RATIO

Ratio Working Paper No. 220

Distance Sensitivity of Export: A Firm-Product Level Approach

Viroj Jienwatcharamongkhol*

*viroj.j@nek.lu.se, The Ratio Institute, P.O. Box 3203, SE-103 64 Stockholm, Sweden and Department of Economics, Lund University, P.O. Box 7082, SE-220 07 Lund, Sweden.

The Ratio Institute P.O. Box 3203, SE-103 64 Stockholm, Sweden, www.ratio.se

Distance Sensitivity of Export: A Firm-Product Level Approach

Viroj Jienwatcharamongkhol
Department of Economics
Lund University

Abstract

Recent literature suggests that the extent to which exports of a product is influenced by distance depends on the product characteristics. Differentiated products with non-standardised attributes are typically claimed to be more distance-sensitive as transactions should involve interactions between buyers and sellers. But the empirical evidence still finds conflicting results. Previous studies have examined the effect of distance on export values across different product groups. This paper employs a gravity model on Swedish firm-product level export data to analyse the effect of distance on the export decisions as well as export values, respectively. The focus is on how the influence of distance varies across differentiated and non-differentiated products. For both export participation and intensity decisions, the results are not in line with the network/search view and suggest that homogeneous products are more sensitive to distance than differentiated products when controlling for annual shocks and industry heterogeneity. Moreover, I find evidence of a learning effect from past trade experience.

Keywords: distance sensitivity, export decisions, gravity model, firm-product level, micro-data
JEL Classification: F12, F14, F41

Address: P.O. Box 7082, SE-220 07 Lund, Sweden, Tel.: +46-46 222 79 17, Fax: +46-46 223 46 13, Viroj.J@nek.lu.se. The author is grateful for the grant from Torsten Söderbergs Stiftelse and the support from Departments of Economics at Lund University and Jönköping International Business School and Ratio Institute. I also thank my supervisors, Prof. Fredrik Sjöholm, Prof. Martin Andersson, Assoc.Prof. Patrik Tingvall, and the colleagues.

Introduction

The negative relationship between distance and trade has been quoted as one of the most robust empirical findings in the international trade literature (Leamer & Levinsohn, 1995). Countries situated close to each other trade more intensively than countries that are farther apart. Such an effect can be caused by the exporter's transportation costs of shipping from home to the destination, so that the greater distance entails greater transportation costs. Large distance can also imply greater transaction costs of having to deal with culturally or institutionally unfamiliar markets, which in turn makes it more difficult for exporters to establish the necessary network of distributors abroad. These costs can arise from procedural differences, communication misalignment, or legal compliance. The more unfamiliar the market, the higher the transaction costs.

The different types of costs associated with distance imply that distance can influence exports in various ways. The basic hypothesis is that, first, distance may affect the decision whether to export or not and decreases the extensive margin of exports by reducing the number of firms present and the number of available products in a particular market abroad. That is, fewer firms are expected to choose to export to more distant markets. This is because distance increases transaction costs, i.e. the fixed and sunk entry costs of setting up contacts and distribution network at unfamiliar host countries abroad. Such increase in costs lessens, in general, the number of firms that can afford these high entry costs (in other words, cross the productivity threshold) to become exporters (Melitz, 2003; Andersson, 2007); and, similarly, the number of products to be exported for each individual firm.

Secondly, distance also affects the intensity decision and decreases the intensive margin of export by reducing the size of export per firm. M. Lawless (2010) considers distance to mainly capture the variable costs. But there might also be some fixed cost element in the distance. This is because after entry, each firm has to incur some unknown per-period fixed costs of maintaining their presence in the market (Segura-Cayuela & Vilarrubia, 2008) and also some market penetration costs of advertising to capture the market share (Arkolakis, 2008). The greater dissimilarity between sellers and buyers tends to increase the uncertainty that leads to the increase of these per-period fixed costs and hence reduces the intensive margin of each firm. Similarly for the marketing costs, producers from far away are less likely to be known to the consumers in the market and a firm has to invest a considerable amount to publicise its products, and will likely result in the reduction of the export capacity per firm.

