1.00	0.82

a) This is the share of time-invariant information in the data.

Table A.3 - Correlation Matrix for Different Concepts of the Grubel-Lloyd Index (OECD Annual Statistics of Foreign Trade Data, 1990-2000)

	Indices based on 5-digit SITC data											
	GL	AGL	CGL_1	$CAGL_1$	CGL_2	$CAGL_2$	CGL_3	$CAGL_3$	CGL_4	$CAGL_4$	CGL_5	$CAGL_5$
5-digit SITC data												
GLI (usual definition; missings=0)	1.00											
AGLI (missings≠ 0)	0.71	1.00										
GLI^C_1 (GLI balance-adjusted)	0.59	0.39	1.00									
$AGLI^C_1$ (AGLI balance-adjusted)	0.28	0.56	0.70	1.00								
GLI^C_2 (GLI export-based)	0.87	0.60	0.53	0.23	1.00							
$AGLI^C_2$ (AGLI export-based)	0.63	0.61	0.35	0.30	0.77	1.00						
GLI^C_3 (GLI^C_2 balance adjusted)	0.36	0.18	0.74	0.51	0.46	0.32	1.00					
$AGLI^C_3$ ($AGLI^C_2$ balance adjusted)	0.14	0.17	0.54	0.61	0.26	0.45	0.79	1.00				
GLI^C_4 (GLI import-based)	0.87	0.56	0.50	0.18	0.88	0.63	0.36	0.14	1.00			
$AGLI^C_4$ (GLI import-based)	0.69	0.62	0.36	0.26	0.73	0.65	0.24	0.17	0.85	1.00		
GLI^C_5 (GLI^C_4 balance adjusted)	0.37	0.18	0.73	0.47	0.40	0.23	0.92	0.69	0.50	0.40	1.00	
$AGLI^C_5$ ($AGLI^C_4$ balance adjusted)	0.16	0.18	0.56	0.59	0.21	0.21	0.75	0.73	0.30	0.46	0.81	1.00
4-digit SITC data												
GLI (usual definition; missings=0)	0.96	0.64	0.56	0.23	0.84	0.58	0.36	0.12	0.84	0.63	0.37	0.14
AGLI (missings≠ 0)	0.80	0.82	0.44	0.38	0.68	0.56	0.24	0.11	0.65	0.58	0.23	0.11
GLI^C_1 (GLI balance-adjusted)	0.59	0.36	0.95	0.64	0.54	0.33	0.73	0.51	0.51	0.34	0.71	0.51
$AGLI^C_1$ (AGLI balance-adjusted)	0.36	0.46	0.75	0.85	0.31	0.29	0.54	0.55	0.26	0.25	0.50	0.52
GLI^C_2 (GLI export-based)	0.83	0.53	0.50	0.19	0.96	0.72	0.46	0.23	0.85	0.67	0.39	0.19
$AGLI^C_2$ (AGLI export-based)	0.70	0.56	0.40	0.24	0.85	0.85	0.37	0.32	0.71	0.66	0.30	0.22
GLI^C_3 (GLI^C_2 balance adjusted)	0.55	0.34	0.27	0.06	0.61	0.44	0.20	0.06	0.71	0.61	0.30	0.17
$AGLI^C_3$ ($AGLI^C_2$ balance adjusted)	0.69	0.53	0.32	0.14	0.74	0.62	0.21	0.09	0.84	0.87	0.34	0.30
GLI^C_4 (GLI import-based)	0.38	0.17	0.72	0.46	0.49	0.31	0.97	0.75	0.38	0.25	0.89	0.70
$AGLI^C_4$ (GLI import-based)	0.21	0.16	0.58	0.55	0.33	0.38	0.83	0.89	0.21	0.21	0.73	0.71
GLI^C_5 (GLI^C_4 balance adjusted)	0.22	0.07	0.57	0.36	0.28	0.16	0.78	0.60	0.35	0.26	0.83	0.67
$AGLI^C_5$ ($AGLI^C_4$ balance adjusted)	0.24	0.17	0.58	0.51	0.30	0.25	0.76	0.69	0.37	0.43	0.82	0.90
3-digit SITC data												
GLI (usual definition; missings=0)	0.93	0.63	0.54	0.20	0.83	0.57	0.35	0.10	0.82	0.61	0.36	0.12
AGLI (missings≠ 0)	0.85	0.76	0.47	0.31	0.73	0.57	0.27	0.10	0.72	0.62	0.27	0.12
GLI^C_1 (GLI balance-adjusted)	0.61	0.37	0.91	0.61	0.56	0.34	0.70	0.48	0.53	0.35	0.68	0.48
$AGLI^C_1$ (AGLI balance-adjusted)	0.43	0.45	0.77	0.79	0.39	0.32	0.56	0.53	0.34	0.31	0.53	0.53
GLI^C_2 (GLI export-based)	0.81	0.51	0.48	0.16	0.93	0.69	0.44	0.21	0.83	0.65	0.38	0.17
$AGLI^C_2$ (AGLI export-based)	0.74	0.55	0.42	0.20	0.87	0.77	0.38	0.26	0.77	0.70	0.33	0.22
GLI^C_3 (GLI^C_2 balance adjusted)	0.83	0.49	0.46	0.11	0.84	0.59	0.35	0.11	0.95	0.78	0.47	0.24
$AGLI^C_3$ ($AGLI^C_2$ balance adjusted)	0.77	0.54	0.40	0.14	0.80	0.63	0.29	0.12	0.90	0.84	0.42	0.28
GLI^C_4 (GLI import-based)	0.41	0.17	0.71	0.44	0.52	0.33	0.93	0.72	0.41	0.27	0.85	0.67
$AGLI^C_4$ (GLI import-based)	0.28	0.18	0.60	0.52	0.39	0.37	0.83	0.83	0.29	0.27	0.74	0.70
GLI^C_5 (GLI^C_4 balance adjusted)	0.43	0.18	0.71	0.41	0.47	0.27	0.87	0.63	0.54	0.40	0.94	0.74
$AGLI^C_5$ ($AGLI^C_4$ balance adjusted)	0.31	0.20	0.62	0.51	0.36	0.27	0.79	0.68	0.43	0.41	0.85	0.85

