



SCHOOL OF ECONOMICS
AND MANAGEMENT
Lund University

Master thesis

July 2006



Intra-Industry Trade

-
An analysis of measurements

Rikard Lindqvist

Senior Advisor
Yves Bourdet
Joakim Gullstrand

Abstract

In the last decades intra-industry trade (IIT) got a lot of attention in economic research. In 1975 Grubel & Lloyd presented a way of measuring IIT, this method is used even today. Later researchers found out that IIT can be disentangled into vertical and horizontal IIT. To calculate vertical and horizontal IIT unit-values are used.

In this paper I want to test different ways of measuring, IIT, vertical and horizontal IIT. Different aggregational levels of data are used, as well as an alternative to the unit-value approach, which I call the quality-defined index. The data set is classified according to the Harmonised system (HS) and covers the passenger car sector. The passenger car sector is used in this paper based on the fact that the data available is very good. To test the different measures correlation coefficients and coefficients of variation are used.

The aggregational level affects the result of the calculations, especially when measuring vertical IIT. Still the measures are significantly positively correlated. When defining quality as done in the quality-defined measure, instead of using unit-values, there are large differences between this measure and the measure based on disaggregated data.

Key-words: International trade, Intra-industry trade, vertical and horizontal intra-industry trade, measurements and correlation coefficient.

ABSTRACT	2
1. INTRODUCTION	4
1.1. PURPOSE.....	4
1.2. DISPOSITION.....	5
2. THEORY	6
2.1. BACKGROUND THEORY	6
2.2. HECKSCHER-OHLIN – AND LATER VERSIONS OF THEIR MODEL	7
2.3. PASSENGER CAR INDUSTRY	8
2.4. INTRA-INDUSTRY TRADE – HORIZONTAL AND VERTICAL	9
3. METHOD	11
3.1. GRUBEL-LLOYD INDEX	12
3.1.1. CRITICISM TOWARDS THE GRUBEL-LLOYD INDEX	13
3.2. CORRELATION BETWEEN MEASURES	13
3.3. INDEXES.....	14
4. CALCULATIONS	15
4.1. UNIT-VALUE CALCULATIONS.....	15
4.2. CLASSIFYING IIT	16
4.2.1. CRITICISM TOWARDS THE UNIT-VALUE APPROACH	17
4.3. INTER INDUSTRY TRADE, IIT, VERTICAL IIT AND HORIZONTAL IIT	17
4.4. MEASURES OF IIT	18
4.4.1 AGGREGATED MEASURE OF IIT	18
4.4.2. DISAGGREGATED MEASURE OF IIT	19
4.5. MEASURES OF VERTICAL IIT.....	19
4.5.1. AGGREGATED MEASURE OF VERTICAL IIT	19
4.5.2. DISAGGREGATED MEASURE OF VERTICAL IIT	20
4.5.3. QUALITY-DEFINED MEASURE OF VERTICAL IIT	21
4.6. CORRELATION COEFFICIENTS BETWEEN MEASURES.....	22
4.6.1. TESTING THE CORRELATION COEFFICIENT	24
5. ANALYSIS	26
5.1. IIT AGGREGATED VS DISAGGREGATED	26
5.2. VIIT AGGREGATED VS DISAGGREGATED.....	27
5.3. VIIT DISAGGREGATED VS QUALITY-DEFINED	28
5.4. CORRELATION COEFFICIENT VS TRADE VOLUMES	29
6. CONCLUSIONS	31
LIST OF REFERENCE	33
APPENDIX	34

1. Introduction

To get a brief introduction to what this paper will deal with, I start with pointing out the questions that I wish to answer with this paper. Then I continue with a short look at what intra-industry trade is all about and why it has got a lot of attention from economists in the last decades. I end this chapter with a brief description of how the paper is structured chapter by chapter

Intra-industry trade (IIT), which is the focus of this paper, has in the last decades drawn much attention from economists around the world, resulting in a large number of studies on IIT. The reason why IIT got this attention was that when the phenomena of IIT were discovered, it called for a modification of the existing theory of international trade. International specialization needed no longer take the form of different countries specializing in different industries.

Most recent studies on IIT focus on the factors that explain the quantity of intra-industry trade, vertical and horizontal. My focus will be on the different methods used when calculating IIT, vertical IIT and horizontal IIT. These measures are important tools when analysing international trade and its effects on the welfare of countries.

1.1. Purpose

The purpose of this paper is to answer the following questions:

- Does the aggregation level of the data affect the result when measuring intra-industry trade?
- Does the aggregation level of the data affect the result when intra-industry trade is divided into vertical and horizontal intra-industry trade?
- What is the effect on the measures of vertical intra-industry trade when the trade data is classified as low or high quality, where the quality levels are self-defined?
- How is trade volume affecting the correlation coefficient of the different measures, do small or no bilateral trade flows affect the conclusions about the different measures?

1.2. Disposition

The first part of this paper focuses on the aggregational level of the data used in the calculations of intra-industry trade. Two different measures are used, Grubel-Lloyd based on aggregated and disaggregated data. The aggregated measure uses data at the 4-digit level, whilst the disaggregated index uses data at the 6-digit level.

The second part of this paper uses the same two measures as in the first part and also a self-defined measure that I name quality-defined (Q-D). In this part the focus is on measuring vertical (and horizontal) IIT. To answer the questions that constitute the purpose of this paper it is sufficient to measure only one of vertical and horizontal IIT. The quality-defined index uses disaggregated data at the 6-digit level. The difference from the disaggregated Grubel-Lloyd measure is that I define trade as low or high quality, and then I calculate the measure from the newly created trade data. The trade data from the different quality-classes is then aggregated into high and low quality trade.

Intra-industry trade is interesting to measure for different reasons. First of all IIT, especially vertical and horizontal IIT, gives valuable information about the composition of the flow of traded goods. Secondly, it reveals information about the industrial structure. Third of all the trade flows are interesting since changes in trade flows can have potential implications for the welfare of a country. Intra-industry trade based on horizontally differentiated products is associated with low adjustment costs whereas vertically differentiated products are associated with significantly higher adjustment costs.

Chapter 2 in this paper consists of the theoretical ground for the analysis. The method and the data used in this paper are presented in chapter 3. Chapter 4 presents the calculations used in the paper. The analysis is included in chapter 5, and finally the conclusions are included in chapter 6.

2. Theory

Before discussing the method and data used, a brief theory introduction is required. I start with an introduction to trade economics, and then the passenger car industry is presented. Vertical and horizontal IIT is discussed in this chapter.

2.1. Background theory

In the history of economic thought intra-industry trade has not always been recognized or accepted. With certain assumptions intra-industry trade is not a probable outcome. When trade depends on comparative advantages, the markets are perfectly competitive and economies of scale are assumed not to be present, then economic theory will conclude that no intra-industry trade exists. The theory predicts country-specific specialization in different sectors. Specialization is based on competitive advantages, which is further based on factor endowments, technology and natural resources. These conditions lead to the assumption that production and consumers spread out geographically, a dispersed industrial structure is created. Countries exported their excess surplus in sectors where the country had a comparative advantage, and imported goods in sectors where they had an excess demand.

A couple of decades ago, economists realized that intra-industry trade is a substantial part of total trade, especially bilateral trade between countries with similar relative factor endowments¹. It's then convenient to draw the conclusion that the assumptions made historically in economic theory is not realistic for today's economic environment. Furthermore a dispersed industrial structure does not exist, today; concentrational forces are much stronger, leading to a spatially concentrated industrial structure. The following is true for most economies; many markets are imperfectly competitive, trade costs have dramatically been reduced, economies of scale exist in many industries and the conclusions made in history about specialization are not correct today. Instead countries can both import and export goods of the same sector, intra-industry trade is then consistent with modern economic theory.