The role of distance on transaction costs has been analysed in recent decades and quickly gained a place in the trade literature. A meta-study by Disdier and Head (2008) confirms the persistence of distance effect and G. M. Grossman (1998) shows that the distance effect is of a greater magnitude than could be accounted to transportation costs alone. Recent studies regard distance not only as a geographical unit. The so-called intangible barriers that are institutional and cultural (dis)similarities are the additional dimensions that have a role in affecting a firm's decision to engage in export. The institutional differences between home and destination market can involve the protection of property rights and contract enforcement (Anderson & Marcouiller, 2002). The im-

perfect alignment would eventually impose additional transaction costs on the exporters due to informational frictions from the uncertainty (Huang, 2007). Accordingly, cultural similarities in terms of language, religion, colonial ties are found to be facilitating export because the trading partners can more easily communicate and share common understanding with each other (Rauch, 1999; G.-J. Linders & De Groot, 2006; Lankhuizen, de Graaff, & de Groot, 2012).

However, the impact of distance is not uniform across products but varies in magnitude due to the product characteristics. The pioneering work by Rauch (1999) introduces the network/search view and argues that differentiated products assert greater sensitivity than homogeneous products. Rauch employs aggregate trade data and find some evidences that the elasticity of trade with respect to distance is greater for differentiated products. But we can still find the conflicting empirical results in other studies, so it is still uncertain whether homogeneous products are more sensitive to distance than differentiated products or the other way around. In the original study by Rauch, moreover, there is no separation between the decision whether to export or not and how much to export. This paper aims to examine the distance effect at the firm-product level in order to provide empirical evidence for this puzzle.

In the case of homogeneous products, which are categorised as the products on an organised exchange and the products with referenced price in Rauch (1999), the products are standardised and can be compared by their prices without having to identify the producer's trademark. The trader can scan and obtain the price information through trade publications (or internet portals nowadays) and then match the buyers and sellers more easily. This makes the search costs lower in comparison to the differentiated products, in which their characteristics vary in many dimensions, e.g. colour or technical features. Matching product characteristics across various markets would also include the identification of the producers[1]. The same connection as in homogeneous products has to arise from a search process instead, so buyers and sellers need to establish network ties in order to match orders. This increases the transaction costs. Therefore, it is expected that differentiated products would assert greater sensitivity towards distance and other intangible barriers. But it can be argued that homogeneous products assert greater distance sensitivity due to the competition of similarly-produced homogeneous products from other competing countries nearby the destination market, whereas the monopolistic nature of differentiated products enables a trade across great distances.

However, the literature that studies the distance effect on product export chiefly looks at the aggregate national level but not how each individual firm behaves. If we take a look from the firm's perspective, we can see that each individual firm faces the following decisions: (i) whether to export or not, (ii) where to export, (iii) which products to export, and (iv) how much to export each of these products. The first two questions are dealt with elsewhere (see for example Bernard and Jensen (2004), K. Lawless Martina; Whelan (2008). Whereas the last two questions will be the main

[1]Take, for example, a case of price comparison for personal notebooks. You would need to gather information on many things, including the screen size, processor speed, RAM capacity, hard-drive capacity and reading technology, graphics card and memory, operating system version, and manufacturer.

focus of this paper, in which it examines the aforementioned participation and intensity decisions of firm's export.

Rauch (1999) classifies products into three categories, organised exchange, reference priced, and differentiated, and estimates a gravity equation of aggregate bilateral trade of selected 63 countries in 1970, 1980, and 1990. The result of higher distance effect for differentiated products is in support of his hypothesis that differentiated products assert higher trade costs besides transportation. Similar result from a different model specification is also obtained in G.-J. M. Linders (2006); Huang (2007).