	Indices based on 4-digit SITC data											
	GL	AGL	CGL_1	$CAGL_1$	CGL_2	$CAGL_2$	CGL_3	$CAGL_3$	CGL_4	$CAGL_4$	CGL_4	$CAGL_4$
4-digit SITC data												
GLI (usual definition; missings=0)	1.00											
AGLI (missings≠ 0)	0.81	1.00										
GLI^C_1 (GLI balance-adjusted)	0.64	0.49	1.00									
$AGLI^C_1$ (AGLI balance-adjusted)	0.37	0.57	0.76	1.00								
GLI^C_2 (GLI export-based)	0.86	0.67	0.57	0.31	1.00							
$AGLI^C_2$ (AGLI export-based)	0.71	0.65	0.44	0.35	0.86	1.00						
GLI^C_3 (GLI^C_2 balance adjusted)	0.56	0.43	0.31	0.14	0.61	0.52	1.00					
$AGLI^C_3$ ($AGLI^C_2$ balance adjusted)	0.67	0.58	0.34	0.20	0.71	0.67	0.84	1.00				
GLI^C_4 (GLI import-based)	0.41	0.27	0.75	0.54	0.53	0.43	0.23	0.23	1.00			
$AGLI^C_4$ (GLI import-based)	0.23	0.20	0.59	0.61	0.35	0.48	0.12	0.15	0.83	1.00		
GLI^C_5 (GLI^C_4 balance adjusted)	0.25	0.14	0.57	0.40	0.30	0.23	0.58	0.41	0.77	0.64	1.00	
$AGLI^C_5$ ($AGLI^C_4$ balance adjusted)	0.24	0.18	0.57	0.52	0.30	0.29	0.41	0.49	0.74	0.71	0.82	1.00
3-digit SITC data												
GLI (usual definition; missings=0)	0.98	0.80	0.62	0.35	0.85	0.70	0.56	0.67	0.41	0.21	0.25	0.23
AGLI (missings≠ 0)	0.87	0.93	0.53	0.48	0.73	0.68	0.48	0.64	0.30	0.21	0.17	0.21
GLI^C_1 (GLI balance-adjusted)	0.66	0.51	0.98	0.74	0.59	0.46	0.34	0.37	0.74	0.57	0.55	0.55
$AGLI^C_1$ (AGLI balance-adjusted)	0.45	0.56	0.79	0.93	0.39	0.39	0.21	0.27	0.57	0.61	0.42	0.54
GLI^C_2 (GLI export-based)	0.85	0.65	0.55	0.28	0.98	0.84	0.61	0.71	0.52	0.34	0.30	0.29
$AGLI^C_2$ (AGLI export-based)	0.76	0.66	0.47	0.32	0.89	0.93	0.56	0.73	0.45	0.42	0.26	0.31
GLI^C_3 (GLI^C_2 balance adjusted)	0.85	0.63	0.52	0.22	0.86	0.72	0.71	0.83	0.40	0.22	0.36	0.35
$AGLI^C_3$ ($AGLI^C_2$ balance adjusted)	0.78	0.66	0.44	0.26	0.80	0.73	0.68	0.86	0.33	0.23	0.31	0.38
GLI^C_4 (GLI import-based)	0.45	0.28	0.75	0.52	0.57	0.46	0.27	0.28	0.98	0.82	0.74	0.73
$AGLI^C_4$ (GLI import-based)	0.30	0.24	0.62	0.60	0.42	0.48	0.18	0.23	0.85	0.94	0.64	0.72
GLI^C_5 (GLI^C_4 balance adjusted)	0.47	0.28	0.74	0.49	0.50	0.38	0.36	0.40	0.89	0.71	0.80	0.80
$AGLI^C_5$ ($AGLI^C_4$ balance adjusted)	0.33	0.25	0.63	0.57	0.37	0.36	0.28	0.38	0.78	0.74	0.72	0.86