¹ See e.g. Gullstrand 2002

IIT is large for trade between countries with relatively similar (often high) income per capita and consumers with similar demand. A lot of studies have focused on what factors affect the amount of IIT, income per capita and demand patterns are just two of these factors.

Intra-industry trade exists since products are differentiated within a sector, consumers has, for most products, a demand for differentiated goods. In addition, consumers are assumed to have a love for variety or a favourite variety behaviour, meaning that a consumer is assumed to prefer to choose between more varieties than less. The love for variety approach (Dixit-Stiglitz) assumes that consumers gain higher utility the more varieties that are available. The favourite (or ideal) variety approach (Lancaster) assume that a consumer has a favourite variety and when the number of varieties increase the probability of a variety close to the favourite increases, leading to a higher utility for the consumer².

The effect of the aggregational level is interesting to measure since at the aggregated level trade within a sector can be horizontal (vertical) even if the trade within all sub-sectors are vertical (horizontal).

Many studies have used hedonic regressions to relate the prices of cars to the characteristics of the cars. A characteristic that always is included is engine volume and/or horsepower. The result of these studies is that engine volume is significantly affecting the price of a car.³

2.2. Heckscher-Ohlin – and later versions of their model

Traditional theory, referring to the theoretical developments of Heckscher-Ohlin, can through comparative advantage explain the existence of vertical intra-industry trade, whilst the more modern theories must be used to explain horizontal IIT. Falvey used a traditional 2x2x2 (goods, factors, countries) approach to show that inter-industry trade and vertical IIT can simultaneously exist.⁴ Through the Chamberlin-Heckscher-Ohlin-model⁵ there can be scope for horizontally differentiated products, which vary across different industries. The difference in scope depends on the minimum efficient scale of production, which also affects the efficient number of firms within an industry. Horizontal IIT increase when one of the

² See e.g. Markusen et al (1995)

³ See e.g. Irandoust (1998)

⁴ Quoted by: Greenaway, Hine, Milner (1995)

⁵ Greenaway, Hine, Milner (1995)

following occurs; decreased minimum efficient scale of production *or* increased number of firms *or* increased number of varieties.

If a product requires large volumes to be produced for the product to be profitable for a firm, then few firms will exist relative to the volume demanded. Few firms lead to less two-way trade flow of that specific product, leading to lower Horizontal IIT.

The Chamberlin-Heckscher-Ohlin-Samuelson model has been the base for most developments in IIT research. The model assumes that a firm within an industry differentiates its products from the product of other firms within the same industry; hence a number of varieties exist within the industry. These varieties are demanded by the consumers who are assumed to have a demand for varieties. Each variety is produced with increasing returns to scale and identical factor intensity. The CHOS-model will lead to the possibility of Inter-Industry trade as well as vertical and horizontal IIT to exist within an industry or a sector.

Free entry in combination with small economies of scale allow for a large number of firms, and monopolistic competition within the industry. Larger economies of scale, *ceteris paribus*, allows for a smaller number of firms, which also implies a smaller share of IIT in the industry.

2.3. Passenger car industry

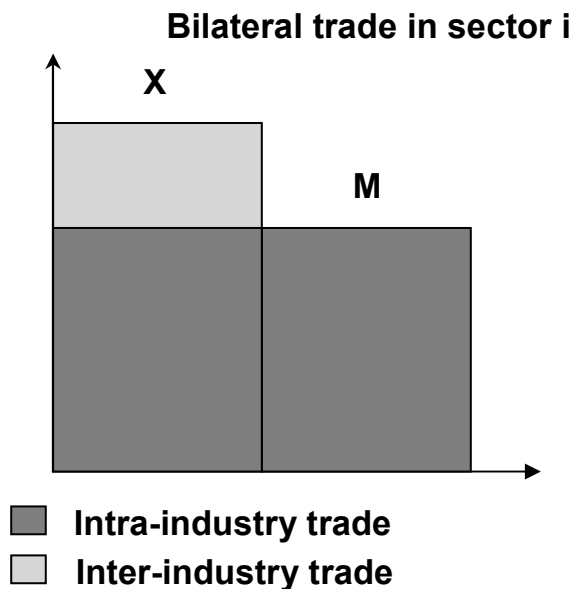
After a short introduction to trade economics, it can be valuable to present the passenger car industry in this context. The passenger car industry is characterized by large minimum efficient scale, and high initial costs. These factors lead to an industry with few firms. Most of these firms have a differentiated product portfolio that matches the consumers demand for different varieties. Even if most producers have a differentiated product portfolio, their products are often classified as belonging to a specific quality-segment.

In this paper, not all included countries has a production of passenger cars e.g. Finland, whereas some countries, like Germany, are amongst the world leading producers of passenger cars.

2.4. Intra-industry trade – Horizontal and vertical

Inter-industry trade is defined as trade between two countries where the goods are from different sectors. Intra-industry trade (IIT) is then when the traded goods are of the same sector. Graphically intra- and inter-industry trade can be presented as follows for any sector i .

Figure 1. Inter- and intra industry trade



Mathematically the same relation can be expressed as:

$$(1) \quad IIT = 2 * \text{Min}(X; M)$$

Where X and M is the value of exports and import of bilateral trade flows. In the graph minimum of export and import is the import, therefore the IIT is twice the amount of the imports.

IIT can further be disentangled in vertical and horizontal intra-industry trade, so that all IIT is either vertical or horizontal. Horizontal intra-industry trade is when imports and exports, within a specific industry during a specific time interval, are composed of products of the same quality. The time interval that is studied is often a year, that is the case in this paper as well. Vertical IIT trade is then when the traded goods, are instead of different quality.

Mathematically the following can be said about the following trade concepts:

$$(2) \text{ TotalTrade} = \text{IntraIT} + \text{InterIT}$$

Where

$$(3) \text{ IntraIT} = \text{HIIT} + \text{VIIT}$$

Measuring vertical (VIIT) and horizontal intra-industry trade (HIIT) requires a definition of quality. Most economists assume that price reflects quality, where price is a measure of value per unit, a unit-value. When relative unit values are outside a specific range, that is defining what is horizontal IIT, any IIT is classified as vertical.⁶

⁶ Abd-El-Rahman (1991)

3. Method

The method and data used in this paper are discussed in this chapter. The Grubel-Lloyd index is presented, together with its weaknesses.

This paper analyses three different measures, aggregated, disaggregated and Quality-defined. The aggregated measure uses data at the 4-digit level, the disaggregated and Quality-defined measure uses data at the 6-digit level.

The data used in this paper is classified according to the Harmonised System (HS).⁷ The reason for using the HS is that for the passenger car sector the HS offers more disaggregated data. The HS separate the subgroups according to the engine volume of the car, a factor that will in the quality-defined measure be used as a determinant of quality. Within the passenger car sector, data is available at the 6-digit level, divided into 8 subgroups. The following subgroups are used in the HS:

Table 1. HS classification

Subgroup	Engine size
8703.21	X < 1000 cc
8703.22	1000 < x < 1500 cc
8703.23	1500 < x < 3000 cc
8703.24	x > 3000 cc
8703.31	x < 1500 cc (diesel)
8703.32	1500 < x < 2500 cc (diesel)
8703.33	x > 2500 cc (diesel)
8703.90	Other

In the SITC-system passenger cars are included in the group 781.2 which corresponds to 8703 in the HS. The SITC-system does not have any subgroups for passenger cars; hence the most disaggregated data is on the 4-digit level for the SITC-system, while it is at the 6-digit level for the HS. The class 8703.90 (other) is not included in this study; since trade in this class is rare and when trade exists, it is of very low volumes and values.