On the other hand, G.-J. Linders, De Groot, and Rietveld (2005) find the opposite results, despite a rather *ad hoc* treatment of zeros in the data while estimating bilateral trade flows of 114 countries with OLS. Möhlmann, Ederveen, de Groot, and Linders (2010) uses an alternative estimation method on the 55 countries for the study, i.e. Heckman selection model with country dummies instead of the standard OLS. Their given explanations are that differentiated products are produced in fewer places and preferably traded over a larger distance, and that the intangible costs are relatively less important for the products on organised exchange. Lankhuizen et al. (2012) extend on these papers by using finite mixture model in order to endogenously group the products into homogeneous segments that are sensitive to geographic distance in various dimensions. The data is from 72 countries in 2000. Among the findings for the eight segments, for example, machinery and transport products are sensitive to high geographic distance, while bulk goods and crude materials are less sensitive to geographic distance.

So far the analysis of distance sensitivity mainly looks at a static picture. But we know that exporting is a dynamic process and should be treated as such. Once the exporters gain access to the foreign markets, the upfront fixed costs have already been paid, so it is reasonable that the costs associated with exporting to the same market should be lower in succeeding years. This is because (i) the institutions are rigid and any procedural changes tend to be slow (ii) as a result, a firm learns to adapt to the market better, e.g. know which forms to submit or whom to contact for tax refund, and be more efficient in later years.

The idea that a firm learns from its past export experience is shared among many scholars. Helpman (1984), G. Grossman and Helpman (1993), and Clerides, Lach, and Tybout (1998) formally show that learning by exporting may lead to higher productivity. The technical or management expertise and the best practices of international buyers lead the exporting firms to increase their stock of knowledge. The increased knowledge then helps them to be more productive later. Besides, the productivity gain is channelled through higher competition in the foreign markets (Verhoogen, 2008).

In terms of import experience, an establishment of contacts in the past also leads a firm to learn. However, the empirical evidence at the disaggregated firm/plant level is mixed. Vogel and Wagner (2010) do not find the evidence of learning-by-importing productivity premia for German manufacturing firms. Whereas Kasahara and Rodrigue (2008) find that importing foreign intermediate goods improves firm's productivity.

The novelty in this paper is the greatly detailed data analysis. This is done by examining the distance sensitivity across product groups based on export decisions at

the disaggregated firm-product level, which has never been done in previous studies. Studying the export decisions at this level yields an insight into the distance effect as reflected by each individual exporting firm and each type of product. I also take into account the various dimensions of distance including cultural and institutional similarities. Moreover, because the process of exporting is dynamic, this paper takes into account the past experiences, both import and export in a country, to provide an evidence of learning effects.

The rest of the paper is organised as follows. The methodology section specifies the empirical strategy and the econometric estimation methods. Then the description of the data follow. The results section presents and discusses the findings. The last section concludes the paper.

Methodology

Estimation

In order to explain the decisions to export, many empirical studies in international trade usually employ the gravity equation. Throughout the years it has been tested and the general concensus is that the gravity equation is robust at exhibiting the negative effect of distance. Since the pioneering work by Tinbergen (1962) some scholars have provided the theoretical foundation for it, including Andersson and van Wincoop (Anderson, 1979; Anderson & Wincoop, 2003), Bergstrand (Bergstrand, 1985), and recently an extension of the model in Egger and Pfaffermayr (Egger & Pfaffermayr, 2011).

The basic equation used in this paper is formulated as the following:

$$X_{ijkt} = \beta_0 Y_{jt}^{\beta_1} \Gamma_j^{\beta_2} \Theta_{i,t-1}^{\beta_3} \delta_{ijk,t-1}^{\beta_4} \varepsilon_{ijkt}, \tag{1}$$

where the dependent variable X_{ijkt} is the export by firm i of product k to country j at time t. The independent variables are vectors of destination country variables, distance variables, lagged firm-specific control variables; lagged import variable. A list of all the variables used and their description is in the Appendix.