	Indices based on 3-digit SITC data											
	GL	AGL	CGL_1	$CAGL_1$	CGL_2	$CAGL_2$	CGL_3	$CAGL_3$	CGL_4	$CAGL_4$	CGL_4	$CAGL_4$
3-digit SITC data												
GLI (usual definition; missings=0)	1.00											
AGLI (missings≠ 0)	0.87	1.00										
GLI^C_1 (GLI balance-adjusted)	0.67	0.56	1.00									
$AGLI^C_1$ (AGLI balance-adjusted)	0.44	0.59	0.80	1.00								
GLI^C_2 (GLI export-based)	0.86	0.73	0.60	0.37	1.00							
$AGLI^C_2$ (AGLI export-based)	0.77	0.74	0.50	0.42	0.89	1.00						
GLI^C_3 (GLI^C_2 balance adjusted)	0.86	0.73	0.56	0.33	0.87	0.80	1.00					
$AGLI^C_3$ ($AGLI^C_2$ balance adjusted)	0.79	0.75	0.48	0.37	0.80	0.80	0.94	1.00				
GLI^C_4 (GLI import-based)	0.46	0.34	0.76	0.58	0.58	0.50	0.46	0.38	1.00			
$AGLI^C_4$ (GLI import-based)	0.30	0.29	0.62	0.65	0.42	0.53	0.32	0.32	0.86	1.00		
GLI^C_5 (GLI^C_4 balance adjusted)	0.48	0.35	0.75	0.56	0.50	0.44	0.59	0.52	0.89	0.75	1.00	
$AGLI^C_5$ ($AGLI^C_4$ balance adjusted)	0.33	0.31	0.63	0.63	0.37	0.39	0.45	0.51	0.77	0.77	0.88	1.00

