Master thesis for the Master of Economic Theory and Econometrics degree

# The gravity in Norwegian manufacturing industry

*Impacts on aggregate exports by adjustments in firms' extensive and intensive margins*

**Espen Nord Eidene**

**May, 2008**

**Department of Economics**

**University of Oslo**

# Preface

This thesis has been written during the first half of 2008. It was completed while I was a research assistant for Professor Karen-Helene Ulltveit-Moe who also was my supervisor during this paper. I would like to thank her for precise and insightful comments and constructive discussions while working with this thesis.

In addition, I am grateful to Andreas Moxnes who has been very helpful with eminent knowledge on estimation programming.

And at last, I would like to thank Finn Christiansen who proofread this thesis.

Of course, remaining errors and weaknesses are my responsibility.

Espen Nord Eidene

Oslo, May 2008.

# Table of contents

# 1. Introduction

*"Nations do not trade, it is firms that trade."* [1]

Historically, international trade flows have been analyzed from the perspective of a country or an industry. However, firms in international trade have received increased attention during the last decades. Recent papers have analyzed the role of firms and corporations in international trade. They find obvious differences between both exporting and non-exporting firms, but also within the group of exporting firms (Chaney (2004); Helpman, Melitz and Rubinstein (2007); Mayer and Ottaviano (2007); Bernard, Jensen and Redding (2007)). Exporting firms tend to be more productive, pay higher profits to owners and wages to workers than non-exporting firms within the same country. There are also major quantitative differences within the group of exporting firms, or internationalized firms as they are called by Mayer and Ottaviano (2007). Exporting statistics show how a relatively exclusive club of exporting firms accounts for a vast share of total exports. 10 % of the biggest exporters in seven European countries account for 70 – 95 % of the respective countries total exports in the manufacturing industry (Mayer and Ottaviano 2007)[2]. The empirical examples above prove the inequality in the performance of exporters and non-exporters.

This is a matter for firms within the same industry but also across industries. It supports the assumption adopted in newer trade theory where heterogeneous firms are important in explaining stylized facts as intra-industry trade and the home market effect. Testing new theses in trade theory and further investigations on which role firms play in international trade are of major importance. Only by doing this it is possible to better understand the role of firms in the global economy.

What are the driving forces behind the aggregated level of trade flows? Is it the number of firms involved in exporting and importing or is it the average volume each firm exports to each destination? I have analyzed the export pattern in Norway from 1999 and 2004 in the manufacturing industry. My hypothesis is that the impact from the number of exporting firms (the extensive margin) dominates the impact from average exports per firm (the intensive margin) on the aggregated trade flows. Recent research papers conclude that the extensive

[1] Quoted from Mayer and Ottaviano (2007), page 11.

[2] The seven countries are Sweden, Norway, Belgium, United Kingdom, Italy, France and Belgium,

margin is the most important one, since most changes in aggregate trade flows takes place along this margin (Bernard et al 2007; Mayer and Ottaviano 2007). I estimate three gravity equations to test if this is the case in Norway; one on the aggregated level of exports, one for the extensive margin and one for the intensive margin.

The role of firms in international trade needs to be studied. It is important to know as much as possible about the exporting firms' environment, but also in order to make the best possible trade policy. But there are several reasons why firms only recently have received increased attention. One is previous lack of data. Reliable and good trade data on firm level have only recently become easily available. Data were previously collected at an aggregated level of the specific industry or the county. The role of the individual firms' extensive and intensive margin could then obviously not be analyzed. The development in the literature of international trade theory is a second reason. Trade theories trying to explain international trade have traditionally not focused on the interplay between internationalized firms. There was a change in the theoretical pace as the traditional trade theories were unable to explain the empirical trade pattern and often were contradicted by stylized facts if a change occurred. A new branch emerged with a new focus on modifications in the traditional trade theory assumptions. The development of new trade theory made it possible to better study the interplay between firms in international trade.

Traditional trade theory has been highly influenced by thoughts from David Ricardo and Heckscher-Ohlin (HO). In their theories trade originates from differences in productivity or relative factor endowments which results in export in the countries' respective competitive advantage industries. This implies trade in goods from different industries between countries (inter-industry trade) and no trade in goods from the same industry between countries (intra-industry trade). The traditional HO framework with a 2-input-industry-country model implied a downplayed focus on the role of internationalized firms in both theoretical and empirical research papers. But as the older theories failed to explain the empirical pattern of trade new assumptions were developed. In this new branch of international trade, called new trade theory, the role of internationalized firms experienced increased attention and importance.

The new trade theory allowed for monopolistic competition, increasing returns to scale and product differentiation. Many of the ideas behind the new trade theory are built on methods from industrial organization (Krugman 1983). New trade theory opens for differences between firms in the same industry and supports findings of intra-industry trade (bilateral trade in differentiated goods between countries). It can also explain how similar sized economies tend to trade more with each other than with other countries (Bernard et al 2007). Intra-industry trade is often found between the OECD (Organization for Economic Co-operation and Development) countries while inter-industry trade is often between developing countries and OECD countries. Empirics show that increasing returns to scale seems to be more of importance in the OECD countries, while factor endowment seems to be an important determinant in the developing countries (Evenett and Keller 2002). Krugman (1983) and other argue that the new trade theory was developed at a later pace because of complexity in calculations resulting from specific underlying assumptions. Yet it predicts concepts which are more general than what the traditional theory did (ibid). The discussion of trade theory based on increasing returns started out in the 60's. But as the economies of scale models had to cope with imperfect competition, which at that time was complex, the development slowed down. Krugman speaks of intra-industry theory trade when referring to models explaining this phenomenon. Together with other trade theorists, he used a Chamberlinian approach with monopolistic competition to formalize the economics of scale models (Dixit and Norman (1980), Krugman 1983)[3].

There is also an increased focus on location and location of production in order to explain stylized facts in international trade, in addition to the modification of perfect competition and constant returns to scale assumptions in old trade theory (Krugman 1991; 2000). By making use of the gravity theory where distance between trading partners is one of two core variables, I link the new trade theory up with gravity model. The gravity equation links flows between two regions by economic mass variables and distance. First it was presented with great empirical success but a lack of strong theoretical foundation, but with time it has also gained a strong theoretical foundation. The equation is also capable of sorting out different economical hypotheses in international trade (Feenstra, Markusen and Rose (1999); Evenett and Keller (2002)). An example is how the gravity equation assumes distance as one of its explanatory variables. The distance variable can be seen as a proxy for different trade costs; shipment costs, melting iceberg cost or similar factors which makes goods more expensive to the importing country (Head 2003).

[3] Imperfect competition also involves oligopoly, in addition to monopolistic competition. However, oligopoly is more complex to model and monopolistic competition gives a good insight after all.

In order to investigate the pattern of exports in the Norwegian manufacturing industry I make use of gravity theory. I start by presenting the new trade theory and how it links to the gravity theory. I then present a theoretical framework for the gravity equation which forms the basis for my empirical gravity. The empirical model is then presented by first showing a basic empirical gravity equation. This is then extended with relevant variables when analyzing the manufacturing exports from Norway. In the same section I decompose the aggregated exports into the extensive and intensive margin in order to determine margins' impact on aggregated trade flow. Results from previously estimated gravity equations are reviewed are the results from regression are presented and discussed at the end of the paper.

The remainder of the paper is as follows. Section 2 presents the theoretical framework which forms the foundation for the empirical analysis and how the gravity equation fits with the new trade theory's assumption about monopolistic competition and heterogeneous firms. Section 3 presents the empirical gravity model being used to analyze the pattern of export over five years. Input data are also commented together with discussion and results of estimations. Section 4 concludes.

# 2. Theoretical framework

In this section I review theoretical literature devoted to the gravity theory. I give a brief overview of the development in the literature by reviewing papers from the last 30 years in section 2.1. Papers address both the previous theoretical challenges in gravity theory and inclusion of modern assumptions from the new trade theory such as increasing returns to scale and focus on location. In section 2.2 I review a theoretical model developed by Anderson and van Wincoop (2004) to give an up-to-date presentation of the theoretical gravity framework. The model ends up with a gravity equation which is point of departure for the empirical gravity equation in section 3.

## 2.1 The gravity and new trade theory

The popularity of applying the gravity equation on bilateral trade flows can be summarized with a list of three factors (Baldwin and Taglioni 2006):

- There is a need to investigate trade flows and gain new insight since trade between nations and regions is a central feature of all economical relationships.
- Data on variables used in a gravity analysis are easily accessible to researchers.
- A list of recent papers have made the gravity theory respected and have developed a set of standard tools addressed to solve challenges in an empirical research.

All three factors listed above are mentioned in the introduction of this thesis, but the last factor can be elaborated upon a bit further. I do this to get a better understanding of how the gravity theory copes with different theoretical models in international trade theory before the empirical analysis is presented.

The gravity characterization stems from Newton's "law of gravity" and is an analogy to this[4]. The most basic gravity equation predicts a country's export to be proportional to the size of the respective trading partners' gross domestic products. The bilateral trade between two countries is predicted to be increasing in their respective gross domestic products and diminishing with increased distance between two trading countries.

[4] The gravity equation has also been used to analyze other types of flows such as migration flows, tourism, equity flows, foreign direct investments and shipping movements.

Despite its rather simplistic mathematical formulation, the equation captures features in the international trade pattern in a remarkably realistic way. Today, it predicts trade equally well with data on both OECD countries and developing countries, though with different theoretical specifications (Evenett and Keller 2002). An example of the first is an estimated gravity equation on trade data from 16 OECD countries (Feenstra 2004). The reported $R^2$ from this regression is 0.40, which implies that the model explains nearly half of the changes in the trade flows.

However, the gravity equation long suffered from lack of strong theoretical microeconomic foundations (Anderson 1979; Bergstrand 1985; Mayer and Ottaviano 2007). Baldwin and Taglioni (2006) describe the earliest theoretical foundations as "*hand-waving underpinnings*". But as more papers trying to develop a stronger theoretical framework for the gravity theory have been published, a richer and more valid theoretical foundation has been built. Recent papers are now focusing on the econometric issues in estimation of the gravity equation (see e.g. Anderson and van Wincoop (2003); Baldwin and Taglioni (2006).

The first papers applying gravity theory on empirical trade data were published in the early 1960s. Anderson (1979) was the first theorist to develop a stronger theoretical foundation which the gravity theory could rest on. His paper is often used as a sort of reference point when reviewing theoretical framework in gravity theory. Almost every gravity paper starts by reviewing his paper. He agreed with the gravity equation's strong explanatory power when applied to trade flows, but it was "(…) severely hampered by its unidentified properties". Anderson (1979) assumes that each nation is specialized in production of one unique good which cannot be perfectly substituted by goods produced in other nations. It is a well-known assumption today, but thought of as a very special assumption at that time.

Other papers claimed, somewhat loosely, that the gravity equation was a reduced form of a system of four equations in a partial equilibrium (Bergstrand 1985). Bergstrand (ibid) solves the gravity equation as a reduced function from general equilibrium model with nationally differentiated goods. In contrast to Anderson (1979), he assumes perfect substitutability between goods across countries.

Bergstrand (1985) presents six assumptions which make it feasible to solve the gravity equation. The first is exogenous explanatory variables. To be defined as a proper gravity equation it must include the exporters and importers GDP exogenously, he argues. If country i is a small open economy in a global perspective, the demand of country i's goods in country j will be very small in relation to the aggregated demand from the other countries. This fits very well with the description of Norway's economy which I am studying. Norway has a small open economy in a global setting relative to the other countries. Assuming this about Norway let us treat the GDP in the exporting and importing country as exogenous. The second is identical preferences amongst consumers and production functions across countries. This ensures that estimated coefficients are constant across countries and are common assumptions in standard present trade models with both intra- and inter-industry trade. These two first assumptions specify what Bergstrand characterizes as a general gravity equation. "General" stems from the exogenous income variables on the right hand side and no restrictions imposed on the variables' coefficients. The last four is though as a bit restrictive in the way "general" on longer holds. They include perfect substitutability of goods internationally in production and consumption, perfect commodity arbitrage and zero transport costs. Bergstrand concludes that coefficient estimates supports the assumption of imperfect competition in production and products being differentiated by production location. His results imply an elasticity of substitution between internationally traded goods which exceed unity, and that substitution between traded and non-traded goods is below unity.