⁷ The trade data is from SourceOECD

For the Quality-defined index, there was a need for defining what is high and low quality. The data set is divided into subgroups according to the engine volumes, and furthermore it is divided into what fuel the car uses, gasoline or diesel. The classification in the HS uses different engine size intervals for gasoline and diesel, making it somewhat more difficult to make a good definition between high and low quality. I decided that cars with engine volume above 1500 cc, are of high quality while those with engine volumes below 1500 cc are of low quality. This means that goods traded in the subgroups 8703.21, .22 and .31 should be considered as low quality goods, while goods traded in the subgroups 8703.23, .24, .32 and .33 should be considered as high quality goods.

3.1. Grubel-Lloyd index

In this paper the Grubel-Lloyd index (GL-index) is used.⁸ The GL-index that Grubel and Lloyd presented in 1975 has been used frequently since then. They presented their measure as follows:

$$(4) B_i = \left[(X_i + M_i) - |X_i - M_i| \right] * 100 / (X_i + M_i)$$

The index can be presented in a number of ways, I prefer the following:

$$(5) B_i = \frac{2 * \text{Min}(X_i, M_i)}{X_i + M_i}$$

The GL-index takes values between 0 and 1, where 0 means that all trade is inter-industry while 1 means that all trade is intra-industry. The GL-index is calculated based on bilateral trade flows. When calculating an index, for a group, based on more disaggregated indexes, from subgroups, the index is weighted with trade shares.

Since the GL-index is calculated as IIT divided by total trade, the GL-index should be interpreted as IIT's share in total trade. The index is therefore in some studies multiplied by a factor 100 to obtain values in percents. In this study I could have used both methods since it is not the actual value of the measure that I am interested in, but the results of the different measures.

⁸ Grubel & Lloyd (1975)

3.1.1. Criticism towards the Grubel-Lloyd index

The Grubel-Lloyd index is not a perfect tool for measuring IIT, it is sometimes criticised but it is still the most frequently used measure. There is no other index that is widely recognised as superior to the Grubel-Lloyd index.

A problem with the Grubel-Lloyd index is the aggregational level of the data.⁹ If the data is highly disaggregated, products that should be included in the same industry are in the data included in different subgroups, and therefore treated as if they are not from the same industry. The opposite is true for aggregated data, different products can be classified as if they were from the same industry even if they are not. In the case of the passenger car industry, the problem with the aggregational level is not a big problem, especially since I choose to exclude trade in class 8703.90 Others.

More criticism has been directed towards the Grubel-Lloyd index, e.g. the effects of multilateral versus bilateral trade flows, the effect of unadjusted trade imbalances. These weaknesses are not important for my paper and are therefore not discussed further.

3.2. Correlation between measures

To evaluate the different measures I will use two statistical measures, the correlation coefficient and the coefficient of variation. The correlation coefficient is a measure that describes how two samples vary together. The correlation coefficient is calculated using both the covariance between the samples as well as the variances within each sample. The coefficient of variation uses the mean correlation coefficient and the standard deviation of the correlation coefficients. I calculate one correlation coefficient and one coefficient of variation for each measure (totally five different measures) for each country.

⁹ See e.g. Markusen et al. (1995)

3.3. Indexes

Index i represents year and index j represents subgroup (6-digit level) within the passenger car group. Index k represents the examined country while index l represents the trading partner country.

$$\begin{aligned} i &= \{1996 - 2002\} \\ j &= \{21, 22, 23, 24, 31, 32, 33\} \\ (6) \quad k &= \{AUS, DEN, ESP, FIN, FRA, GBR, GER, GRC, IRL, ITA, NLD, PRT, SWE, XBL\} \\ l \neq k, \quad l &= \{AUS, DEN, ESP, FIN, FRA, GBR, GER, GRC, IRL, ITA, NLD, PRT, SWE, XBL\} \end{aligned}$$

The index $j = 21$ corresponds to 8703.21, $j = 22$ corresponds to 8703.22 and so on. XBL represents Belgium and Luxembourg, the original data set is split up between Belgium and Luxembourg between 1999-2002 while it is presented together in 1996-1998. Therefore I choose to treat Belgium and Luxembourg as one country in this study. In addition to these indexes I also use High and Low as indexes in the calculations leading to the Quality-defined measure.

4. Calculations

The analysis of this paper is based on a large data set, but the data also requires extensive calculations before it can be used; therefore I devote this chapter to present the equations for those calculations that has been done.

Index k represents the country, in the initial calculations (or formulas) the k-index is left out due to simplicity. First let's imagine that we only calculate values for one country. Index l represents the trading partner country, the l-index is also left out initially for the same reason.

$$(7) \quad TT_i = EXP_i + IMP_i$$

$$(8) \quad TT_{ij} = EXP_{ij} + IMP_{ij}$$

TT_i (Total trade) is calculated for all years and for all pair of countries, TT_{ij} is calculated for all years, for all countries and for all classes.

$$(9) \quad IIT_i = 2 * MIN(EXP_i, IMP_i)$$

$$(10) \quad IIT_{ij} = 2 * MIN(EXP_{ij}, IMP_{ij})$$

IIT_i is calculated for all years and for all pair of countries, IIT_{ij} is calculated for all years, for all countries and for all classes. IIT_i is calculated based on aggregated data, it is not the sum of a number of IIT_{ij}'s. When calculating for example unit-values (see equation 12) it doesn't matter if you use aggregated data or you sum the disaggregated data, but when calculating IIT the result differs between the two alternatives.

4.1. Unit-value calculations

Unit-value calculations are based on the data where values are measured in thousands of dollars and units are measured in tonnes.

$$(11) \quad UV_{EXP_{ij}} = \frac{Value_{EXP_{ij}}}{Units_{EXP_{ij}}}, UV_{IMP_{ij}} = \frac{Value_{IMP_{ij}}}{Units_{IMP_{ij}}}$$

Disaggregated unit-values are calculated for both export and import, for all classes, for all years and for all pair of countries.

$$(12) \quad UV_{EXP_i} = \frac{\sum_j Value_{EXP_{ij}}}{\sum_j Units_{EXP_{ij}}} = \frac{Value_{EXP_i}}{Units_{EXP_i}}$$

Aggregated unit-values are calculated for both export and import for all years and all pair of countries.

4.2. Classifying IIT

The following equation is used to measure whether the intra-industry trade is vertical or horizontal:

$$(13) \quad 1 - \alpha \leq \frac{UV_{EXP}}{UV_{IMP}} \leq 1 + \alpha$$

where UV_{EXP} and UV_{IMP} are unit-value of export and import. There is no obvious value for α that is correct, in research 0.15 and 0.25 is frequently used. In this paper I use 0.15, which is the most frequently used. I also use a self-defined method of classifying trade into high and low quality goods through engine size. Greenaway, Hine and Milner¹⁰ make the following definition:

“Horizontal IIT was first defined as the simultaneous export and import of a 5 digit SITC product where the unit-value of exports (measured f.o.b.) relative to the unit-value of imports (measured c.i.f.) was within the range of $\pm 15\%$.”