The analysis focuses on the influence of distance on firm's decisions whether to export a product to a given market or not and how much to export. Therefore, each individual firm encounters the following choice problem:

i) **Participation:** each individual firm chooses to export a certain product, out of its product portfolio, to a certain country, out of its established networks, in each year. Hence, the set of possible products and countries is constructed from each respective firm's possibility set of products and countries, defined from its experience throughout the period of study. Instead of constructing the possibility set for each firm from all manufactured products and all countries, this approach means that a shoes company, for example, would not consider to export automobile spare parts to foreign countries where it never has any past or future contacts. This way reduces the possibility set tremendously and allow me to make the analyses manageable[2].

[2]Alternatively, the possibility set will explode as we add more dimensions into the consideration.

ii) **Intensity:** at any given year, each individual firm that decides to export a particular product to a particular country also has another decision to make, that is how much to export.

In regarding the export decisions, I consider three different cases:

1. Export participation and export intensity are separate decisions. In order to estimate the two, I use probit regression and ordinary least square (OLS) estimator. The dependent variable in the participation is a dummy indicating whether a firm export a product to the destination country in current year or not, whereas it is total export value in the intensity. The full dataset is used in the participation while only the observed positive exports are included in the intensity estimation. Including all zero exports in the intensity estimation will lead to downward bias otherwise.

2. Export is modelled as a one-step decision, in which there is no distinction between participation and intensity decisions. This means that both zero exports and the export amount are used in the estimation[3]. Wagner (2001) argues that firms take the *ex ante* costs of entering into export into calculation and choose the profit-maximising amount of exports accordingly. Then, the occurrence of zero export quantity is merely the result of the average total costs exceeding the profit-maximising price. So there is no participation decision involved. The export quantity, in this case, is a ratio between total export value in the current year of a product to a country and total sales of the firm. Due to the fractional response nature of the variable, I use quasi-maximum likelihood estimator (QMLE), which is proposed by Papke and Wooldridge (1996) and discussed in Wagner (2001). The traditional tobit estimator is not suitable in this case because the non-export zeros here are not censored (Baum, 2008). The range of the dependent variable is strictly from 0 to 1. Censoring the zeros away from the sample leads to an upward bias.

3. The two export decisions are jointly estimated. Because the zeros can be a consequence from an economic reason that a firm with low productivity cannot afford the high cost of export and choose to self-select themselves not to participate in export. Therefore these zeros can provide an important information for the analysis. Unlike in the second case where there is no distinction between the two export decisions, the zeros in this case are generated from a self-selection process that is different from the intensity decision. The zero-inflated beta (ZOIB) estimator, developed by Cook, Kieschnick, and McCullough (2008), takes into account such an inflation process in estimating the export intensity. Alternatively, I use Heckman selection model (Heckman, 1979) as a robustness

---

Consider a set of only 500 firms with 100 possible products shipping to 165 countries in a 10-year period. The total number of observations in the dataset to work with is 82.5 million.

[3]One common issue that arises from this approach is the frequent occurrence of zero observations. The problem of frequent zeros is typical in trade data including this one, in which the zeros account for 94.7% of total observations. The problem arises because the estimation model is in a linearised form (by logarithmic transformation), which lead any zeros in the original dataset undefined. There are several alternative estimation methods that deal with data with frequent zeros, for example zero-Inflated Poisson (ZIP), zero-inflated negative binomial (ZINB), pseudo poisson maximum likelihood (PPML) (Santos Silva & Tenreyro, 2006) but such models are mainly appropriate for count data and an evidence of superiority over Heckman is still debatable (Martínez-Zarzoso, 2013; Martin & Pham, 2008).

check[4]. Normally, in order to consistently estimate a Heckman selection model, there should be an addition of at least one variable in the selection equation. I choose to include the human capital variable, which is the fraction of employees with more than three years of university studies. The argument for human capital as the exclusion variable is as follows. In order for the firm to start an export in a new market, it requires certain specific knowledge or network connection, usually embedded in high-ranking personnels in a managerial position, to establish a contractual transaction. The push factor, such as profitability goal or expansion strategy, from the firm is predominant. Whereas after the initial entry, it becomes routine process which can be executed by lower-ranking administrative personnels to fulfill the export orders. The pull factor, such as the consumer demand in the country, becomes more important. Hence, the human capital is more important in the entry decision and becomes not so important in determining how much to export.