Anderson and van Wincoop (2003) argue, in a paper resting much on Anderson's (1979) first contribution, that there is no theoretical foundation for the gravity equation. In their opinion, the missing theoretical framework lead to a situation with biased estimates due to omitted variables and an inconsistent comparative static. The goal of their paper is to develop a valid theoretical equation and methods that can estimate the gravity equation without omitted variable bias, giving consistent comparative static and solving the border puzzle presented by McCallum (1995) when he analyzed the trade between the US and Canada. The omitted variable bias has also been discussed by other economists such as Bergstrand (1985) and Baldwin and Taglioni (2007)[5].

By the emergence of the new trade theory in the late 70s and the 80s gravity theory almost got too many theoretical foundations to rest on (Evenett and Keller (2002); Baldwin and Taglioni (2007)). The new trade theory utilizes economies of scale, imperfect competition and product differentiation into trade models and is consistent with stylized facts from the empirical trade pattern. Increasing returns capture how some industries tend to be dominated

[5] A presentation of how to solve the different biases is beyond the scope of this thesis, but can be studied in depth in papers such as Baldwin and Taglioni (2007).

by multinational corporations as Nestlé. Product differentiation and preferences for variety of goods in consumptions ensures bilateral trade within the same industry across countries as in the car industry between Germany and United States (Markusen and Venables 1995). It also captures how similar countries tend to trade more with each other than the traditional theory predicts as the stylized fact that most OECD countries trade mostly with each other. The amount of intra-industry trade is higher the more similar trading partners are in their industrial structure (Krugman 1983). The fact that much trade takes place between similar countries is also in line with the gravity theory. The inclusion of distance is predicted to dampen trade the further countries are from each other. Other papers also use new trade specifications when applying gravity on trade data.

Feenstra, Markusen and Rose (1999) use different estimated gravity equations to diversify between the different trade models. They test if the home market effect depends on different trade theories by applying the gravity equation on the different models. They compare traditional HO framework with increasing return and a differentiated good model versus a homogeneous good model. Their results from estimation vary consistently and confirm the gravity equation's applicability to different trade models. An example is how a differentiated good model predicts the home market effect and how a homogenous good model is without the effect due to entry costs. To why the gravity equation has gained success across different trade models, they conclude that even if the empirical specification is quite specific, the theoretical specification is general. A similar paper to Feenstra et al (1999) is the paper by Evenett and Keller (2002), where they investigate the consistency of the gravity equation in different trade theories.

However, the two papers differ in how they treat different types of trade. Evenett and Keller (ibid) use an index to separate intra-industry from inter-industry trade, while Feenstra et al. measure if goods are differentiated or homogeneous to separate their samples. Evenett and Keller's try to identify which theory actually accounts for the empirical success of the gravity equation, which they call the *identification problem* of the gravity model. They develop two gravity equations, one for perfect and one for imperfect specialization in production to investigate how trade is determined in the respective models. Their estimates find that the size of the differentiated goods sector and the share of intra-industry trade of total trade are positively correlated. This findig may ease the explanation of North- North trade by using scale economics and imperfect specialization in differentiated goods. However, they underline the overestimated amount of bilateral trade when a model is specified with perfect specialization.

Melitz (2003) incorporates an assumption of heterogeneous firms into a model with monopolistic competition and increasing return. The model is an extension to Krugman's model (Krugman 1980) and it incorporates productivity differences between firms and it shows how reallocation takes place within and across the comparative advantage and disadvantage sectors when a country is more exposed to trade. The trade exposure is followed by a higher competition level between firms which next imply higher export activity only from the most productive firms. This results in better economical performance in the competitive advantage industry. The performance can be measured by higher aggregate productivity level, increased profits and bigger market shares both domestically and abroad to the internationalized firms. He argues that the changes must be interpreted as long-run effects since it is a general equilibrium model. Melitz' concludes that increasing exposure to trade leads to increased welfare gains, but he reminds us of how losers will be found since the less effective ones will run out of business and trade will lead do reallocation of resources.

Chaney (2004) makes use of the gravity equation in a model where countries and firms are identical. But firms are heterogeneous in production and consumers have a preference for variety. He argues that the volume per existing exporter will increase when trade barriers decline. But in addition new firms that can afford to start exporting will results in an increase in the number of exporters as well. The effect on trade flows are dominated by changes in the extensive margin rather than the intensive margin when trade barriers decrease. Thus, Chaney rejects predictions from models with a representative firm.

The Melitz-model (Melitz 2003) is developed further by Bernard, Redding and Schott (2004). They develop a model with heterogonous firms, consumer love for varieties and relative endowment differences across countries. In their model they unite a traditional HO-model with monopolistic competition with increasing returns to scale assumptions from the new theory. Their model is able to explain traditional inter-industry trade, newer intra-industry trade and the firm's selection effect into export markets. They also claim that theorems such as factor price equalization, the Stolper-Samuelson and the Rybczynski theorem from a traditional HO model continue to apply with slight adjustments[6].

As shown above the gravity equation can be derived from different theoretical specifications. But the gravity equation has also been used to investigate anomalies in trade. A

[6] Stolper-Samuleson theorem says that an increase in the relative price of a good will lead to an increase in the return on the factor used intensively in the production of that good, and a decline in the other good. Rybczynski says with constant relative prices will an increase in factor endowment lead to a more than proportional increase in production of the industry using it intensively. Factor price equalization occurs when the relative price of two identical goods in the same market equalizes due to competition.

subcategory of the new trade theory called new economic geography has emerged related to different anomalies. This theory has been highly influenced by economists such as Krugman (1991) and Markusen and Venables (1995). This theory put much weight on how geography of production and transport costs can explain the pattern of trade. Two examples illustrate this very well. The first example is the fact that small countries like Belgium and the Netherlands export much more than what their economic size and traditional theory predict. The explanation is not found in comparative advantages but their geographical location. Both countries have Europe's biggest ports, Antwerp and Rotterdam which attract significant volumes of goods to be exported from all over Europe since most trade is transported overseas. Their ports are closer to foreign markets and shipping from Antwerp or Rotterdam lowers transport costs in trade. The last example concentrates on trade volume between 2 pairs of cities with equal distance between each other. Statistics show that London and Paris trade far less with each other than Chicago and New York which contradicts predictions by traditional trade theory. What make this example interesting are the different types of trade flow. Between Paris and London there is international trade, but between New York and Chicago there is intra-national trade. The intranational trade avoids trade barriers such as higher transport costs, language cost and borders compared with the international trade. The costs mentioned reduced the volume of trade. It might be obvious that trade between two cities within a country is larger than trade between two cities in separate countries. However, it is interesting to investigate which factors actually cause the decline in trade and create the anomalies in international trade theory. Common to both examples is how location is central in explaining the trade. In the first example location increases trade due to advantages in the industrial structure (ports are close to foreign markets and attract goods to be exported more cheaply). The latter explain how two pairs of trading partners can give different trade volumes due to trade barriers or other trade dampening determinants. The economic geography theory explains trade through location and market structures. And the gravity equation is able to capture these features by the inclusion of distance and economies of scale assumptions.

The last contribution from Anderson and Wincoop (2003) makes use of a complete expenditure system when estimating key parameters in the gravity equation. I adopt their approximation when I next present a theoretical foundation for the gravity equation.

## 2.2 A theoretical gravity model

I follow the structure from Anderson & Van Wincoop (2003) when presenting the theoretical model. The model is presented in without free trade since this introduces trade cost originating from tariffs or transport costs. It is an up-to-date and efficient general framework which includes costs directly and fits my analysis well. Since it aims to solve McCallum's border puzzle the model is set up to explain both intra-national and international trade. The model describes trade between *region*s since it analyzes trade within and between countries, but I will only use one *country* since I am researching trade between Norway and other countries. However, it shows that the gravity equation is a reduced form from a partial equilibrium of a general equilibrium model with products differentiated by place of origin (Anderson (1979), Bergstrand (1985), Anderson and van Wincoop (2003)). The set of assumptions are similar to Bergstrands specification of a general gravity equation mentioned in section 3.1, and the income level of the exporting and importing countries are treated as fixed.

**Consumers' preferences**

A constant elasticity of substitution (CES) function represents the consumers identical and homothetic preferences. This means that as consumer income increases, the consumer's consumption of goods will increase equally which in turn makes the income elasticity of demand for each good constant and equal to one. The consumers in country j maximize

$$(1) \qquad \left( \sum_i \beta_i^{(1-\sigma)/\sigma} c_{ij}^{(\sigma-1)/\sigma} \right)^{\sigma/(\sigma-1)},$$

relative to their budget constraint

$$(2) \qquad \sum_i p_{ij} c_{ij} = y_j .$$

Equation (2) expresses how all income equals the sum of all consumed goods, with no saving possible. The consumption in country j of country i's goods is given by $c_{ij}$. The $\sigma$ is elasticity of substitution between goods. $\beta_i$ is the positive distribution parameter. $y_j$ is nominal income in country j, and $p_{ij}$ is the price consumers from region j have to pay for goods from region i (import prices). All supply prices are set equal to one.

There also exists a price difference between countries. The price difference is due to unobservable trade costs, which are imposed on the exporters. The cost results in lower profits to exporting firms and is important to identify. $p_i$ is the price of goods in country i and $t_{ij}$ is the trade cost from country i to country j. By multiplying these two factors, we obtain real costs of goods from country i to country j. $p_{ij}= t_ij p_i$ as used in equation (2).

(3) $p_{ij}=t_{ij}p_i$

**Demand for region i's goods**

In addition a cost occurs when goods are transported from country i to country j which is passed on to the importer. Exporters face large costs rising from sources such as transport costs, regulatory costs, linguistic problems and acquiring information about the foreign market (Bernard, Redding & Schott 2004). This type of cost can be formulated as a melting iceberg, meaning that a share of the good is lost during shipment, represented by $(t_{ij}-1)/t_{ij}$. The unit cost for each shipped good between i and j can be represented by $(t_{ij} – 1)$. Total costs passed on to the importer are $(t_{ij}-1)p_i c_{ij}$. The value of trade between i and j in nominal terms is $x_{ij}=p_{ij}c_{ij}$. Income of region i is $y_i=\Sigma_j(x_{ij})$. The demand in country j for goods produced in country i is satisfying the maximization of (1) subject to (2) and solves as follows:

(4) $$x_{ij} = \left( \frac{\beta_i p_i t_{ij}}{P_j} \right)^{(1-\sigma)} y_j$$

The denominator in (4), $P_j$, is the consumer price index in country j. It is a function of price in region i, the positive distribution parameter, and the trade cost between i and j

(5) $$P_j = \left[ \sum_i (\beta_i p_i t_{ij})^{(1-\sigma)} \right]^{1/(1-\sigma)}$$

The market clearance condition in (6) is imposed due to the general equilibrium model[7]. This implies that the income in country i must equal the total demand of goods produced in country i from other j countries. The nation's income, $y_i$, solves by using properties of previous functions;

(6) $$y_i = \sum_j x_{ij}$$

Inserting $x_j$ from (4) on the RHS in (6) gives us

(7) $$y_i = \sum_j (\beta_i p_i t_{ij} / P_j)^{(1-\sigma)} y_j$$

Since we are summarizing over all the *j* terms we can separate out the *i* terms and move the *i* terms, $\beta_i p_i$, in front of the $\Sigma$ sign.