If the equation above (13) holds then the intra-industry trade is defined as horizontal, if it does not hold then IIT is vertical. Whether 15 % is enough or not has been discussed frequently, especially since the theoretical assumption of perfectly informed consumers is not realistic in most cases. With this assumption, we require consumers to have full information of all goods prior to the purchase; this is especially unrealistic in a sector like the car sector with numerous varieties. The assumption of perfectly informed consumers has been discussed in a number of

¹⁰ Greenaway, Hine & Milner (1995)

paper, Caves and Greene¹¹ made a study of quality and prices. With the assumption of perfect information the correlation between price and quality should be 1. The study finds that the correlation is positive but far from 1, this can be explained by a lot of factors. For this paper the discussion if 15% or 25% is the best alternative, is not critical since I am not interested in the actual values of IIT but the correlation between the different ways of measuring IIT.

One factor that I find interesting for my sector, passenger cars, is how one is supposed to measure quality. Since a passenger car has got so many different functions in addition to working as a means of transportation. Another factor that is of interest to my paper is the cost of acquiring information about all varieties of goods. With perfect information you, as a consumer, are assumed to be able to rank all varieties of the goods according to their quality. It would be very expensive for each consumer to acquire all information about all models of passenger cars available. Therefore it is necessary to keep in mind that the correlation between quality and price is not perfect for the passenger car sector, but it is reasonable to believe that the correlation is positive.

4.2.1. Criticism towards the unit-value approach

If quality is positively correlated with weight, then high quality goods create underestimated unit-values since the denominator in the calculations are biased upwards.¹² Therefore the ratio of import and export unit-value can be underestimated, which result in an overestimation of horizontal IIT. I have not found any studies that have examined the relation between quality and weight for the passenger car industry but I think that it is reasonable to think that a positive correlation exists. If so, the values of vertical IIT can be somewhat biased downwards, which can have some effect when comparing vertical IIT measured with disaggregated data and the quality-defined measure.

4.3. Inter industry trade, IIT, Vertical IIT and Horizontal IIT

Calculations are carried out both on the aggregated and the disaggregated level, but for simplicity the disaggregated calculations are shown below.

¹¹ Caves and Greene (1996)

¹² Greenaway, Hine & Milner (1994)

$$(14) \text{ if } IIT_{ij} = 0 \rightarrow TT_{ij} = InterIT_{ij}$$

If there is no IIT in a class, a specific year then total trade must equal InterIT. This is all the attention that InterIT will get in this paper. From now, I only focus on IIT.

$$\text{if } \frac{UV_{ij}^{EXP}}{UV_{ij}^{IMP}} > 1 + \alpha$$

$$(15) \text{ or } \frac{UV_{ij}^{EXP}}{UV_{ij}^{IMP}} < 1 - \alpha$$

$$\text{then } IIT_{ij} = VIIT_{ij} \text{ otherwise } IIT_{ij} = HIIT_{ij}$$

If the difference between unit-value of exports and imports is outside the interval of $1 \pm \alpha$ then IIT is classified as vertical IIT, otherwise it is classified as horizontal IIT. A result from this definition is that in disaggregated data calculations all IIT within a subgroup (IIT_{ij}) is either vertical or horizontal, for aggregated data calculations all IIT within a group (IIT_i) is either vertical or horizontal.

$$(16) HIIT_{ij} = TT_{ij} - InterIT_{ij} - VIIT_{ij}$$

HIIT can be calculated in a number of ways, but with all previous calculations the easiest way to get HIIT is to subtract InterIT and VIIT from TT. For subgroups either vertical or horizontal IIT is zero, but if one would aggregate disaggregated data then both horizontal and vertical IIT could be larger than zero.

4.4. Measures of IIT

Now we have to include the k, and l-indexes, where k is the country and l is the partner country.

4.4.1 Aggregated measure of IIT

The aggregated measure does not use the disaggregated trade data that in this case is available. The aggregation level is 4-digit (8703). The Grubel-Lloyd index is calculated as follows:

$$B_{ikl}^{Agg} = \frac{IIT_{ikl}}{TT_{ikl}}$$

(17) where

$$IIT_{ikl} = 2 * \text{Min}(EXP_{ikl}, IMP_{ikl})$$

B_{ikl}^{Agg} , or the trade weighted Grubel-Lloyd index, which I call Aggregated IIT (IIT^{Agg}) is calculated for all years and for all bilateral trade flows. For each country I get 91 (13*7 (number of trading partners*number of years)) different IIT^{Agg} .

4.4.2. Disaggregated measure of IIT

The disaggregated calculations use the 6-digit level data.

$$(18) \quad B_{ijkl} = \frac{IIT_{ijkl}}{TT_{ijkl}}$$

$$\text{where } IIT_{ijkl} = 2 * \text{Min}(EXP_{ijkl}, IMP_{ijkl})$$

B_{ijkl} is calculated for all years, for all classes and for all pair of countries.

$$(19) \quad B_{ikl}^{Disagg} = \sum_j \left(B_{ijkl} * \frac{TT_{ijkl}}{TT_{ilk}} \right)$$

B_{ikl}^{Disagg} , which I call Disaggregated IIT (IIT^{Dis}) is calculated for all years and for all bilateral trade flows. The B_{ijkl} are weighted, where the weight is the total trade within class z between country a and country b a specific year divided by the total trade between country a and b that year. To get IIT^{Dis} the weighted indexes are summed up over all classes, this summation result in 91 different IIT^{Dis} .

4.5. Measures of vertical IIT

To separate indexes for IIT and vertical IIT, I choose to use the letter C for indexes of vertical IIT instead of B, as used in indexes for IIT.

4.5.1. Aggregated measure of vertical IIT

The aggregated measure of vertical IIT is calculated as follows:

if

$$\frac{UV_{ikl}^{EXP}}{UV_{ikl}^{IMP}} > 1 + \alpha$$

or

(20) $\frac{UV_{ikl}^{EXP}}{UV_{ikl}^{IMP}} < 1 - \alpha$

then

$$IIT^{Agg} = VIIT^{Agg}$$

$VIIT_{ikl}^{Agg}$ is calculated for all years and for all bilateral trade flows. If the trade is classified as vertical then the Grubel-Lloyd index for vertical IIT is equal to the aggregated Grubel-Lloyd index for IIT. This is true since at the aggregated level all IIT is either vertical or horizontal; hence if IIT is classified as vertical then all IIT is vertical.

To get the index the following calculation is performed:

(21) $C_{ikl}^{Agg} = \frac{IIT_{ikl}}{TT_{ikl}}$

The index is calculated for all years and all bilateral trade flows.

4.5.2. Disaggregated measure of vertical IIT

The disaggregated measure of vertical IIT is calculated as follows:

if

$$\frac{UV_{ijkl}^{EXP}}{UV_{ijkl}^{IMP}} > 1 + \alpha$$

or

(22) $\frac{UV_{ijkl}^{EXP}}{UV_{ijkl}^{IMP}} < 1 - \alpha$

then

$$IIT_{ijkl} = VIIT_{ijkl}$$

$$C_{ikl}^{Disagg} = \sum_j \left(\frac{VIIT_{ijkl}}{TT_{ijkl}} * \frac{TT_{ijkl}}{TT_{ikl}} \right)$$

C_{ikl}^{Disagg} is calculated for all years and for all bilateral trade flows. All VIIT in subgroups within a specific group is weighted and then added together to the disaggregated vertical IIT measure.