Empirical strategy

In preparation for the dataset, I follow the approach from a paper on local export spillovers in France by Koenig, Mayneris, and Poncet (2010). The main advantage is the focus on the within transformation of each of firm's decision possibility set in order to avoid exhausting the analyses with explosively large dataset.

To begin with, I exclude self-employed firms - firms with zero employee, zero or negative sales and value-added due to the log-linearised model that renders these observations undefined, which leaves a total of 23,943 manufacturing firms, of which 6,007 firms are exporters. Next, I include only the active firms that appear throughout the ten-year period of study and with at least one export start. The reason not to include firms with temporary exits is due to computational constraint[5]. The result from excluding them reduces the total from 6,007 to 2,151 export firms in the end.

Next, I match the firms with countries and products with positive export value to form a triad (firm-country-product). I then exclude the triads that persist for the whole ten-year period, which accounts for 22 out of the total 31,375 triads or roughly 0.07%. So the possibility set of a triad will include at least one start during the entire period. Lastly, I fill the possibility sets for each individual firm with its respective possible countries and products and years. This procedure leaves a total of 395,900 observations for the analyses. The justification is that firms that persistently export the same product to the same country are already paying the upfront fixed entry costs so the comparison to new entrants would be incorrect. To control for any experience a firm has in the country, regardless of products, I also include a dummy with 1 if a firm has been exporting to that country before and 0 if it is a new export country.

[4]Both outcome and selection equations can be either jointly estimated with maximum likelihood or as a two-step approach, with maximum likelihood in the first stage and normal OLS in the second. I rely on the first approach to follow Verbeek (2008) who points out that the OLS standard errors from the two-step estimator are incorrect, whereas the maximum likelihood provides a consistent and asymptotically efficient estimator.

[5]Otherwise, using all exporters will lead to total observations of 98,375,860, which deem impossible to perform an analysis.

From the basic equation to estimate in the previous subsection, there are sets of country, distance, firm-specific, and previous import variables. The country variables include the nominal Gross Domestic Products (GDP) and the GDP per capita of only the destination countries. Usually in the gravity equation, the model also include these variables from the home country. But since the analysis is executed for exporters registered in Sweden only, there is no variation across firms in the dataset. The gravity equation here is therefore one-sided.

There are several distance variables included in the estimation. These variables constitute both the tangible and intangible barriers to trade. Firstly, the main variable of interest, i.e. the geographical distance, measured in kilometres away from Sweden. The measurement is a weighted great-circle distance that takes into account the main trading and financial cities of each country. Secondly, the contiguity or common border dummy. This variable takes a value of 1 if a destination country shares a border with Sweden, and 0 otherwise. It controls for neighbor trades, which tend to be disproportionally high and potentially bias the result. Thirdly, I include the landlocked dummy due to the fact that transportation costs are higher for countries without direct access to sea. Fourthly, a dummy indicating countries with English as the official language is included. Since there is no other countries sharing Swedish as the official language, the traders have to use English as the main *lingua franca*. Lastly, I control for the regional trade agreement. This is because the streamlined institutional system and the abandon of tariffs within the common trade area will induce gross trade creation through integration, as evidenced by Aitken (1973) and Carrère (2006) among others [6].