[7] This is not a traditional equilibrium model since the production and income are given, but a more narrow trade model.

$$(8) \qquad y_i = (\beta_i p_i)^{(1-\sigma)} \sum_j (t_{ij} / P_j)^{(1-\sigma)} y_j \ , \quad \forall i$$

Anderson and van Wincoop (2004) follow Anderson's (1979) technique to solve $\beta_i$ from the last equation by imposing all prices in origin i to equal one and inserts this into the $x_{ij}$ equation. They define a region's income share of world income as $\theta_j = y_j / y^W$, where world's total income is $y^W \equiv \Sigma_j y_j$. Export between region i and j can therefore be expressed as

$$(9) \qquad x_{ij} = \frac{y_i y_j}{y_w} \left( \frac{t_{ij}}{\Pi_i P_j} \right)^{(1-\sigma)} ,$$

where the denominator $\Pi_i$ represents;

$$(10) \qquad \Pi_i \equiv \left( \sum_j (t_{ij} / P_j)^{(1-\sigma)} \theta_j \right)^{1/(1-\sigma)} .$$

By substituting the equilibrium scaled prices into (7), the price index in the importing country, $P_j$, is specified as

$$(11) \qquad P_j = \left( \sum_i (t_{ij} / \Pi_i)^{(1-\sigma)} \theta_i \right)^{1/(1-\sigma)} .$$

It is now possible to solve Π by combining the two last equations
By assuming symmetric trade barriers, $t_{ij}=t_{ji}$, this provides the model with a helpful simplification, by $\Pi_i=P_i$.

$$(12) \qquad P_j^{1-\sigma} = \sum_i P_i^{\sigma-1} \theta_i t_{ij}^{1-\sigma} \quad \forall j.$$

The price indices are now a function of all bilateral trade barriers, transport costs and income shares between country j and rest of the world. The gravity equation then turns out to solve for

$$(13) \qquad x_{ij} = \frac{y_i y_j}{y^w} \left( \frac{t_{ij}}{P_i P_j} \right)^{(1-\sigma)} .$$

When following the framework developed by Anderson and van Wincoop, a gravity equation will be possible to estimate and make operational since they makes simultaneous use of the market clearance properties to find condition to find $P_j^{(1-\sigma)}$ . This, they argue, simplifies the equation significantly compared to other propositions of the gravity equation (Anderson and Van Wincoop 2003).

A last important factor to include is the unobservable factor $t_{ij}$. Anderson and van Wincoop (2003) assume that $t_{ij}$ is a log linear function of distance $d_{ij}$ and a joint border between two regions, $b_{ij}$, given by

$$(14) \quad t_{ij} = d_{ij} b_{ij}^{\rho} .$$

If $b_{ij}=1$ then both regions are located within the same country. If not located within the same county $b_{ij}$ is one plus the bilateral trade tariff.

With the system of equations presented above, the theory implies a log transformed gravity equation like (13) derived from equation (11) where $t_{ij}$ is substituted from equation (12):

$$(15) \quad \ln x_{ij} = k + \ln y_i + \ln y_j + (1-\sigma)\rho \ln d_{ij} + (1-\sigma)\ln b_{ij} - (1-\sigma)\ln P_i - (1-\sigma)\ln P_j$$

The variables from left to right represent trade export from i to j, a constant k, GDP in region i, GDP region j, distance between i and j, the border measure and lastly the regions' two price indices. A gravity equation's explanatory variables can be grouped into three categories:

i) domestic economic forces in the trading countries, i.e. exporters income
ii) economic forces in the importing country; i.e. importers income and population
iii) forces helping or dampening the export between origin and destination; i.e. membership in mutual regional trade agreements, common language or historical linkages

An important insight from a theoretical gravity equation is which impact changes in the relative trade barriers between countries bring to the traded amount. The traded amount of goods between region i and j depend on the trade barriers between them, but also the trade barriers each region face with other trading partners. An increase of trade barriers between region j and the other regions than i – a rise in the import price on goods from other regions than i -, will improve the terms of trade between region i and j. In other words, export from i to j will increase. The dependency on relative trade barriers gives rise to three implications (Anderson & van Wincoop 2004)[8].

[8] A formal justification of the implications is given in Anderson & van Wincoop from 2003, page 176 - 178

After controlling for the economic size, which can be seen as a proxy for factor endowment in the trading countries, Anderson and van Wincoop draw some implications in the pattern of trade. The first is how trade barriers between two small countries is less affected than the trade between two large economies. The second is that intra-national trade is increased more within small economies than large economies.

With Anderson and van Wincoop (2003) as a baseline in this section, I have presented a gravity equation where trade between two regions depends on the income of both regions, the distance between regions, the regions' respective price level and an existence of a common border between countries. This solves for omitted variable bias and makes it possible to conduct a comparative static due to the inclusion of relevant variables other than just income and distance and solving for a general-equilibrium model. The next sections make use of the theoretical framework just presented, and use it to analyze the Norwegian export between 1999 and 2004.

# 3. Empirical gravity model

I present the empirical gravity model in this section. The empirical gravity model I make use of is similar to equation (13) in section 2.2, but it is extended with relevant variables which are relevant to determine the Norwegian export pattern.

## 3.1 An augmented gravity model

The most simplified gravity equation can be expressed as

(14) $$X_{ij} = AY_i^{\beta_1} Y_j^{\beta_2} dist_{ij}^{\beta_3}$$

Typically, equation (14) is specified as log-linear function and used on either cross-sectional data or panel data. This equation would predict a trade share in a country equal to the country's share of total GDP in the world. When used on panel data, as I will do in this thesis, a time subscript τ is added, except on time invariant terms. Replacing Y with GDP and taking log on both sides provides this equation;

(15) $$\ln X_{ij\tau} = \ln A + \beta_1 \ln GDP_{i\tau} + \beta_2 \ln GDP_{j\tau} + \beta_3 \ln dist_{\mathrm{ij}}.$$

$X_{ij}$ represents the flows of export from country i to country j. To be a valid gravity equation, it must include the countries national income. A country's gross domestic product (GDP) is most often used to measure a nation's income. The $GDP_i$ and $GDP_j$ represents the economical mass of country i or j and *dist* is the distance between the two economies[9]. The GDP's coefficients are expected to be positive and the distance coefficient is predicted to be negative, which is the central implication from the gravity equation, as I will discuss later.

Equation (15) represents a basic gravity equation in an economic environment, but explains trade flows with numerical success. This is the point of departure in most of the papers applying the gravity equation on trade flows. As the gravity equation has been widely used, more determinants have been added. These have been tested and found significant to give impact on aggregate trade flows. Dummy variables are very often included to indicate an existence of previous colonial linkage, common border or common currency. The use of the latter has increased over the last ten years as more countries have joined the European Monetary Union (Rose 2000).

[9] Note the special form of (14) if $\beta_1 = \beta_2 = 1$ and $\beta_3 = -2$. The equation is then identical to Newton's original gravity equation.

I keep the standard variables in my augmented gravity equation, except for the exporters GDP term. This term is excluded as a variable since it would be constant each year. Instead the exporters' national GDP is captured by the constant term, lnA, in the regressions. In addition to these I also include 7 other variables in the augmented gravity equation. My final equation is given by:

$$(16)\quad \ln X_{ij\tau} = \ln A + \beta_1 \ln GDP_{j\tau} + \beta_2 \ln dist_{ij} + \beta_3 \ln POP_{\tau} + \beta_4 SCAND + \beta_5 COMBOR + \beta_6 EFTA + \beta_7 WTO + \beta_8 LDC + \beta_9 EU + \beta_{10} YEAR + \ln \varepsilon_{\tau}\,.$$

I keep the standard gravity variables representing distance between countries as proxy for trade barriers, the importing country's GDP as a proxy for the nation's income and I include the importer's population. I continue to include a dummy representing a common land linked border between Norway and its trading partner. Three different dummies indicate if the trading partners are members of bilateral or multilateral trade and policy agreements of which Norway is also a member. A dummy representing a Scandinavian language is also included and a dummy included indicating if an importing country is defined as a least developed country (LDC) by United Nations.
The variables are defined as

- $X_{ij}$ denotes the value of export from country i to country j in logarithmic values. I.e. from Norway to its trading partners.
- *lnA* is a constant, which in this situation also includes Norway's GDP. I include it since it will act as a constant to all firms, just adjusted from year to year. lnA is replaced by a from now on.
- $GDP_{ij}$ is the gross domestic product of the importing countries, measured in logarithmic terms. Both GDP coefficients are expected to take a positive value. The GDPs are included as a measure of importers' demand of Norwegian goods.
- *dist* is a measure of the distance between country i and j. The coefficient is expected to be negative. Increasing distance decreases trade. The distance parameter seems to capture the fact that intra-industry trade decreases more in relation to inter-industry trade when distance between two trading economies increases.
- *POP* is the population in the importing country j, the coefficient's sign is somewhat undetermined as it can go both ways. I'll return to this in the result and discussion part.

- *SCAND* is a dummy variable set up to indicate the existence of a Scandinavian language in country j. It takes the value "1" if there is a Scandinavian language in the importing country and "0" if not[10]. This is expected to be positive due to decreased trade barriers, i.e. no need to translate agreements, fewer costly misunderstandings etc.
- *COMBOR* is a binary variable for the existence of a common geographical border, expected to be positive due to smaller iceberg costs on the way and lower transport costs from Norway to the importing destination[11].
- *WTO, EU* and *EFTA* are dummy variables indicating if Norway's trading partners are members of the European Union (EU), European Free Trade Agreement (EFTA) or the World Trade Organization (WTO). If a country is a member in any of these it takes the value "1", if not it is given the value"0".
- *LDC* is an abbreviation for least developed countries. The group contains 50 countries which fall within a definition set by the UN. These countries are mostly found in Africa, Asia and Latin America. LDC countries have an agreement with the WTO which says that the countries can import goods from other WTO members without paying taxes and dues. If a country is defined as a LDC it is given the value "1", if not it is defined as "0".
- *YEAR* is a time dummy representing which year the data is collected from, either 1999, 2000, 2001, 2002, 2003 or 2004.
- $ln\varepsilon$ is a log-normally distributed error term with $E(\ln\varepsilon)=0$. I assume the error term to only be correlated with the left hand side variable, $X_{ij}$, implying homoscedasiticity. $\ln\varepsilon$ is from now substituted with $u$ for simplicity.

I drop *i* as subscript on variables since all exports are from Norway to other countries indicated by *j*. *i* could be replaced by N or dropped. I choose the latter for simplicity.

The first three coefficients can be interpreted as how much the aggregated export changes in percent due to a marginal change in first three variables. The EU, WTO and EFTA dummies indicate regional trade agreements and bilateral trade agreements between Norway and other countries. They are treated separately to get a more detailed picture of how the different agreements can impact different margins. The RTA coefficients can be interpreted as how

[10] Other Scandinavian languages, besides Norwegian, are Swedish and Danish.
[11] Norway shares a contingent border with Russia, Finland and Sweden.

much more Norway will export in total if a trading partner joins one of the three RTAs. Helpman, Melitz and Rubinstein (2007) show that countries which share a common mainland border, regional trade agreement, a common language or a close location tend to trade more with each other than other countries.

In section 2 I presented and reviewed a theoretical framework for the gravity theory. I have so far presented an augmented empirical gravity equation on an aggregated level of exports. But the aggregated export equation is just one of three equations I will estimate in this paper. As mentioned in the introduction of this thesis, I estimate three gravity equations. So in addition, to the aggregated equation I also estimate the firm's extensive and intensive margins. These two margins can be decomposed from the aggregated export. The goal of this paper is to determine which of the two margins that is the dominant driving force behind the aggregated trade flows. It is now time to decompose the aggregated gravity equation into the extensive and intensive margin to analyze the driving forces in trade at firm level.

The firms' extensive margin in trade is the numbers of firms exporting and the intensive margin in trade is the average export value per firm. The aggregate trade flows between two economies can be broken down into

(17) $X_{ij} = N_{ij}^{f} \times x_{ij}^{-f}$

The first term on the right hand side, $N_{ij}^{f}$, indicates the number of firms exporting from country i to country j and the latter term, $x_{ij}^{-f}$, is the average value of exports from i to j. By taking the natural logarithm on both sides I obtain

(18) $\ln X_{ij} = \ln N_{ij}^{f} + \ln x_{ij}^{-f}$

Combining this with our final gravity equation discussed earlier, this gives us the relation between the margins and the gravity equation before further investigation of the trade flows.

(19) $$\ln X_{ij}^{f} = \alpha_o + \beta_1 \ln GDP_{j\tau} + \beta_2 \ln dist_{ij} + \beta_3 \ln POP_{\tau} + \beta_4 SCAND + \beta_5 COMBOR + \beta_6 EFTA + \beta_7 WTO + \beta_8 LDC + \beta_9 EU + \beta_{10} YEAR + u_{\tau}$$

When analyzing the extensive and intensive margins in the regression, $N_{ij}^{f}$ and $x_{ij}^{-f}$ will alternate in being the dependent LHS variable.

## 3.2 Data

The Norwegian export data are collected through Statistics Norway (hereafter abbreviated to SSB). They cover Norwegian export firms for the period from 1999 to 2004 and cover 90 % of the manufacturing industry in total per year. A firm is defined as an exporter when its exports exceed 1000 Norwegian kroner (NOK). The data set also specifies how much a firm exports to each importing country. The sample contains sector facts about the firm down to a NACE 5 level and the value of exports to each country the firms' trade with. The exports have been converted from Norwegian kroner to US dollars with average exchange rates for each year.