4.5.3. Quality-defined measure of vertical IIT

The quality-defined calculations use the 6-digit level data. Total trade is divided up in low and high quality as follows:

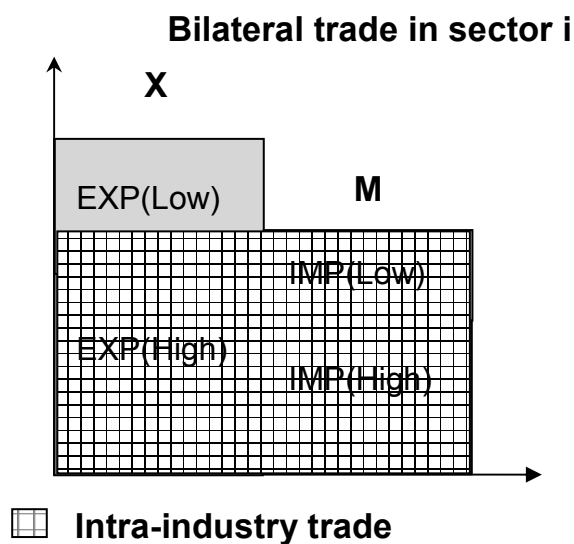
$$TT_{ikl}^{Low} = TT_{i21kl} + TT_{i22kl} + TT_{i31kl}$$

(23) and

$$TT_{ikl}^{High} = TT_{i23kl} + TT_{i24kl} + TT_{i32kl} + TT_{i33kl}$$

IIT for the Quality-Defined measure is calculated in the same way as for the aggregated measure. The figure below illustrates this calculation, where imports and exports from the different quality levels do not affect the IIT value.

Figure 2. Imports and exports according to the quality-defined trade data



The Quality-defined measure of vertical intra-industry trade is based on the following calculations:

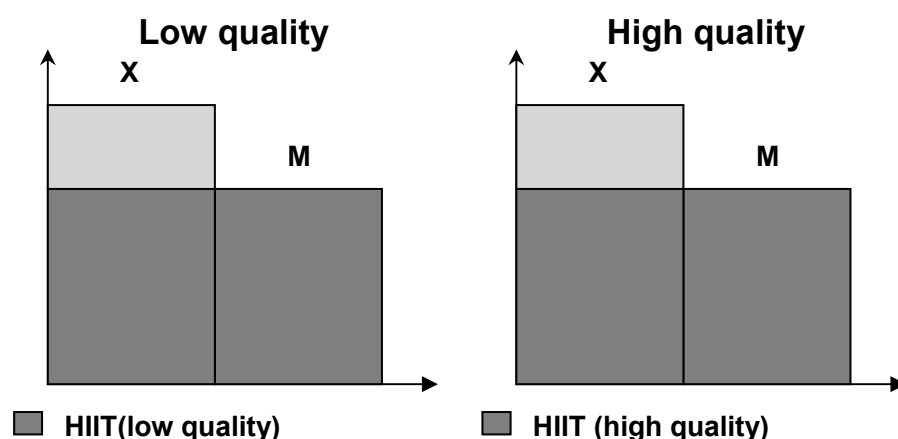
$$VIIT_{ikl}^{Q-D} = IIT_{ikl} - HIIT_{ikl}^{Q-D}$$

(24) where

$$HIIT_{ikl}^{Q-D} = 2 * \text{Min}(IMP_{ikl}^{Low}, EXP_{ikl}^{Low}) + 2 * \text{Min}(IMP_{ikl}^{High}, EXP_{ikl}^{High})$$

The $HIIT_{ikl}^{Q-D}$ is the sum of the HIIT in each of the quality groups, HIIT (low quality) and HIIT (high quality). Graphically the $HIIT_{ikl}^{Q-D}$ can be describes as done below.

Figure 3. Horizontal IIT according to the quality-defined method



The index is calculated with the following equation:

$$(25) C_{ikl}^{Q-D} = \frac{VIIT_{ikl}^{Q-D}}{TT_{ikl}}$$

C_{ikl}^{Q-D} is calculated for all years and for all bilateral trade flows.

4.6. Correlation coefficients between measures¹³

If two variables vary positively and perfectly, the correlation coefficient will equal 1.00. On the other hand, if two variables vary oppositely and perfectly, then the correlation coefficient will equal -1.00. If there is no correlation between the variables the correlation coefficient is zero.

The correlation coefficient is calculated as follows:

¹³ (26), (27) and (28) Körner (2000)

(29) http://economics.about.com/cs/economicsglossary/g/coefficient_v.htm

(31) www.shsu.edu/~icc_cmf/cj_685/mod10.doc

$$(26) \rho_{x,y} = \frac{Cov(X,Y)}{\sigma_x * \sigma_y}$$

Where $Cov(X,Y)$ is the covariance between the two samples (called x and y) and σ_i is the standard deviation in sample i (where i is either x or y). The standard deviation is calculated as follows:

$$(27) \sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$$

and the covariance is calculated as follows:

$$(28) Cov(X,Y) = E(X*Y) - E(X)*E(Y)$$

The standard deviation is a method for describing the spread in a population. The standard deviation is a measure that for a sample describes the average deviation from its mean, while the covariance is a measure describing the linear dependence between two variables.

The covariance and the correlation coefficient are both measures of how two variables vary together, there are some advantages with using the correlation coefficient. First, the value of the correlation coefficient is always between -1.0 and 1.0. Second, the correlation coefficient is the same no matter what units you are using; if you use dollars or thousands of dollars does not matter they return the same correlation coefficient.

The coefficient of variation is:

$$(29) \tau = \frac{\sigma}{\mu}$$

hence the coefficient of variation (CV) is the standard deviation divided by the mean. It is a measure that relates the standard deviation of the data to the mean value. The advantage with the CV is that it makes it possible to compare the variations of populations where the mean value differs significantly. The disadvantage with the measure is that the CV is sensitive to low mean values, leading to very high values of CV.

$$(30) \rho_i^k = [B_{ikl}^{Agg}; B_{ikl}^{Disagg}]$$

The equation above should be interpreted as the correlation coefficients, between the measure based on disaggregated data and the measure based on the aggregated data, are calculated based on the Grubel-Lloyd indexes for a certain country, a specific year. Therefore we get one correlation coefficient for each country, each year and for each test.

The same calculations are repeated for the following tests; vertical IIT based on disaggregated versus aggregated data and vertical IIT based on disaggregated data versus the quality-defined measure.

When these calculations are done, a mean of the correlation coefficients for each country is calculated, as well as the standard deviation and the coefficient of variation for each country's correlation coefficients. It is these figures, mean correlation coefficient and the coefficient of variation, that this study bases its conclusions on.

Values of correlation coefficient close to 1 means that the two measures produce almost the same result. A low value of correlation of variance implies low standard deviation and high mean correlation coefficient, which is also an outcome of similar results from the two measures.

4.6.1. Testing the correlation coefficient

To test if the correlation is significantly above zero, i.e. a positive correlation exist, we can use a t-test.

The hypotheses are as follows:

Null hypothesis: $H_0 : \rho = 0.0$

Alternative hypothesis: $H_1 : \rho > 0$

$$(31) t = \frac{r\sqrt{N-2}}{\sqrt{1-r^2}} \quad ^{14}$$

¹⁴ Rho is replaced by r to separate sample values from the population values.

Table 2. Critical values for testing correlation coefficients¹⁵

Significance	N=14	N=11
5% (*)	0.457	0.521
2,5% (**)	0.612	0.685
1% (***)	0.75	0.82

¹⁵ Körner (2000), in chapter 5 I explain why both N=14 and N=11 is included in the table.