The firm-specific set of variables is to control for firm's heterogeneity. I include value-added at year's end, human capital in the estimation equation. Human capital is calculated as the share of highly-educated, i.e. graduated above secondary education level, workforce within a firm. In order to avoid an endogeneity problem, I lag these variables by one year. To control for corporate affiliation, I also include dummies denoting a firm's affiliation to domestic or multinational corporation. Non-affiliation is used as a reference group to avoid the dummy trap problem. Initially, I aim to include dummies for destination countries in the full model, as is recommended by many scholars (Anderson & Wincoop, 2003; Feenstra, 2002; Mátyás, 1997; Redding & Venables, 2004) to account for country-specific unobserved attributes. But due to computational difficulties, the probit estimation and Heckman do not converge. I report the OLS results in Table 10.

Although the original dataset contains all firms in Sweden, the manufacturing sector is chosen while leaving away the service sector. This is because I want to focus only on firms that export what they are producing. Many firms within the service sector, on the other hand, are intermediaries or trading firms. The manufacturing sector includes the industries indicated by the two-digit NACE[7] revision 1.1 codes 15 to 36.

[6] I also considered including a dummy indicating EU membership states to take into account the reporting policy that excludes any firms with annual imports from or exports to EU members below 1 million euros from the database. But due to a high collinearity between the EU dummy and the regional trade agreement dummy, I decide to drop the EU dummy.

[7] NACE is abbreviated for *Nomenclature des Activités Économiques dans la Communauté Européenne* or Classification of Economic Activities in the European Union.

The distribution of exporters per total producers and the share of export per total sales for each industry is presented in the Appendix.

Data

The data for this analysis is a merge from three datasets. First, the firm-level export-import data contains the export value and weight of products defined at 8-digit Swedish equivalent of Harmonised System (HS) for each individual firm. Second, the firm characteristics variables, including value-added, affiliation and several other variables. Both datasets are linked by a unique firm identification, encoded by Statistics Sweden, and contain the population of all firms which are registered at the Swedish Tax Agency[8]. These two data are complemented by country and distance variables, available from *Centre d'Étude Prospectives et d'Information Internationales* (CEPII). The period of analysis is ten years from 1997 to 2006 and includes in total 2,151 manufacturing firms and 2,553 unique products. The variables in use and descriptive statistics are listed in the Appendix.

The product classification used in this paper refers to Rauch (1999), in which the categorisation procedure can be summarised as the following. The products on organised exchanges and referenced prices according to Standard International Trade Classification (SITC) at the five-digit level are identified by looking up the trade publications. The largest share of global trade value determines the category the products belong at the four-digit aggregation[9]. Rauch's classification[10] then is converted to EU Combined Nomenclature (CN)[11]. The share of both products in my sample over time is presented in Table 1. Similar to Rauch (1999), most of the exported products are differentiated and its share is increasing over the years.

Results

The results are presented in three cases. The first case separates the export participation apart from export intensity decision. In the second case, firm's exports are decided in one step. Lastly, both export decisions are jointly determined. Throughout this section, the discussion will focus on two main topics, namely the statistical and economic significance. Due to space, the results on display will be the marginal effects for interpretation of economic significance, whereas the summary of statistical results

[8]The databases are part of Microdata On-line Access (MONA) service provided by Statistics Sweden. All analyses are executed via a remote access to the station. For information regarding the access to the database, please refer to Statistics Sweden at `www.scb.se`.

[9]For this division, there are conservative and liberal classifications, in which the former minimises the number of products belonging to either organised exchange or reference priced and the latter maximises this number. This does not affect my results because I mainly look at homogeneous products as a whole.

[10]The classification is available through Jon Haveman's website at `www.macalester.edu/research/economics/PAGE/HAVEMAN/Trade.Resources/TradeData.html#Rauch`.

[11]The conversion is from SITC to Harmonized System (HS) and lastly to CN, similar to Persson (2013) but the classification in this paper is based on SITC rev. 2 while Persson's study is based on SITC rev. 3.