The distance variables are collected from the website of Centre d'Etudes Prospectives et d'Informations Internationales (CEPII). The distances are termed "simple distances" and are measured in kilometres. The distance is measured between the two trading countries' most populated city. Since Oslo is most populated and also the capital in Norway, it is where the distances are measured from.

Yearly gross domestic product data is collected from the website of The United Nations Statistics Database. The countries' gross domestic product GDP data are measured in current US dollars and estimated by The World Bank[12].

Statistics of population from 1999 to 2004 are downloaded from Penn World Table 6.1 and are measured in absolute numbers. In the data set population is measured in whole numbers.

I define Scandinavian languages as Danish, Swedish and Norwegian. Two more languages could be included, Icelandic and Finnish even though the countries belong to the Nordic and not Scandinavia, but that would not necessary mean that the Norwegian firms could speak Norwegian in those countries, so the including of language as a positive determinant in trade would lose its meaning[13].

Countries Norway shares a land-linked border with are found in any encyclopaedia or atlas. Neighbouring countries overseas are not included in this group. Shipping to overseas countries involves higher transport costs and countries such as Denmark are not included. A key point of diversifying the countries is to get a picture of how distance and geographical location influence trade.[14]

[12] GDP data available online at http://unstats.un.org/unsd/cdb/cdb_quick_data_extract.asp (Visited 2.2.2008)
[13] For a discussion of Scandinavian visit http://en.wiktionary.org/wiki/Scandinavian (Visited 2.2.2008)
[14] I used http://en.wikipedia.org/wiki/Norway as a source. (Visited 2.2.2008)

The list of countries that are members of the different regional trade agreements are gathered from the websites of the respective organizations. Today 28 countries are members of the European Union. During the six years in question the European Union has expanded. In 2004 Estonia, Lithuania, Latvia, Cyprus, Poland, Slovenia, Slovakia, Hungary, Malta and the Czech Republic joined as new members. The European Free Trade Agreement comprises four countries - Iceland, Norway, Switzerland, and Liechtenstein. The member list of the World Trade Organization today consists of 151 countries.[15] Also the WTO enlarged its member list between 1999 and 2004. In 2000 Oman, Jordan, Georgia, Croatia and Albania joined the WTO. In 2001 China and Taiwan joined. Macedonia joined in 2003 and Nepal was the last country to join in 2004. The complete member lists are presented in Appendix A.

The list of countries defined as a LDC is available from the United Nations website and today comprises 50 countries.[16] The group contains countries from Africa, Asia and Latin-America.

Due to lack of data on some countries my final data set contains fewer observations than initially. In 1999 and 2000 152 countries, in 2001 and 2002 153 countries received goods from the manufacturing industry. In 2003 the firms exported to 150 countries, while in 2004 the firms exported to 152 countries. Most of the countries lacking data are countries like Afghanistan, Iraq or other countries with little importance to Norwegian export patterns.

## 3.3 Norwegian export patterns from 1999 - 2004

Before I estimate the three different gravity equations I will take time to present descriptive statistics from the data set during 1999 – 2004. Table 3.1 below shows the number of firms, average export values per firm and the number of countries importing from Norwegian manufacturers. The number of Norwegian manufacturers, the extensive margin represented by $N_{ij}^{f}$ firms in (19), engaged in exports was 3930 in 1999. This number continued to increase in 5 years before it reached its peak in 2003 with just over 4500 firms and declined 1,5% in 2004. From 1999 to 2004 the number of firms increased with 13.1 %. However, the number of countries receiving goods from these firms had nearly no reduction. It varied between 207 countries in 1999 and 201 importing countries in 2003.

Despite the small reduction in importing countries the average export value per firm increased continuously over the last 6 years. This illustrates that there has been an increase in both the

[15] Online member lists found at: WTO - http://www.wto.org, EFTA member list; http://www.efta.int EU - http://europa.eu/abc/european_countries/index_en.htm (Visited 2.2.2008)

[16] UN's LDC website online; http://www.unohrlls.org/en/ldc/related/62/ (Visited 2.2.2008)

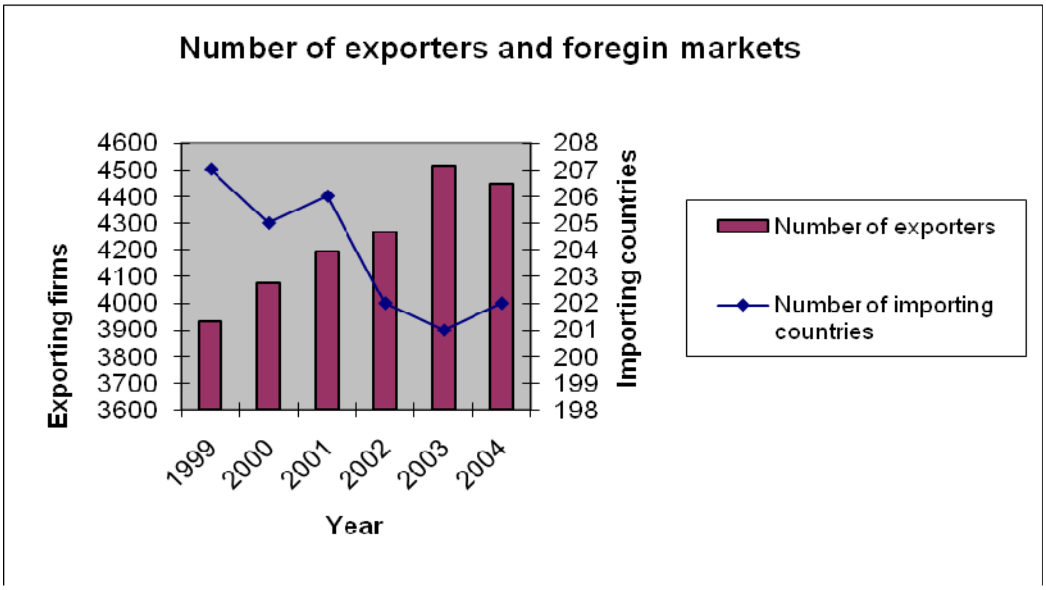
extensive and intensive margin the last 6 years which naturally imply a higher level of aggregate exports.

**Table 3.1 Statistics from the manufacturing industry 1999 – 2004 in Norway.**

| Year | Number of exporters | Importing countries | Average export value per firm |
|---|---|---|---|
| 1999 | 3930 | 207 | 9223494 |
| 2000 | 4074 | 205 | 9485908 |
| 2001 | 4193 | 206 | 9929833 |
| 2002 | 4267 | 202 | 10400000 |
| 2003 | 4516 | 201 | 11000000 |
| 2004 | 4445 | 202 | 11200000 |

Graph 3.1 plots the development in exporters (the left y-axis) and the number of importing countries (right y-axis) in the six years of interest. Out of curiosity I estimated the correlation between them. The period is too short to draw any valid conclusions, but the correlation between number of exporters and the number of importing countries is -0,9.

**Graph 3.1 Development in exporting firms and importing countries**




The reason to look more closely into the intensive and extensive margins is the different magnitudes in change when the right hand side variables are changed. Several papers shows that the effect on aggregated trade flows from changes in the firms' margins are dominated by the numbers of exporters rather than the amount each firm exports on average (e.g. Chaney (2004); Bernard et.al (2007). But how does this work out in Norway? Does an increase in my variables increase the numbers of exporters or do some of them actually make it worse to be an exporter? Is it so that an increase between Norway and its trading partners limits the number of exporters or will it do the opposite? Or on the other hand, if distance increases, will the average export value shipped from Norway increase or will it limit the value shipped from Norway so that the average export value per firm will actually decline? These and other questions I will try to answer in the next two sections. There are several papers showing that the magnitude of change in the firm's extensive margin is larger on the aggregated trade level than a change in the firm's intensive margin.

## 3.4 Results

In section 3.1 I reviewed an empirical gravity equation and developed a gravity equation suitable to analyze Norwegian exports. Results from estimates of the aggregated firms' extensive and intensive margins are presented in two different versions. The first version of the gravity equation contains only the importers GDP and distance which are the original variables in the gravity equation (equations M1 – M3). The last version is extended with all variables (equations M4 – M6).

(M1) $\ln X_{ij} = \alpha_o + \beta_1 \ln GDP_{j\tau} + \beta_2 \ln dist_{ij} + \beta_3 \ln POP_{\tau}$

(M2) $\ln N^{f}_{ij} = \alpha_o + \beta_1 \ln GDP_{j\tau} + \beta_2 \ln dist_{ij} + \beta_3 \ln POP_{\tau}$

(M3) $\ln x^{-f}_{ij} = \alpha_o + \beta_1 \ln GDP_{j\tau} + \beta_2 \ln dist_{ij} + \beta_3 \ln POP_{\tau}$

These first three equations are the simplest version. As mentioned earlier, exporter's GDP is normally included in gravity equations, but since I am only analyzing the exports from Norway to other countries Norway's GDP will be treated as a yearly constant. Thus, the GDP term is defined in the equation and as the constant term, $\alpha_0$.

The next three equations make up the augmented gravity equation. Importers' GDP, population and distance are included as in M1-M3. In addition, these equations have the factors common contingent border, Scandinavian language, membership in RTAs and if the country is defined as a LDC. M4 is the aggregate exports equation. M5 represents the extensive margin equation. M6 represents the intensive margin equation.

(M4) $\ln X_{ij}^{f} = \alpha_o + \beta_1 \ln GDP_{j\tau} + \beta_2 \ln dist_{ij} + \beta_3 \ln POP_{\tau} + \beta_4 SCAND + \beta_5 COMBOR + \beta_6 EFTA + \beta_7 WTO + \beta_8 LDC + \beta_9 EU + \beta_{10} YEAR + u_{\tau}$

(M5) $\ln N_{ij}^{f} = \alpha_o + \beta_1 \ln GDP_{j\tau} + \beta_2 \ln dist_{ij} + \beta_3 \ln POP_{\tau} + \beta_4 SCAND + \beta_5 COMBOR + \beta_6 EFTA + \beta_7 WTO + \beta_8 LDC + \beta_9 EU + \beta_{10} YEAR + u_{\tau}$

(M6) $\ln x_{ij}^{-f} = \alpha_o + \beta_1 \ln GDP_{j\tau} + \beta_2 \ln dist_{ij} + \beta_3 \ln POP_{\tau} + \beta_4 SCAND + \beta_5 COMBOR + \beta_6 EFTA + \beta_7 WTO + \beta_8 LDC + \beta_9 EU + \beta_{10} YEAR + u$

The following table show estimated values of ordinary least squares regressions on M1 – M6. Yearly dummies are included to capture variation in values from year to year. The aim of comparing M1 - 3 with M4 - 6 is to analyze any potential change in the estimated coefficients when more explanatory variables are added and check the development of the reported $R^2$ and the $R^2$.