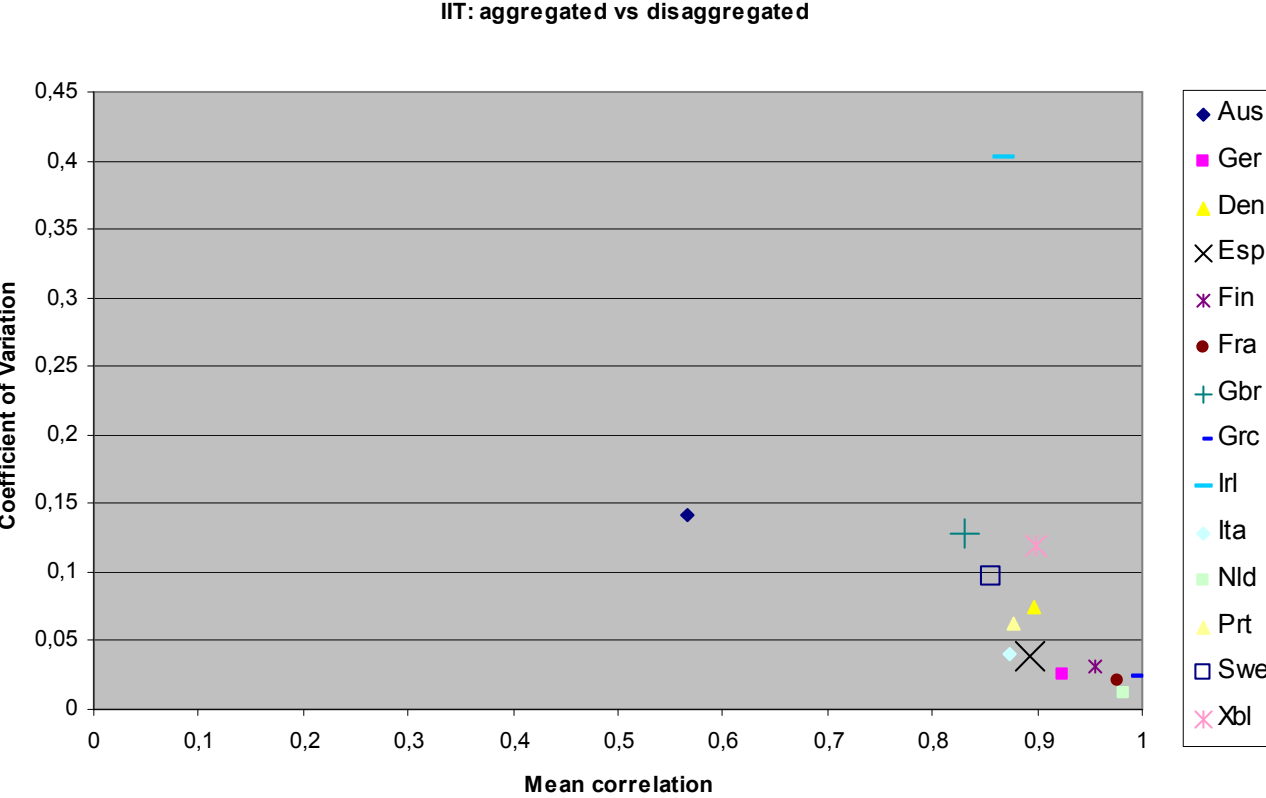
5. Analysis

In this chapter I analyse the different measures through values of mean correlation coefficient and coefficient of variation. Important to note is that the y-scale of the graphs are not identical. The fourth graph compares mean correlation with total trade volumes.

5.1. IIT aggregated vs disaggregated

Two different measures have been used to calculate intra-industry trade, based on different aggregation levels of the data.

Figure 4. IIT measure – aggregated versus disaggregated data



From the graph it is possible to draw the conclusion that for most countries the aggregated and disaggregated measure follows the same pattern, their mean correlation is close to one and the coefficient of variation is close to zero. The exceptions are Austria and Ireland, where Austria has a low mean correlation while Ireland has a high coefficient of variation. A reason

to why these countries have extreme values has not been found. All other countries have a mean correlation above 0.80 and a coefficient of variation below 0.20.

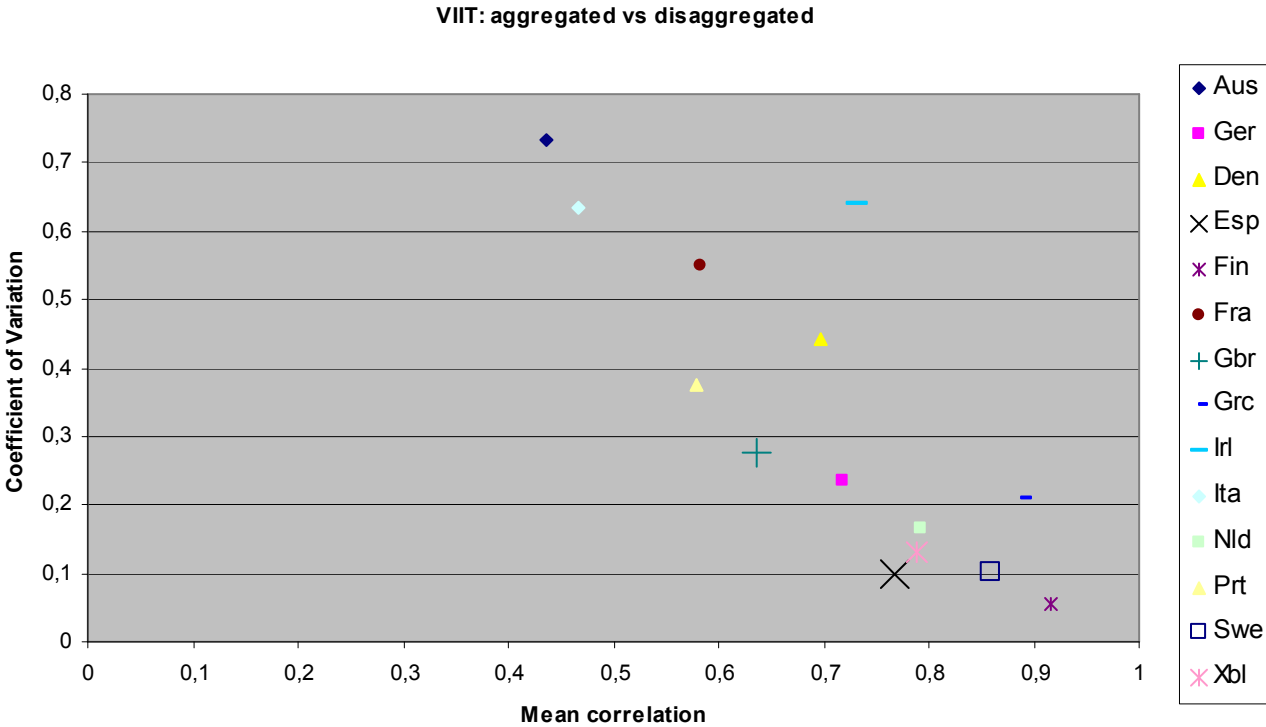
When testing the correlation coefficient as described in section 4.6.1, the following value is achieved:

$$t = 6.552***$$

we can therefore conclude that the correlation coefficient is significantly above zero, at the 1% level (three stars).

5.2. VIIT aggregated vs disaggregated

Figure 5. Vertical IIT measure – aggregated versus disaggregated data



When it comes to vertical IIT the aggregation level of the data affects the result more than in the previous test. For some countries, e.g. Finland, the two measures is highly correlated with a low coefficient of variation, whilst for some, e.g. Austria, the correlation is low and the coefficient of variation is high. In this case there are only two countries that have a mean correlation above 0.8 and a coefficient of variation below 0.2. This can be compared with the previous test, where all countries except two fulfilled those conditions.