Table 1:: Share of Swedish exported products in value

Year	Homogeneous Products	Differentiated Products
1997	12.51	87.49
1998	12.05	87.95
1999	10.26	89.74
2000	7.46	92.54
2001	12.69	87.31
2002	11.90	88.10
2003	12.42	87.58
2004	13.25	86.75
2005	9.60	90.40
2006	9.37	90.63

and the estimated coefficients of the distance variable are displayed in tables 10 and 11 in the Appendix.

Separate export participation and export intensity decisions

Table 2 displays the regression results using standard probit for the participation decision (columns 1-3) and OLS for export intensity, measured as total export value (columns 4-6). The OLS results are limited only to observations with positive exports. The result of the overall products is displayed first (columns 1 and 4) and followed by separate regressions for homogeneous and differentiated products. Here I control for both annual shocks and industry heterogeneity. The probit results are presented with the marginal effects for a convenient interpretation.

For the GDP variable, it shows positive and significant in almost all regressions, meaning that the market size positively affects both export decisions. Concerning the magnitude of the GDP effect on export participation, a 1% increase of GDP of the destination country leads to an increase of 0.1 - 0.5 percentage point in the probability to export, holding other variables at mean value. This is a rather small impact. On the other hand, the positive effect on export intensity, measured as total value, is approximately 8.5% - 20%, a sizable increase. Moreover, differentiated products assert greater GDP effect than homogeneous products. While it exhibits the effect of GDP per capita only on the intensity of differentiated products, roughly 19.3% for a 1% increase in the GDP per capita of the destination country.

Regarding various distance and intangible barrier variables, distance, landlockedness, and regional trade agreement have negative impacts on the participation, whereas the negative impact on the intensity comes from distance, contiguity, and landlockedness. This suggests the high costs of exporting and the competition. Whereas the positive impact of English language, contiguity (on the participation), and trade agreement (on the intensity) suggests the institutional similarity. In terms of magnitude, the

effect on exports is greatest when the border is shared between the trading countries, i.e. approximately 3 percentage point increase in probability of export and a huge decrease of 53%-94% in the total value of differentiated products. Comparing across products, homogeneous products display greater distance effects in both decisions, i.e. a 0.2 percentage point difference in the probability to export and a 15% difference in total value, while the results are mixed for other variables. This contradicts the network/search view hypothesis that exports of differentiated products depends more on the familiarity than homogeneous products. Note also that the size of these distance coefficients for the export intensity are lower than in other studies, 0.1 - 0.9 in most variables, compared to 0.6 - 1.2 in Rauch (1999) and G.-J. M. Linders (2006) for example. This is because the gravity that I employ here is one-sided that includes unilateral trade flows from Sweden to the rest of the world only, unlike the bilateral trade flows used in other studies.

The value-added variable confirms Melitz's model on firm's productivity and export. The positive coefficients imply that higher productivity leads to a slightly greater probability to export and greater extent to export size, especially for differentiated products (a 23% increase in total value for every 1% increase in value added). Affiliation to a corporate group, either domestic or multinationals (MNEs), negatively relate to the decision to participate but positively relate to export intensity. Belonging to MNEs increase export value by roughly 40%. These negative results contradict the expectation, i.e. most international trade flows usually occur within MNEs and thus affiliation to multinationals should induce the probability of firms to participate in exports. One possible explanation is that the permanent exports, which consist mainly of firms affiliated to corporations and MNEs, are excluded from the dataset.

For trade experience variables, when a firm has imported any products from the destination a year before, it has a complementary effect for both types of products, inducing the export participation by 0.2-0.3 percentage point and the intensity to export by 11%-16%. Moreover, if a firm has a market presence by exporting, regardless of products, in the destination country before, it positively impacts the decision whether to export or not for homogeneous products only (0.5 percentage point). But the value of export in succeeding years for both types of products is unaffected by its export experience in the country possibly due to the capacity limitation.

One-step export decision

In this case, a firm only considers how much to export and zero export is a result of profit-maximising calculation. However, the problem from the presence of frequent zeros prompts a consideration of an alternative estimator. The calculated marginal effects from the Papke-Wooldridge quasi-maximum likelihood estimator (QMLE) is presented in table 3.