**TABLE 3.2 Regression result when estimating M1 – M6**

| | M1 | M2 | M3 | M4 | M5 | M6 |
|---|---|---|---|---|---|---|
| | $X_{ij}$ | $N_{ij}^{f}$ | $x_{ij}^{-f}$ | $X_{ij}$ | $N_{ij}^{f}$ | $x_{ij}^{-f}$ |
| **ln_n_gdp** | 1.306*** | .807*** | .499*** | 1.250*** | .747*** | .503*** |
| | (.029) | (.016) | (.022) | (.040) | (.021) | (.030) |
| **lndist** | -.739*** | -.480*** | -.259*** | -.661*** | -.432*** | -.228*** |
| | (.059) | (.032) | (.044) | (.076) | (.040) | (.057) |
| **lnpop** | -.421*** | -.286*** | -.134*** | -.373*** | -.230*** | -.143*** |
| | (.033) | (.018) | (.025) | (.040) | (.021) | (.030) |
| **efta** | - | - | - | .632 | 1.174*** | -.542 |
| | | | | (.410) | (.218) | (.311) |
| **SCAND** | - | - | - | .717 | .854*** | -.136 |
| | | | | .(431) | (.229) | (.327) |
| **COMBOR** | - | - | - | .535 | .456*** | .078 |
| | | | | (.347) | (.184) | (.263) |
| **wto** | - | - | - | .268* | .402*** | -.133 |
| | | | | (.123) | (.065) | (.093) |
| **eu** | - | - | - | .180 | .050 | .129 |
| | | | | (.185) | (.098) | (.140) |
| **ldc** | - | - | - | -.0592 | -.014 | -.044 |
| | | | | (.135) | (.071) | (.102) |
| **_laar_2000** | -4.296*** | .0481 | -4.344*** | -4.307*** | .031 | -4.338*** |
| | (.149) *** | (.083) | (.113) | (.148) | (.079) | (.113) |
| **_laar_2001** | -4.165*** | .101 | -4.266*** | -4.179*** | .086 | -4.261*** |
| | (.149) | (.082) | (.113) | (.148) | (.079) | (.112) |
| **_laar_2002** | -4.151*** | .0395 | -4.191*** | -4.158*** | .0277 | -4.186*** |
| | (.149) | (.082) | (.113) | (.148) | (.079) | (.130) |
| **_laar_2003** | -4.169*** | -.0001 | -4.169*** | -4.177*** | -.0175 | -4.159*** |
| | (.150) | (.083) | (.113) | (.150) | (.079) | (.114) |
| **_laar_2004** | -4.176*** | -.0674 | -4.109*** | -4.189*** | -.080 | -4.108*** |
| | (.149) | (.083) | (.113) | (.149) | (.079) | (.114) |
| **_cons** | 8.667*** | -6.758*** | 15.426*** | 8.315*** | -6.994*** | 15.311*** |
| | (.786) | (.436) | (.594) | (.861) | (.458) | (.653) |
| R^2 | 0.85 | 0.85 | 0.79 | 0.86 | 0.86 | 0.79 |
| Adj. R^2 | 0.85 | 0.84 | 0.78 | 0.85 | 0.85 | 0.78 |
| N | 918 | 918 | 918 | 918 | 918 | 918 |

*Source; Stata, each regression out prints is presented in appendix B.*
*Note; Numbers in parenthesis is the coefficients' standard errors. *, ** and ***, respectively indicate levels of significance at 10 %, 5 % and 1 %.*

All equations are log-log specified which means that the explanatory variables coefficients can be interpreted as percentage changes in the LHS variable when RHS variable is increased marginally (Hill, Griffiths and Judge 2001). This is also consistent with the constant elasticity of substitution assumption from section 2 (ibid). Hence, the coefficient predicts how much the dependent variable is changed in percentage by a marginal – one unit - increase in one of the explanatory variables while the other right hand side variables are kept fixed. Shown in previous subsections, the Cobb-Douglas type of function which specifies the aggregated exports allows us to decompose the aggregated export ($lnX_{ij}$) equation into the extensive ($lnN_{ij}^{f}$ ) and intensive ($x_{ij}^{-f}$) margins of exports. Adding together a specific variable's coefficients from the extensive (M2 or M5) and the intensive (M3 or M6) margin will equal the coefficient on aggregated exports (M1 or M4).

As expected and in line with previous results from other countries, the result from a change in the right hand side variables on the aggregated export flow comes from the firms' extensive margins and not the intensive margin papers (Bernard et al. 2007; Mayer and Ottaviano 2007). This implies that what really matters as the driving force of these two margins is the number of firms. The effect on M1 by a change in the intensive margin is outweighed by the effect on the extensive margin by the change in the same variable in every case except for the *EU* and *LDC* dummies in the augmented gravity equations. I will return to this in section 3.4.

A high standard error can cause the estimate to be statistically insignificant. I have investigated the reported p-values in the regression out print, column termed "P>t", to check if any estimates can be stated as not statistically significant. In M1 - M6 the three first coefficients are statistically significant. In M4 and M6 the dummies are reported with p-values higher than a 0.025 level of significance. This means that the t-test failed to reject the hypothesis of the dummies being insignificant. In M5 however, only the EU and LDC coefficients are reported to be statistically insignificant. There may be some kind of economical relationship between the left-hand side variable and the explanatory variables, even though some of these variables are tested to be statistically insignificant.

I have focused on the explanatory validity and consistency of my estimates separately and commented possibility of mutual dependency between the explanatory variables up until now. There is however, a way to test the overall significance of a model. This is to test if the model depends on one particular or several variables. The test used is the F-test. In the regression output the F-statistic is reported to the top right. Since we have two versions of the different equations we have to deal with two different F statistics.

In M1 – M3 it is $F_{(8,909)}$ and in M4 - M6 it is $F_{(14, 903)}$. $F_{(8,909)}$ tests $H_0$: $\beta_i=0$, i=1-6 against $H_1$: at least one or more $\beta_i \neq 0$. With a significance level of 0.05 and $F_c$= 1.94 resulting from the $F_{(8,909)}$ distribution. The decision rule to reject $H_0$ is if $F_c<F_{(8,909)}$. In M1 – M3 is $F_c<F_{(8,909)}$, and $H_o$ is rejected in M1, M2 and M3. The same conclusion can be drawn in M4 –M6 where $F_{(14, 903)}$ tests $H_0$: $\beta_i=0$, i=1-14, against $H_1$: at least one or more $\beta_i \neq 0$. In this augmented version is $F_{(14, 903)} \approx 1.69$ and we can still reject the $H_0$.

To find out the status of colinearity of the right-hand side variables I estimate a correlation matrix presented as table 3.3 below.

**Table 3.3 Correlation matrix of explanatory variables**

| | ln_n_gdp | lndist | lnpop | COMBOR | SCAND | EFTA | EU | WTO | LDC |
|---|---|---|---|---|---|---|---|---|---|
| **ln_n_gdp** | 1.000 | | | | | | | | |
| **lndist** | -0.324 | 1.000 | | | | | | | |
| **lnpop** | 0.724 | -0.070 | 1.000 | | | | | | |
| **COMBOR** | 0.164 | -0.312 | 0.072 | 1.000 | | | | | |
| **SCAND** | 0.131 | -0.340 | -0.001 | 0.398 | 1.000 | | | | |
| **EFTA** | 0.090 | -0.157 | -0.095 | -0.016 | -0.013 | 1.000 | | | |
| **EU** | 0.411 | -0.621 | 0.111 | 0.227 | 0.299 | -0.044 | 1.000 | | |
| **WTO** | 0.186 | 0.069 | 0.010 | -0.065 | 0.052 | 0.052 | 0.122 | 1.000 | |
| **LDC** | -0.429 | 0.212 | -0.029 | -0.072 | -0.059 | -0.059 | -0.197 | -0.225 | 1.000 |

*Source*; Stata

The correlations reported in table 3.3 are of normal values in most cases. Hill et al. (2001) uses correlations from 0.8 and up as strong indicators of multicolinearity. The only correlation close to this is the strong positive correlation between GDP and population. The correlations worth commenting relates to geographical features. *Ln_n_GDP* is reported with a strong positive relationship with *EU* and *lnpop* population and EU and a negative relationship with distance. The positive relationship between EU and GDP is explained mainly by looking at the EU members' GDP levels. Their level of GDP can be treated as relatively high compared to the rest of the world. GDP and population have a positive relationship from the fact that the higher population a country has, the bigger the chance is of it having a high GDP given stable political and economical conditions. It must also be added to the explanation that the distance is measured from Norway. And there is a stylized fact that Norway is located together with other OECD countries which also have a high GDP, in the western part of Europe. The poorer countries are often found more distant to Norway in Africa, Latin America and some parts of Asia causing the negative relationship between distance and GDP

*Lndist* is negative correlated with COMBOR, *SCAND* and *EU*. This is again explained by the measuring of distance between Norway and its trading countries. The countries who speak a Scandinavian language are our closest neighbours and imply low distance values. The possibility of speaking a Scandinavian language decreases with distance. EU-members are located close to Norway but not outside Europe, giving a negative relationship between distance and EU-membership.

*COMBOR* and *SCAND* have a positive relationship measured to be 0.40. This strong relationship is not surprising considering Sweden's common border and Scandinavian language. The positive relation between speaking Scandinavian and being a member of the European Union is the last relationship worth commenting. Sweden and Denmark, the two other countries speaking Scandinavian, are both EU members and create the positive relationship.

To check out how much of the variation in exports around its mean the six equations explain I look at the reported $R^2$s and adjusted $R^2$s in the previous table. The $R^2$s are reported to be around 0.80 in the first three equations and are only slightly increased in value compared with the three augmented equations. The same pattern applies to adjusted $R^2$s. A high $R^2$ may be a result of explanatory variables moving together, as phenomenon often referred to as multicolinearity. The consequences can be large standard errors in the estimates and not possible to define OLS estimators if exact collinearity. Checking all the standard errors I find two coefficients in M5 which have standard errors larger than their estimated value. The most serious in equation M5 is the LDC coefficient where the large standard error can give it a positive effect rather than a negative. Of the estimated coefficients in M5 there are two estimates which are statistically insignificant. In M6 *SCAND*, *COMBOR*, *EU* and *LDC* have standard errors larger than their estimated coefficients. As a result of the imprecise estimates, *EU* can have a positive and a negative effect and *SCAND* can have a positive effect on the intensive margin, which is the opposite sign reported in M6 in table 3.2. In M4 *EU* and *LDC* have standard errors larger than the values of their estimated coefficients.

There can also be observed some kind of pattern in the coefficients values in the extensive and intensive margin. The estimated coefficients have the same signs in M2 and M3. This implies an amplifying force on the aggregated export level from a change in one of the gravity variables along the extensive margin. This tendency is also observed with *ln_n_GDP*, *lndist* and *lnpop* in equation M4 - M6. Table 3.2 reports that a change in GDP, distance, population and controlling for WTO or LDC gives the same effect in both M5 and M6. However, controlling for Scandinavian language, common border and EFTA-membership results in opposite signs in M5 and M6. The three dummies give a positive effect on aggregated exports through the extensive margin and a negative effect through the intensive margin. Hence, the positive impact on total exports through the extensive margin is reduced by the negative impact from reduction in the intensive margin.

Since some of the coefficients were found insignificant, I tested the equations' overall significance with F-tests. The F-tests showed that all equations were significant.

## 3.5 Discussion

In this section I will discuss the different estimated coefficients in the light of expected values and compare my estimates with previous papers. I treat each variable separately and discuss its estimated coefficients in light of theory and seek to explain how it can affect aggregate exports through the intensive and extensive margins.

A general tendency can be observed by looking at the three first estimates in M1 – M6. 7 of 9 coefficients are reduced in absolute value when the equations include dummy variables as well. Only the GDP and population coefficients in M3 experience an increase in absolute value when it is enlarged to M6.

**Importers gross domestic product**

Importers' GDP are expected to have a positive effect on all three dependent variables. Table 3.2 confirms this in my data set as well with positive values in all six estimations. On the aggregated level the estimated coefficients are 1.30 in M1 and 1.25 in M4. The results support other regression results where the importers' GDP coefficient is positive and varies around unity when the gravity equation is applied on trade data from OECD countries. Feenstra (2004) reviews different regressions done on gravity equations by McCallum (1995) and Anderson and van Wincoop (2004). They apply the gravity theory on aggregate trade data between the U.S. and Canada in order to solve the border-effect puzzle presented by McCallum (1995). They estimate the GDP coefficient to vary around unity depending on estimating methods. Mayer and Ottaviano (2007) estimate different gravity equations on exports and foreign direct investments data from countries such as Norway, Belgium and Italy. As done in this thesis, they also decompose the aggregate exports into the intensive and extensive margin. The estimated aggregate exports coefficient is 0.93 in the simple version consisting of only distance, exporters' GDP and importers' GDP. And estimated to 0.96 in the extended version where language, colonial linkage and different RTAs are included. In the same paper the firms' extensive margin is reported to dominate the effect on the aggregated export level more than the intensive margin. This is also the conclusion from this table 3.2. The extensive margin accounts for about 60 % of the increase in aggregate exports when importers GDP increase the margin.