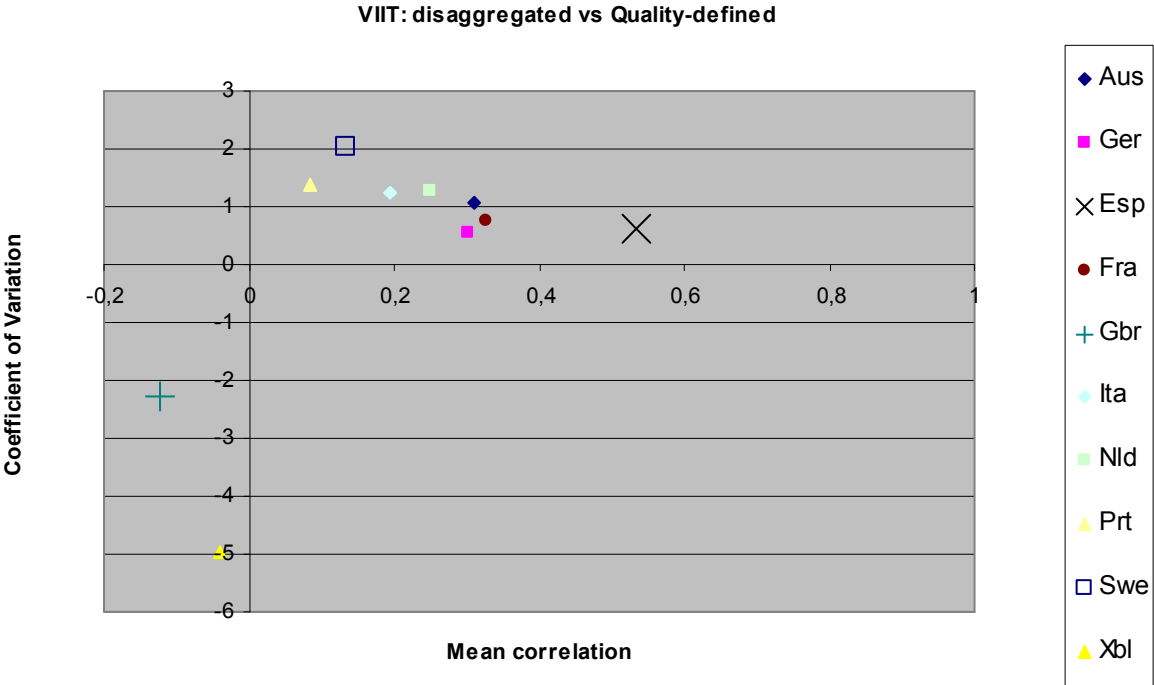
The result of the t-test is:

$$t = 3.443^{***}$$

we can therefore conclude that, at the 1% significance level, the correlation coefficient is above zero.

5.3. VIIT disaggregated vs quality-defined

Figure 6. Vertical IIT measure –disaggregated versus quality-defined data



Finland, Greece and Ireland are not included in the graph. Those countries had at least one year without a correlation coefficient. Correlation coefficients are missing when for a country all Vertical-IIT one year is zero for a measure. Denmark is also excluded from the graph since its coefficient of variation equals -73.2.

When comparing vertical IIT based on disaggregated data and data sorted according to my definition of high and low quality, the plots are more spread out. Spain has a high value of mean correlation and rather low coefficient of variation, while Great Britain and Belgium/Luxembourg has negative correlation coefficients (and coefficient of variations).

Defining quality as done in the Quality-defined measure gives a different result from what we get when data is sorted in subgroups according to engine size and quality is defined through unit-values.

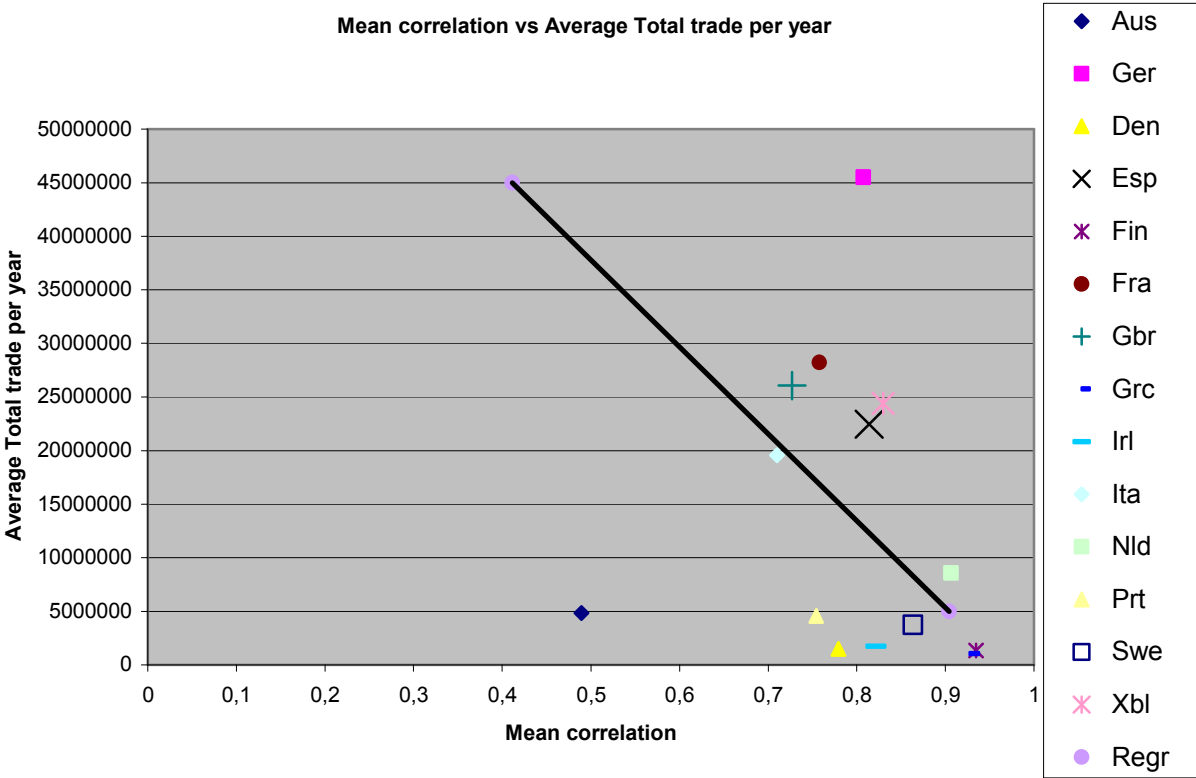
The result from the t-test when Finland, Ireland and Greece are excluded is:¹⁶

$$t = 0.545^*$$

Even with the more dispersed pattern we can through the t-test conclude that the correlation coefficient is significantly above zero, but only at the 5% level (one star).

5.4. Correlation coefficient vs trade volumes

Figure 7. Relation between correlation coefficient and total trade



A country where many bilateral trade flows are zero, no trade exists, has a high correlation, since both measures return the value zero when trade does not exist. Therefore it is interesting to see if there is a relation between mean correlation and average total trade per year. To investigate the relation a linear regression is inserted in the graph. There is a negative relation

¹⁶ If these countries are included the result of the t-test is: $t=0.725^{**}$

between trade volume and correlation, low trade volumes mean high correlation, and large trade volumes mean low correlation coefficients.

In the linear regression, that is based on the least-square method the both extremes Austria and Germany are left out¹⁷. The regression result in the following equation:

$$(27) \quad y = 78376038 - 81142292x$$

The determination coefficient, that describes the proportion of the total variation for the dependent variable that is explained by the linear relation with the independent variable. The determination coefficient for this linear regression is:

$$(28) \quad R^2 = 0.312329$$

which means that 31 % of the total variation in the sample is described by the linear relation, other variables, that are not included in this model, explain the remaining 69 %. Since the only independent variable in the regression is Total bilateral trade all of the 31% of the variation is described by total bilateral trade.