Table A.4 - Quantifying the Various Sources of Bias in Intra-Industry Trade Shares
(Bias Figures are Averaged over Time, Bilateral Relationships and the Three Aggregation Levels)

	Transport cost level bias	Transport cost difference bias	Trade im-balance bias	Missing value interpret. bias	Finger (1975) bias per digit	Total if 5-digit data used[a)]
GLI (usual definition; missings=0)	0.017		-0.104		0.039	-0.104
AGLI (missings≠ 0)	0.016		-0.146	0.059	0.023	-0.031
GLI^C_1 (GLI balance-adjusted)	-0.005				0.055	-0.013
$AGLI^C_1$ (AGLI balance-adjusted)	-0.003			0.101	0.030	0.110
GLI^C_2 (GLI export-based)			-0.128		0.035	-0.117
$AGLI^C_2$ (AGLI export-based)			-0.167	0.057	0.020	-0.044
GLI^C_3 (GLI^C_2 balance adjusted)					0.047	0.000
$AGLI^C_3$ ($AGLI^C_2$ balance adjusted)				0.096	0.023	0.120
GLI^C_4 (GLI import-based)		0.001	-0.127		0.037	-0.108
$AGLI^C_4$ (GLI import-based)		0.011	-0.168	0.070	0.022	-0.045
GLI^C_5 (GLI^C_4 balance adjusted)		-0.004			0.049	0.007
$AGLI^C_5$ ($AGLI^C_4$ balance adjusted)		0.007		0.110	0.026	0.112
Weighted average (# of obs. weight)	0.006	0.004	-0.139	0.082	0.034	-0.009

All figures assume in accordance with our theoretical model that CLI^C_3 is the correct index. - a) Excluding the Finger bias.

Table A.5 - Correlation Matrix and Descriptive Statistics of Explanatory Variables (Variables in Logs)

	Max GDP	Min GDP	Max K/L	Min K/L	Max S/L	Min S/L	Max INVC	Min INVC	Max TC	Min TC
Maximum GDP: max{ln(GDP_i),ln(GDP_j)}	1.00									
Minimum GDP: min{ln(GDP_i),ln(GDP_j)}	0.45	1.00								
Maximum Capital-Labor Ratio: max{ln(K_i/L_i),ln(Kj/L_j)}	0.28	0.16	1.00							
Minimum Capital-Labor Ratio: min{ln(K_i/L_i),ln(Kj/L_j)}	0.13	0.11	0.45	1.00						
Maximum Endowment with Skilled Labor: max{ln(S_i/L_i),ln(S_j/L_j)}	0.41	0.27	0.44	0.36	1.00					
Minimum Endowment with Skilled Labor: min{ln(S_i/L_i),ln(S_j/L_j)}	0.17	-0.05	0.37	0.70	0.50	1.00				
Maximum Investment Costs: max{ln($INVC_i$),ln($INVC_j$)}	0.05	0.20	-0.23	-0.02	-0.26	-0.20	1.00			
Minimum Investment Costs: min{ln($INVC_i$),ln($INVC_j$)}	-0.02	-0.06	-0.39	-0.29	-0.37	-0.23	0.56	1.00		
Maximum Transport Costs: max{ln(TC_{ij}),ln(TC_{ji})}	-0.09	-0.15	-0.12	-0.24	-0.15	-0.26	-0.12	-0.01	1.00	
Minimum Transport Costs: min{ln(TC_{ij}),ln(TC_{ji})}	-0.14	-0.17	-0.12	-0.17	-0.13	-0.17	-0.07	-0.01	0.71	1.00
	Descriptive Statistics									
Mean	27.04	25.46	11.13	10.23	1.90	1.77	3.58	3.34	-1.46	-2.09
Standard Deviation	1.28	1.21	0.57	1.04	0.08	0.14	0.29	0.30	1.21	1.29