The gravity equation applied in export data from OECD countries gives unitary values as just shown. However, there are papers contradicting the expected positive effect on aggregate exports by a change in importer's GDP. A special case worth mentioning is gravity theory applied on trade data from non-OECD countries. Feenstra (2004) shows that the estimated values of importer's GDP coefficient are negative, which contradicts the gravity equation. The explanation Feenstra finds in the gravity equation's specification. . The explanation finds Feenstra in the gravity equation's specification. The gravity equation is often specified with specialization in differentiated goods. This assumption fits the OECD counties better than developing countries. Developing countries often trade in basic agricultural goods or commodities with low-skilled labour. This implies a much lower degree of differentiation in production and more homogenous trade flows. Hence, the gravity theory fails to explain trade lows in developing (Feenstra 2004).

As mentioned earlier, the gravity function is a type of reduced expenditure function where the importers GDP is entered as a proxy for a country's income and is said to be the importing potential of country j. This is the same as saying that the GDP coefficients in M1 and M4 represents country j's marginal propensity to consume imported goods. I.e. how much of the unit increase in GDP that would be used on importing goods. Table 3.2 reports that an increase in importers GDP will have a positive effect on the number of exporters but also on the average amount exported. Higher income results in higher demand on foreign goods. As the demand on foreign goods increases more firms find it profitable to begin exporting and the number of exporters rises. Low productivity firms are given an incentive to start exporting since the threshold value of making positive profit is reduced. In the end the number of firms increases. The positive effect on the intensive margin could be explained by looking at the increased purchasing power an increase in income brings. Increased purchasing power will also increase the demand on foreign goods which in turn will affect exports from Norway in a positive direction. The positive effect from a marginal increase in GDP on aggregate exports is also strengthened through the average exports shipped per firm. Not only has the number of exporters increased - the increase in income also increases the ability to purchase a larger variety of products which lets each firm export more volume as well. One of the assumptions from the new trade theory is that consumers have preferences for a variety of goods. The assumption ensures all goods which are produced in the exporting industry are actually exported. If each firm produces a unique variety of a product then the increase in number of firms equals a similar increase in the number of varieties exported as well. In addition the increase on the intensive margin can support the hypothesis of increasing returns to scale.

**Distance between exporter and importers**

In table 3.2, increased distance gives an expected negative effect on all three dependent variables. It is estimated to be -0.73 in M1 and -0.66 in M4. These results are normal when compared to other estimations. Mayer and Ottaviano (2007) estimate it to be -0.86 in their simple and extended versions of the gravity equation. The reported coefficients in Feenstra are somewhat higher than reported by table 3.2 and others. The distance coefficients are reported with values ranging from -0.79 to -1.42. As mentioned in 3.3, distance affects the aggregate trade mainly through the extensive margin in both the simple and the augmented version. In table 3.2 is the decrease in aggregate export by a change in distance comes 6 5% from the reduction in number of exporters. This is also the fact in Mayer and Ottaviano's (2007) regression analysis. The extensive margin accounts for 75% of the effect on aggregate exports by increased distance in their regression report.

The distance causes a higher variable cost in trade. Higher transport costs demand higher productivity in order to generate the same profit as for the same products exported to closer markets. As a result, firms are demanded to increase efficiency in production to earn positive profit. The least productive exporters quit exporting and switch to producing solely for the domestic market rather than go out of business. My results are in line with the expected negative value. The magnitudes on the aggregated level of exports are dominated by the effect from the extensive margin rather than the intensive margin in both simple and extended model. The coefficient measures how much exports are reduced in percent on aggregated exports, the extensive margin or the intensive margin if distance is increased marginally. An interpretation of the extensive margin dominance is that distance also captures fixed costs. The decrease in average exported value per firm relates to what I called a melting iceberg cost in section 2.2. When distance increases the transport costs rise. The longer the transport, the higher costs an exporter has to pay in order to get the products to the final market. This will reduce the value of each shipment, especially if the price of transport depends on quantity or weight rather than quality. Hence, the intensive margin can be negatively affected by increased distance.

**Population**

In table 3.2 increased populations are estimated to have a negative effect on aggregate exports. The reduction is mostly driven by the decrease in the extensive margin when population increases. The extensive margin accounts for 61 % and the reduction in average exports per firm is responsible for the rest. The interpretation of population coefficient is straightforward. It measures the marginal change in one of the three dependent variables resulting from a marginal increase in importers population. However, as already mentioned, the sign is somewhat ambiguous. Population is a type of factor endowment which from traditional trade theory can give rise to comparative advantages. If a country has a big population this can imply a large labour force at hand. A big labour force brings down the wages and lowers the price on labour as input. If this is the case, the country can be thought of as being self-sufficient with products produced in labour-intensive industries. Thus, the estimated coefficient of population is expected to take negative value, as in table 3.2. On the other hand, an isolated increase in population implies higher purchasing power in the country. Given that the increased population is given the same income level the estimate coefficients can be expected to be positive. More people buy more goods. This can also imply higher demand on goods produced abroad. The manufacturing industry is labour-intensive. In my data set increased population is estimated to have a negative effect on exports in the manufacturing industry through both reduced number of exporters and through reduced average exported value. As mentioned earlier, increased population may ultimately reduce the cost where labour is used in production. A result is lower priced domestic goods and the amount sold rises which again results in a decreased demand on imported goods.

**Common Scandinavian language**

Helpman, Melitz and Rubinstein (2007) show that sharing a common language increase the trade between two countries. Norwegian is a distinct language, but it is similar to and can be understood by other Scandinavians. The *SCAND* dummy in M4 – M6 captures this fact. Table 3.2 shows that sharing a Scandinavian language gives a positive effect on aggregate exports. This effect is driven entirely by the increased number of exporters, since the effect on average export per firm is negative.

A common language or the possibility of communicating with trading partners through a common language results in a reduction of fixed costs. It simplifies reading law documents, avoids translation of business documents and reduces the costs of establishing a business in the importing country. This is captured by the high estimated value of the extensive margin's coefficient, rather than the intensive margin. The large increases in the number of exporting firms suggest that having a common language is of major importance when deciding where to export. An explanation of why the *SCAND* coefficient in M6 is negative can be explained by how a common language does not reduce variable costs such as transport. As mentioned in 3.2, it is also statistically insignificant.

**Common contingent border**

The variable transport cost can be captured by the contingent border dummy, *COMBOR*. The estimated coefficients in M4 - M6 are all positive, but the effect through the extensive margin on aggregated exports is largest, though amplified by the intensive margin. The coefficients are estimated to be 0.53 in M3, 0.45 in M5 and 0.07 in M6. The coefficients could also be negative, depending on which country the paper takes as point of departure. Bernard et al (2007) explains this by historical arguments. If countries which share a contingent border have recently been at war or had some kind of conflict this is likely to affect trade negatively. A conflict can make firms hesitant with regards to export and the importers' inhabitants could possibly boycott the imported goods. The situation illustrated by Bernard et al. is not representative for the harmonic relationship between Norway and Sweden. In addition Sweden is the country with by far the longest contingent border and the country receiving the largest share of exports in my data set. Russia and Finland have only a minor contingent border compared with Sweden. The large dominance on total exports from the number of exporters again confirms the pattern found in other papers.

**Regional Trade Agreements (RTAs)**

The interpretation of the EU, WTO and ETA coefficients is the percentage change in the dependent variable resulting from the importer being a member of one or more of the RTA's. The effects of being a member of one of the three RTA's are all reported to be positive on the aggregated level of exports (M4). Helpman, Melitz and Rubinstein (2007) also show that countries sharing a common regional trade agreement, like EU, EFTA or WTO, tend to trade more with each other. Their paper differs slightly in how they treat RTA as a controlling variable in the gravity equation. I control for EU, WTO and EFTA separately, but they have one free trade agreement dummy and one WTO dummy instead. However, they estimate the WTO and FTA coefficients to be positive.

The coefficients in table 3.2 range from 0.18 to 0.63 in M4. This seems natural since RTA's often are created with the purpose to dampen trade frictions and lower trade costs by the liberalization of trade regimes. The effect on $lnX_{ij}$ in M4 is mainly through the increased number of internationalized firms, except when it comes to being a member of the EU. The estimated EU coefficient in M5 and M6 are both positive, but the intensive margin accounts for 72 % on the effect on aggregate exports if the importer is a EU- member. This contradicts the traditional pattern where change in the extensive margin is expected to dominate the change in the intensive margin. As commented in section 3.2, some of the variables are reported with high standard errors and reported insignificant in the equations. EU was one of the variables which were reported to have no statistical significance on the intensive margin and it is also insignificant in M5.

The lowering of cost also reduces the internationalized firms' threshold level of earning positive profit. The increased possibility to earn positive profits attracts more firms and the number of exporters increases. The negative effect on average export per firm may be a result of the international competition they face with firms from other member countries which on average are cheaper or for some other reason more attractive than the Norwegian exported goods. Head (2004) reports normal values of an RTA coefficient around 0.5 on the aggregated exports, M4 in table 3.2. This is in line with the estimated values of EFTA and WTO coefficients. The regression finds an export premium of 63 % in aggregate exports if a firm exports to EFTA-members, a mildest increase of 26 % in total if the country is a WTO member and 18 % increased exports if it is to a WTO-member. During the 6 year analysis period only a few countries joined the WTO. Countries of interest are Croatia, joining in 2000, China in 2001 and the former republic of Yugoslavia, in 2003. There was no change in the EFTA member list, while EU gained 10 new countries in 2004. The complete list of members is presented in Appendix A.

**Least developed countries**

If an importing country is defined as a LDC the coefficient says how much effect this will have on each of the three explanatory variables. From table 3.2 this implies a negative impact on the aggregated exports in the manufacturing industry. The impact is driven through both the extensive and intensive channels of trade. As in the EU variable, the intensive margin accounts for most of the impact on aggregate exports. The number of exporters is reported to decline with 1.4 % and the average amount exported is reported to be reduced with 4.5 %. Again, this contradicts expected results in gravity theory as in the case with EU in M5 and M6. As with EU in M6 and M5, the coefficient is reported to be insignificant by a t-test with 5 % significance level.

The reduction in intensive margin can be explained by looking at the characteristics of a LDC country. LDC's are often characterized as poor countries with low income levels, unstable national economies and lack of stable infrastructure in the industries and public services. The correlation matrix reports a negative relationship between being a LDC and GDP and positive correlation with LDC and distance. Looking at where the LDC countries are located relative to Norway it is clear that LDC's are quite remote located. Low levels of income imply lower demand on foreign goods. Shipping to distant markets increases costs in transportation. Low levels of income also reduce the LDC inhabitants' possibility to consume a larger variety of goods which results in lower priority on manufactured products which the data set covers. It might be no surprise that the number of exporters and average export value will decline seeing these factors in relation to each other. It creates larger costs in production, lower demand on foreign goods and a higher threshold level of earning positive profits which reduces the number of firms.

I have now analyzed the regression report on the six equations, M1 – M6 and tried to find an explanation as to why the margin's coefficients either take on positive or negative values by a change in the explanatory variables. The effect on aggregate exports is mainly influenced by the extensive margin by a change in importers' GDP, importers' population, distance or controlling for, EFTA, WTO, language and border. What contradicts expected results in gravity theory is how the effect from the intensive margin dominates the effect from the extensive margin by controlling by LDC and EU-membership. Testing if EU and LDC are statistical significance in M6 shows that they are statistically insignificant with a 5 % level of significance

# 4. Concluding remarks

The point of departure of this thesis was to examine the interplay between the number of exporters and the average exported value per firm by applying the gravity theory on export data from the manufacturing industry from 1999 to 2004. Mayer and Ottaviano (2007) and Bernard et al. (2007) illustrate the difference between domestic and internationalized firms. The firms differ in size, performance and return on inputs. The gravity theory has been highly respected for its ability to explain empirical facts in trade. The gravity equation has been shown to apply equally well for both the new and the old trade theory (Feenstra, Markusen and Rose (1999); Evenett and Keller (2002). It also serves remarkably well when internationalized firms in trade are studied and not only aggregated trade flows.

I have found a significant dominance by the extensive margin from a change in importers' GDP, distance and importers' GDP in both the simple and the augmented version. That is, the effect on numbers of exporters from a change in the mentioned RHS variables is larger on the aggregate exports than the effect on average exported value per firm. The extensive margin is also found to dominate the intensive margin when controlling for language, contingent border, EFTA, WTO. The findings strongly support the fact that the number of firms are of importance in international trade. What contradicts the traditional gravity analyzes of extensive and intensive margins in trade are how the effect from EU-membership or being defined as a LDC. This is a somewhat surprising result. Both coefficients have the largest effect on aggregate exports through the intensive margin. But by testing the coefficients with a 5 % level of significance, I find both to be insignificant in M6. An almost similar regression is done by Mayer and Ottaviano (2007) on French and Belgian trade data. They end up concluding that the effect on aggregate exports is mostly driven by effects from the number of exporters.