If the two extremes are not left out the determination coefficient decreases to:

$$(29) \quad R^2 = 0.017042$$

The extremes dramatically change the determination coefficient and the result of the analysis. In the case without the extremes approximately 30% of the variation is described by total bilateral trade, with the extremes that figure is approximately 2%.

¹⁷ When examining samples from a large population, extremes are often left out to avoid misleading results due to a few extremes. In this case it looks, in the graph, as if the regression would not be very different from the one in (27).

6. Conclusions

The theoretical ground that this paper utilizes has been presented, as well as the data, the method and all calculations. The previous chapter analysed the result, this final chapter summarizes the conclusions that can be drawn from this paper.

Measures of IIT based on aggregated compared to measures based on disaggregated data are highly correlated. Most countries have a correlation coefficient above 0.8, which means that the two measures are closely and positively correlated. The t-test results in a very high t-value which should be interpreted as a very high level of significance, when examining if the correlation coefficient is above zero. The aggregational level is of little importance since both methods return closely related values.

When measuring Vertical IIT with aggregated and disaggregated data the result is more dispersed, the correlation coefficients are on average lower than in the previous test. Still correlation coefficients are positively correlated. The t-test result in a high t-value, hence a high level of significance, which lead to the conclusion that the correlation coefficient is above zero. The conclusions that are made about the composition of bilateral trade differ depending on what level of aggregation that is chosen.

In the comparison of measurements between the measure based on disaggregated data and that based on the quality-defined data the result is even more dispersed. In this test the correlation coefficient for two countries were negative, even though the t-test result in a t-value that is significant, at the 0.05 level. I believe that the fact that I chose only two levels of quality is a great contributor to the large differences between the measures. Therefore it would be interesting to make a study where more quality levels are included, and where the quality levels are suited for the specific products.

When analysing the effect of total bilateral trade on the correlation, the result depends on whether the extremes are included or not in the regression. If they are not included, much of the total variation can be described by differences in total bilateral trade.

It could be interesting to further analyse the relation between total bilateral trade and mean correlation coefficient, then more sectors could be analysed to see if a high determination coefficient is rare or frequent.

List of reference

Published sources

Abd-el-Rahman, K.S. (1991) "Firms' Competitive and National Comparative Advantages as Joint Determinants of Trade Composition" *Weltwirtschaftliches Archiv*

Caves, Richard & Greene, David (1996), "Brands' quality levels, prices, and advertising outlays: empirical evidence on signals and information costs", *International Journal of Industrial Organisation*

Greenaway, David & Hine, Robert & Milner Chris (1994), "Country-specific factors and the pattern of horizontal and vertical Intra-Industry trade in UK", *Weltwirtschaftliches Archive*

Greenaway, David & Hine, Robert & Milner Chris (1995), "Vertical and horizontal intra-industry trade: A cross industry analysis for the United Kingdom", *The economic Journal*

Grubel, Herbert & Lloyd, P.J. (1975), *Intra industry trade: The theory and measurement of international trade in differentiated products*, Halsted Press, New York

Gullstrand, Joakim (2002) *Does the Measurement of Intra-Industry Trade Matter?*, *Weltwirtschaftliches Archiv*

Irاندoust, Manuchehr (1998), *Pricing Policy in the Swedish Automobile Market*, *Journal of Economics and Business*, New York

Körner, Svante (2000), *Tabeller och formler för statistiska beräkningar*, Studentlitteratur, Lund

Markusen, James & Melvin, James & Kaempfer, William & Maskus, Keith (1995) "International trade – Theory and evidence", McGraw-Hill, Singapore

Electronic sources

www.shsu.edu/~icc_cmf/cj_685/mod10.doc, 2006-05-08

http://economics.about.com/cs/economicsglossary/g/coefficient_v.htm, 2006-05-06

Data sources

SourceOECD, Organisation of Economic Co-operation and Development, Statistical Database, www.sourceoecd.org

Appendix

A1 – Mean correlation coefficient, standard deviation and coefficient of variation

		IIT: dis vs agg	VIIT: dis vs agg	VIIT: dis vs Q-D
Aus	Mean	0,566108218	0,437088781	0,310591615
	Stddev	0,080358431	0,320403307	0,333717743
	CV	0,141948887	0,73303942	1,074458312
Ger	Mean	0,923991915	0,718685253	0,301820856
	Stddev	0,023086628	0,169668891	0,167427346
	CV	0,024985747	0,236082333	0,554724244
Den	Mean	0,897328531	0,697686362	-0,00468231
	Stddev	0,067276508	0,308912671	0,342799453
	CV	0,074974221	0,442767248	-73,21160729
Esp	Mean	0,892861584	0,768023008	0,533582913
	Stddev	0,034942434	0,075725157	0,328291129
	CV	0,039135331	0,098597511	0,61525795
Fin	Mean	0,955552063	0,916961757	*0,175574855
	Stddev	0,030206136	0,050652958	*0,301083111
	CV	0,031611189	0,05523999	*1,714841863
Fra	Mean	0,976218382	0,582689499	0,327001992
	Stddev	0,02047979	0,319591306	0,25064274
	CV	0,020978698	0,548476172	0,766486892
Gbr	Mean	0,830591954	0,636559956	-0,1238648
	Stddev	0,106158451	0,175720815	0,280249623
	CV	0,127810594	0,276047548	-2,262544512
Grc	Mean	0,991136475	0,887314312	*-
	Stddev	0,023450682	0,186794987	*-
	CV	0,023660397	0,21051727	*-
Irl	Mean	0,867994425	0,730918123	*0,313499464
	Stddev	0,348751281	0,467382834	*0,596231224
	CV	0,401789771	0,639446224	*1,901857239
Ita	Mean	0,874261251	0,466951002	0,193176055
	Stddev	0,035718065	0,295952827	0,237253723
	CV	0,04085514	0,633798462	1,228173557
Nld	Mean	0,98192769	0,792752262	0,250790641
	Stddev	0,010989858	0,13032308	0,315707581
	CV	0,011192125	0,164393199	1,258849133

Prt	Mean	0,877960866	0,579649962	0,084745813
	Stddev	0,054964125	0,218285539	0,115703136
	CV	0,0626043	0,37658165	1,365296188
Swe	Mean	0,855873951	0,858855473	0,133260853
	Stddev	0,08286056	0,087980385	0,271443065
	CV	0,096813976	0,102439104	2,036930261
Xbl	Mean	0,897979575	0,788379284	-0,040933139
	Stddev	0,107059948	0,103518068	0,202909829
	CV	0,119223143	0,131304906	-4,957104004

* Values are based on calculations with missing values

A2 – T-test values

	IIT: dis vs agg	VIIT: dis vs agg	VIIT: dis vs Q-D	VIIT: dis vs Q-D
Mean	0,884035258	0,704986821	0,20493726	0,178680954
N	14	14	14	11
t-test	6,551680531	3,443423185	0,725318292	0,544810459
Significance	***	***	**	*

A3 – List of figures and tables

Figures

Figure 1. Inter- and intra industry trade

Figure 2. Imports and exports according to the quality-defined trade data

Figure 3. Horizontal IIT according to the quality-defined method

Figure 4. IIT measure – aggregated versus disaggregated data

Figure 5. Vertical IIT measure – aggregated versus disaggregated data

Figure 6. Vertical IIT measure –disaggregated versus quality-defined data

Figure 7. Relation between correlation coefficient and total trade

Tables

Table 1. HS classification

Table 2. Critical values for testing correlation coefficients