From the estimation it can be concluded that the number of exports in the manufacturing industry seem to increase in importer's GDP, presence of Scandinavian language and common border, membership in WTO, EU or EFTA. The firm's extensive margin is decreasing in importers' population when the importer is defined as a LDC and in the distance between Norway and the importing country. The average exported value per firm has the

same pattern, but is reported to be decreasing if the importer is a member of EU, WTO or EFTA.
However, the dominance of extensive margin in aggregated exports is statistically significant only through the GDP, population and distance variables.

As Bernard et al. (2007) claim, my findings support the *newer* theories of trade. The large influence on aggregate exports by number of exporters supports the hypothesis of heterogeneous firms in trade. When distance increases, only the most efficient firms find it profitable to export since the distance implies higher transport costs. On the other hand, if importers' GDP increases, new exporters enter the market. The new trade theory assumes love of variety which implies that adjustments in export happen on the intensive margin if distance increases. I find no support for this hypothesis, except when controlling for EU and LDC. But then again, their coefficients are found statistically insignificant.

The literature regarding internationalized firms is still young. Further investigation is necessary but also exciting. The information about how the firms' intensive and extensive margins impact aggregate exports draws a more detailed picture of the relationship between determinants in trade. Further investigations of manufacturing industry could be to decompose the aggregated exports in the manufacturing industry in a more sophisticated way. For example, it could be possible to divide into the number of exporting firms, number of exporting product per firm and average value per exported product per firm. The decomposing of number of products could be defined from the NACE system (a European industry standard classification system). Bernard et al. (2007) and Mayer and Ottaviano (2007) have done this on trade data from European countries. They argue with strong evidence that the extensive margins – both number of exporters and number of exported products – play a significant role when the aggregated level of exports are determined.

# References


- Anderson, J. E. (1979): "A theoretical foundation for the gravity theory", The American Economic Review, vol 69, No. 1, pp. 106 – 116.
- Anderson, J.E. and E.van Wincoop (2001) "Borders, Trade and Welfare", NBER Working Paper No. 8515, pp. 1 – 33.
- Anderson, J.E. and E.van Wincoop (2003): "Gravity with gravitas: A solution to the Border Puzzle", The American Economic Review, Vol93, No.1, pp. 170 – 192.
- Dixit, A. and V. Norman (2002): "Theory of International Trade" - Oxford University Press, ch 9
- Baldwin, R. and D. Taglioni (2006): "Gravity for dummies and dummies for gravity equations", CEPR Discussion Paper No. 5850, pp. 1 – 24.
- Bergstrand, J. H. (1985): "The gravity equation in International Trade: Some Microeconomic Foundations And Empirical Evidence", The Review of Economics and Statistics, Vol. 67, No.3, pp 474-481
- Bernard, A.B., S.J. Redding and P.K. Schott (2004); "Comparative Advantage and Heterogeneous Firms", NBER Working Papers No. 10668, pp. 2 – 37.
- Bernard, A.B., J.Jensen, S. J. Redding, and P. K. Schott (2007): "Firms in International Trade". NBER Working Paper No. W13054, pp. 1- 29.
- Chaney, T. (2004): "Distorted Gravity: Heterogeneous Firms, Market Structure and the Geography of international trade", http://www.crest.fr/seminaires/lmi/chaney.pdf, downloaded 1.2.2008.
- Evenett, S. J., W. Keller (2002); "On theories explaining the success of The Gravity Equation", Journal of Political Economy, Vol. 110, No 2, pp. 281 – 297.
- Feenstra, R. (2004); "Advanced International Trade – Theory and Evidence", Princeton University Press
- Feenstra, R., J.A. Markusen and A.K Rose (1999): "Using the Gravity Equation To Differentiate Among Alternative Theories Of Trade", downloaded 1.2.2008, http://haas.berkeley.edu/~aroseS, pp. 1 - 28
- Head, K (2003): "Gravity for beginners", downloaded 1.2.200: http://economics.ca/keith/gravity.pdf, pp. 1 – 11.

- Hill, R.C., W.E. Griffiths, G.G. Judge (2001): "Undergraduate Econometrics - 2nd edition", John Wiley & Son
- Helpman E., M.J. Melitz and Y. Rubinstein (2007): "Estimating trade flows: Trading partners and trading volumes", NBER Working Paper No. 12927, pp. 1- 43.
- Krugman, P. R. (1980); "Scale Economics, Product Differentiation and The pattern of Trade", American Economic Review, Vol. 70, No. 5, pp 950 – 959.
- Krugman, P.R.(1983): "New Theories of Trade Among Industrial Countries", The American Economic Review, Vol. 73, No. 2, Papers and Proceedings of the Ninety-Fifth Annual Meeting of the American Economic Association, pp. 343 - 347.
- Krugman, P.R.(1991): "Geography and Trade", The MIT Press Cambridge
- Krugman, P.R.(2000): "Rethinking International Trade" ,6.th edition The MIT Press, Cambridge
- Krugman, P.R. and M. Obstfeld (2006): "International Economics – Theory and Policy", 7th edition, Person Addison Wesley.
- Markusen, J.R. and A.J Venables (1995): "Multinational firms and the new trade theory", NBER Working Paper No. 5036, pp. 1 – 27.
- Mayer, T. and T.Ottaviano (2007): "EFIM Report 2007 – The internationalization of European Firms: New facts based on firm-level evidence", pp. 1 – 48.
- McCallum,J. (1995): "National Borders Matter: Canada-US. Regional Trade Patterns" The American Economic Review, Vol.85, No.3, pp. 615 – 623.
- Melitz, M.J. (2003): "The Impact of International Trade on Intra-Industry Reallocations and Aggregate Industry Productivity", Econometrica, Vol. 71, No.6 (Nov 2003), pp 1695 – 1725.
- Rose, A.K. (2000): "One Money, One Market: Estimating the effect of Common Currencies on Trade", NBER Working Paper No. 7432, downloaded 1.2.2008, http://www.nber.org/papers/w7432

# Appendix A

**Member list of the European Union per 31.12.2004 :**

1957: Belgium, France, Luxembourg, Italy and the Netherlands.

1973: Denmark, United Kingdom, and, Ireland.

1981: Greece

1986: Portugal and Spain

1995: Austria, Finland and Sweden

2004: Cyprus, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia and Slovenia

**Member list of the EFTA:**

Norway, Lichtenstein, Iceland and Switzerland

**Member list of WTO and the countries' joining date**

Albania 8 September 2000
Angola 23 November 1996
Antigua and Barbuda 1 January 1995
Argentina 1 January 1995
Armenia 5 February 2003
Australia 1 January 1995
Austria 1 January 1995
Bahrain, 1 January 1995
Bangladesh 1 January 1995
Barbados 1 January 1995
Belgium 1 January 1995
Belize 1 January 1995
Benin 22 February 1996
Bolivia 12 September 1995
Botswana 31 May 1995
Brazil 1 January 1995
Brunei Darussalam 1 January 1995
Bulgaria 1 December 1996
Burkina Faso 3 June 1995
Burundi 23 July 1995
Cambodia 13 October 2004
Cameroon 13 December 1995
Canada 1 January 1995
Central African Republic 31 May 1995
Chad 19 October 1996
Chile 1 January 1995
China 11 December 2001
Colombia 30 April 1995
Congo 27 March 1997
Costa Rica 1 January 1995
Côte d'Ivoire 1 January 1995
Croatia 30 November 2000
Cuba 20 April 1995
Cyprus 30 July 1995
Czech Republic 1 January 1995
Democratic Republic of the Congo 1 January 1997
Denmark 1 January 1995
Djibouti 31 May 1995
Dominica 1 January 1995
Dominican Republic 9 March 1995
Ecuador 21 January 1996
Egypt 30 June 1995
El Salvador 7 May 1995
Estonia 13 November 1999
European Communities 1 January 1995
Fiji 14 January 1996
Finland 1 January 1995
Former Yugoslav Republic of Macedonia (FYROM) 4 April 2003
France 1 January 1995
Gabon 1 January 1995
The Gambia 23 October 1996
Georgia 14 June 2000
Germany 1 January 1995
Ghana 1 January 1995
Greece 1 January 1995
Grenada 22 February 1996
Guatemala 21 July 1995
Guinea 25 October 1995
Guinea Bissau 31 May 1995
Guyana 1 January 1995
Haiti 30 January 1996
Honduras 1 January 1995
Hong Kong, China 1 January 1995
Hungary 1 January 1995
Iceland 1 January 1995
India 1 January 1995

Indonesia 1 January 1995
Ireland 1 January 1995
Israel 21 April 1995
Italy 1 January 1995
Jamaica 9 March 1995
Japan 1 January 1995
Jordan 11 April 2000
Kenya 1 January 1995
Korea, Republic of 1 January 1995
Kuwait 1 January 1995
Kyrgyz 20 December 1998
Latvia 10 February 1999
Lesotho 31 May 1995
Liechtenstein 1 September 1995
Lithuania 31 May 2001
Luxembourg 1 January 1995
Macao, China 1 January 1995
Madagascar 17 November 1995
Malawi 31 May 1995
Malaysia 1 January 1995
Maldives 31 May 1995
Mali 31 May 1995
Malta 1 January 1995
Mauritania 31 May 1995
Mauritius 1 January 1995
Mexico 1 January 1995
Moldova 26 July 2001
Mongolia 29 January 1997
Morocco 1 January 1995
Mozambique 26 August 1995
Myanmar 1 January 1995
Namibia 1 January 1995
Nepal 23 April 2004
Netherlands — For the Kingdom in Europe and for the Netherlands Antilles 1 January 1995
New Zealand 1 January 1995
Nicaragua 3 September 1995
Niger 13 December 1996
Nigeria 1 January 1995
Norway 1 January 1995
Oman 9 November 2000
Pakistan 1 January 1995
Panama 6 September 1997
Papua New Guinea 9 June 1996
Paraguay 1 January 1995
Peru 1 January 1995
Philippines 1 January 1995
Poland 1 July 1995
Portugal 1 January 1995
Qatar 13 January 1996
Romania 1 January 1995
Rwanda 22 May 1996
Saint Kitts and Nevis 21 February 1996
Saint Lucia 1 January 1995
Saint Vincent & the Grenadines 1 January 1995
Saudi Arabia 11 December 2005
Senegal 1 January 1995
Sierra Leone 23 July 1995
Singapore 1 January 1995
Slovak Republic 1 January 1995
Slovenia 30 July 1995
Solomon Islands 26 July 1996
South Africa 1 January 1995
Spain 1 January 1995
Sri Lanka 1 January 1995
Suriname 1 January 1995
Swaziland 1 January 1995
Sweden 1 January 1995
Switzerland 1 July 1995
Chinese Taipei 1 January 2002
Tanzania 1 January 1995
Thailand 1 January 1995
Togo 31 May 1995
Tonga 27 July 2007
Trinidad and Tobago 1 March 1995
Tunisia 29 March 1995
Turkey 26 March 1995
Uganda 1 January 1995
United Arab Emirates 10 April 1996
United Kingdom 1 January 1995
United States of America 1 January 1995
Uruguay 1 January 1995
Venezuela (Bolivarian Republic of) 1 January 1995
Viet Nam 11 January 2007
Zambia 1 January 1995
Zimbabwe 5 March 1995

**Countries defined as LDC by the UN:**

Angola
Benin
Burkina Faso
Burundi
Central African Republic
Chad
Comoros
Democratic Republic of the Congo
Djibouti
Equatorial Guinea
Eritrea
Ethiopia
The Gambia
Guinea
Guinea-Bissau
Lesotho
Liberia
Madagascar
Malawi
Mali

Mauritania
Mozambique
Niger
Rwanda
São Tomé and Príncipe
Senegal
Sierra Leone
Somalia
Sudan
Tanzania
Togo
Uganda
Zambia
Haiti

Afghanistan
Bhutan
Cambodia
East Timor
Lao People's Democratic Republic
Burma
Nepal
Yemen
Kiribati
Samoa
The Solomon Islands
Tuvalu
Vanuatu

# Appendix B

## M1 regression

| Source | SS | df MS | | Number of obs | 918 |
|---|---|---|---|---|---|
| | | | | F( 8, 909) | = 658.74 |
| Model | 9016.30615 | 8 1127.03827 | | Prob > F | = 0.0000 |
| Residual | 1555.21986 | 909 1.71091293 | | R-squared | = 0.8529 |
| | | | | Adj R-squared | = 0.8516 |
| Total | 10571.526 | 917 11.5283817 | | Root MSE | = 1.308 |
| | | | | | |
| **ln_ex** | **Coef.** | **Std. Err. t** | **P>t** | **[95% Conf.** | **Interval]** |
| | | | | | |
| **ln_n_gdp** | 1.306521 | .0296928 44.00 | 0.000 | 1.248246 | 1.364795 |
| **lndist** | -.7393062 | .0591884 -12.49 | 0.000 | -.855468 | -.6231444 |
| **lnpop** | -.4213419 | .0337817 -12.47 | 0.000 | -.4876412 | -.3550427 |
| **_Iaar_2000** | -4.29626 | .149567 -28.72 | 0.000 | -4.589797 | -4.002723 |
| **_Iaar_2001** | -4.165297 | .1493244 -27.89 | 0.000 | -4.458358 | -3.872237 |
| **_Iaar_2002** | -4.151488 | .1493293 -27.80 | 0.000 | -4.444559 | -3.858418 |
| **_Iaar_2003** | -4.169666 | .1501847 -27.76 | 0.000 | -4.464415 | -3.874917 |
| **_Iaar_2004** | -4.176861 | .1498339 -27.88 | 0.000 | -4.470922 | -3.882801 |
| **_cons** | 8.667974 | .7867797 11.02 | 0.000 | 7.123858 | 10.21209 |

## M2 regression

| Source | SS | Df MS | Number of obs = | 918 |
|---|---|---|---|---|
| | | | F( 8, 909) | = 622.49 |
| Model | 2621.62385 | 8 327.702981 | Prob > F | = 0.0000 |
| Residual | 478.529649 | 909 .526435258 | R-squared | = 0.8456 |
| | | | Adj R-squared | = 0.8443 |
| Total | 3100.1535 | 917 3.38075627 | Root MSE | = .72556 |
| | | | | |
| **ln_no_ex** | **Coef.** | **Std. Err. t** | **P>t [95% Conf.** | **Interval]** |
| | | | | |
| **ln_n_gdp** | .8071184 | .0164706 49.00 | 0.000 .7747935 | .8394432 |
| **lndist** | -.4799101 | .0328318 -14.62 | 0.000 -.5443451 | -.4154751 |
| **lnpop** | -.286521 | .0187387 -15.29 | 0.000 -.3232972 | -.2497448 |
| **_Iaar_2000** | .0481656 | .0829649 0.58 | 0.562 -.1146594 | .2109906 |
| **_Iaar_2001** | .1013345 | .0828303 1.22 | 0.221 -.0612263 | .2638954 |

| | | | | |
|---|---|---|---|---|
| **_Iaar_2002** | .0395455 | .082833 0.48 | 0.633 -.1230207 | .2021118 |
| **_Iaar_2003** | -.0006451 | .0833075 -0.01 | 0.994 -.1641425 | .1628524 |
| **_Iaar_2004** | -.0674768 | .083113 -0.81 | 0.417 -.2305924 | .0956388 |
| **_cons** | -6.758895 | .4364271 -15.49 | 0.000 -7.615417 | -5.902373 |

## M3 regression

| **Source** | **SS** | **df MS** | | **Number of obs** | **918** |
|---|---|---|---|---|---|
| | | | | F( 8, 909) | = 415.08 |
| **Model** | 3240.54726 | 8 405.06 | | Prob > F | = 0.0000 |
| **Residual** | 887.075735 | 90 9 .9758 | | R-squared | = 0.7851 |
| | | | | Adj R-squared | = 0.7832 |
| **Total** | 4127.623 | 917 4.50122 | | Root MSE | = .98787 |
| **ln_avg_ex** | **Coef.** | **Std. Err. t** | **P>t** | **[95% Conf.** | **Interval]** |
| | | | | | |
| **ln_n_gdp** | .4994025 | .0224252 22.27 | 0.000 | .4553914 | .5434137 |
| **lndist** | -.2593961 | .0447014 -5.80 | 0.000 | -.3471261 | -.1716662 |
| **lnpop** | -.1348209 | .0255133 -5.28 | 0.000 | -.1848927 | -.0847492 |
| **_Iaar_2000** | -4.344426 | .1129589 -38.46 | 0.000 | -4.566116 | -4.122735 |
| **_Iaar_2001** | -4.266632 | .1127756 -37.83 | 0.000 | -4.487963 | -4.045301 |
| **_Iaar_2002** | -4.191034 | .1127793 -37.16 | 0.000 | -4.412372 | -3.969696 |
| **_Iaar_2003** | -4.169021 | .1134254 -36.76 | 0.000 | -4.391627 | -3.946415 |
| **_Iaar_2004** | -4.109384 | .1131604 -36.31 | 0.000 | -4.33147 | -3.887298 |
| **_cons** | 15.42687 | .59420 25.96 | 0.000 | 14.26069 | 16.59305 |

## M4 regression

| **Source** | **SS** | **df MS** | | **Number of obs** | **918** |
|---|---|---|---|---|---|
| | | | | F( 14, 903) | = 381.60 |
| **Model** | 9043.02437 | 14 645.930312 | | Prob > F | = 0.0000 |
| **Residual** | 1528.50164 | 903 1.69269285 | | R-squared | = 0.8554 |
| | | | | Adj R-squared | = 0.8532 |
| **Total** | 10571.526 | 917 11.5283817 | | Root MSE | = 1.301 |
| | | | | | |
| **ln_ex** | **Coef.** | **Std. Err. t** | **P>t** | **[95% Conf.** | **Interval]** |
| | | | | | |
| **ln_n_gdp** | 1.250677 | .0403658 30.98 | 0.000 | 1.171456 | 1.329899 |
| **lndist** | -.6610513 | .0760423 -8.69 | 0.000 | -.8102915 | -.5118111 |
| **lnpop** | -.3739574 | .0407252 -9.18 | 0.000 | -.4538844 | -.2940304 |

| | | | | | |
|---|---|---|---|---|---|
| **efta** | .6322745 | .410895 1.54 | 0.124 | -.1741458 | 1.438695 |
| **SCAND** | .7178188 | .4314892 1.66 | 0.097 | -.1290196 | 1.564657 |
| **COMBOR** | .535328 | .3475878 1.54 | 0.124 | -.146846 | 1.217502 |
| **wto** | .26868 | .1232519 2.18 | 0.030 | .0267865 | .5105735 |
| **eu** | .1800278 | .1851804 0.97 | 0.331 | -.1834063 | .5434619 |
| **ldc** | -.0592139 | .1351416 -0.44 | 0.661 | -.324442 | .2060143 |
| **_Iaar_2000** | -4.307154 | .1488517 -28.94 | 0.000 | -4.599289 | -4.015018 |
| **_Iaar_2001** | -4.17492 | .1485859 -28.10 | 0.000 | -4.466534 | -3.883306 |
| **_Iaar_2002** | -4.158456 | .1485871 -27.99 | 0.000 | -4.450072 | -3.86684 |
| **_Iaar_2003** | -4.177217 | .1496164 -27.92 | 0.000 | -4.470854 | -3.883581 |
| **_Iaar_2004** | -4.189402 | .1494435 -28.03 | 0.000 | -4.482699 | -3.896105 |
| **_cons** | 8.315937 | .8615322 9.65 | 0.000 | 6.625099 | 10.00678 |

## M5 regression

| **Source** | **SS** | **df MS** | **Number of obs** | **918** |
|---|---|---|---|---|
| | | | F( 14, 903) | = 397.89 |
| **Model** | 2667.70223 | 14 190.550159 | Prob > F | = 0.0000 |
| **Residual** | 432.451275 | 903 .478905067 | R-squared | = 0.8605 |
| | | | Adj R-squared | = 0.8583 |
| **Total** | 3100.1535 | 917 3.38075627 | Root MSE | = .69203 |

| **ln_no_ex** | **Coef.** | **Std. Err. t** | **P>t** | **[95% Conf.** | **Interval]** |
|---|---|---|---|---|---|
| **ln_n_gdp** | .7471042 | .0214708 34.80 | 0.000 | .7049657 | .7892428 |
| **lndist** | -.4326626 | .0404474 -10.70 | 0.000 | -.5120444 | -.3532807 |
| **lnpop** | -.2303915 | .021662 -10.64 | 0.000 | -.2729052 | -.1878777 |
| **efta** | 1.17434 | .2185579 5.37 | 0.000 | .7453998 | 1.603281 |
| **SCAND** | .8540807 | .2295121 3.72 | 0.000 | .4036415 | 1.30452 |
| **COMBOR** | .4564953 | .1848844 2.47 | 0.014 | .0936422 | .8193483 |
| **wto** | .4022629 | .0655585 6.14 | 0.000 | .2735981 | .5309278 |
| **eu** | .0502983 | .0984988 0.51 | 0.610 | -.1430149 | .2436114 |
| **ldc** | -.0145339 | .0718827 -0.20 | 0.840 | -.1556105 | .1265428 |
| **_Iaar_2000** | .0315686 | .0791752 0.40 | 0.690 | -.1238203 | .1869575 |
| **_Iaar_2001** | .0867091 | .0790338 1.10 | 0.273 | -.0684022 | .2418205 |
| **_Iaar_2002** | .0277007 | .0790345 0.35 | 0.726 | -.127412 | .1828133 |
| **_Iaar_2003** | -.0175134 | .079582 -0.22 | 0.826 | -.1737007 | .1386738 |
| **_Iaar_2004** | -.0806932 | .07949 -1.02 | 0.310 | -.2366999 | .0753135 |
| **_cons** | -6.994954 | .4582549 -15.26 | 0.000 | -7.894323 | -6.095586 |

## M6 regressions

| Source | SS | df MS | | Number of obs | 918 |
|---|---|---|---|---|---|
| | | | | F( 14, 903) | = 238.08 |
| **Model** | 3247.73741 | 14 231.981243 | | Prob > F | = 0.0000 |
| **Residual** | 879.885591 | 903 .974402648 | | R-squared | = 0.7868 |
| | | | | Adj R-squared | = 0.7835 |
| **Total** | 4127.623 | 917 4.50122465 | | Root MSE | = .98712 |

| ln_avg_ex | Coef. | Std. Err. t | P>t | [95% Conf. | Interval] |
|---|---|---|---|---|---|
| **ln_n_gdp** | .5035733 | .0306262 16.44 | 0.000 | .4434664 | .5636801 |
| **lndist** | -.2283888 | .0576946 -3.96 | 0.000 | -.3416199 | -.1151576 |
| **lnpop** | -.1435659 | .0308989 -4.65 | 0.000 | -.204208 | -.0829239 |
| **efta** | -.542066 | .3117533 -1.74 | 0.082 | -1.153911 | .0697792 |
| **SCAND** | -.1362619 | .3273784 -0.42 | 0.677 | -.7787731 | .5062492 |
| **COMBOR** | .0788325 | .263721 0.30 | 0.765 | -.4387449 | .5964099 |
| **wto** | -.133583 | .0935134 -1.43 | 0.153 | -.3171118 | .0499459 |
| **eu** | .1297295 | .1404997 0.92 | 0.356 | -.1460143 | .4054734 |
| **ldc** | -.0446799 | .1025343 -0.44 | 0.663 | -.2459131 | .1565533 |
| **_Iaar_2000** | -4.338722 | .1129364 -38.42 | 0.000 | -4.560371 | -4.117074 |
| **_Iaar_2001** | -4.261629 | .1127347 -37.80 | 0.000 | -4.482881 | -4.040376 |
| **_Iaar_2002** | -4.186157 | .1127356 -37.13 | 0.000 | -4.407411 | -3.964902 |
| **_Iaar_2003** | -4.159704 | .1135166 -36.64 | 0.000 | -4.382491 | -3.936917 |
| **_Iaar_2004** | -4.108709 | .1133854 -36.24 | 0.000 | -4.331238 | -3.886179 |
| **_cons** | 15.31089 | .6536596 23.42 | 0.000 | 14.02802 | 16.59376 |