Statistically speaking, when we look at the raw coefficients in table 10, the GDP variable exhibits a positive effect, similar to previous result. But GDP per capita shows negative sign and only slightly significant for homogeneous products. Similar to the previous results, homogeneous products exhibit greater negative effect on exports, measured as value per total sales, for the distance variable. In contrast, contiguity is not significant

Table 2:: Probit and OLS results by product groups

Variables	(1)	(2)	(3)	(4)	(5)	(6)
	Probit (Export Dummy)			OLS (Export Value)		
	All	Homog.	Diff.	All	Homog.	Diff.
GDP	0.004***	0.001	0.005***	0.183***	0.085***	0.200***
(log)	(0.000)	(0.001)	(0.000)	(0.013)	(0.029)	(0.014)
GDP Per Capita	0.001	0.000	0.000	0.172***	-0.016	0.193***
(log)	(0.001)	(0.001)	(0.001)	(0.023)	(0.055)	(0.025)
Distance	-0.025***	-0.027***	-0.025***	-0.249***	-0.400***	-0.248***
(log)	(0.001)	(0.002)	(0.001)	(0.031)	(0.078)	(0.034)
Contiguity	0.026***	0.028***	0.025***	-0.889***	-0.529***	-0.938***
	(0.001)	(0.002)	(0.001)	(0.042)	(0.106)	(0.046)
Landlocked	-0.013***	-0.005	-0.015***	-0.402***	-0.455***	-0.341***
	(0.002)	(0.003)	(0.002)	(0.064)	(0.170)	(0.068)
English Dummy	0.007***	0.000	0.008***	0.377***	0.661***	0.313***
	(0.001)	(0.003)	(0.002)	(0.058)	(0.171)	(0.061)
Regional Trade	-0.018***	-0.018***	-0.017***	0.114*	-0.052	0.131**
Agreement	(0.002)	(0.003)	(0.002)	(0.060)	(0.140)	(0.066)
Value Added	0.003***	0.001	0.003***	0.213***	0.131*	0.229***
(log)(lag)	(0.001)	(0.002)	(0.001)	(0.031)	(0.074)	(0.034)
Domestic Corporation	-0.011***	-0.008***	-0.013***	0.023	0.038	0.036
	(0.001)	(0.002)	(0.001)	(0.031)	(0.069)	(0.035)
Multinationals	-0.013***	-0.015***	-0.013***	0.419***	0.393***	0.410***
	(0.001)	(0.002)	(0.001)	(0.039)	(0.097)	(0.043)
Import Dummy	0.002***	0.003*	0.003***	0.139***	0.108	0.162***
(lag)	(0.001)	(0.002)	(0.001)	(0.028)	(0.068)	(0.031)
Country Experience	-0.000	0.005***	-0.001	0.021	0.034	0.038
	(0.001)	(0.002)	(0.001)	(0.030)	(0.072)	(0.032)
Observations	355,612	85,284	270,328	19,021	3,688	15,333
R-squared	0.055	0.078	0.052	0.159	0.121	0.181

Note: The number in the table represents the marginal effects. All regressions include constants, year and industry dummies but are not reported. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3:: Papke-Wooldridge QMLE results

Variables	(7) QMLE (Export/Total Sales) All	(8) Homog.	(9) Diff.
GDP	0.274***	0.097***	0.333***
(log)	(0.023)	(0.032)	(0.029)
GDP Per Capita	-0.053	-0.075*	-0.050
(log)	(0.052)	(0.039)	(0.070)
Distance	-0.973***	-0.628***	-1.088***
(log)	(0.055)	(0.073)	(0.070)
Contiguity	0.053	0.026	0.063
	(0.085)	(0.069)	(0.114)
Landlocked	-0.169*	0.232	-0.294***
	(0.092)	(0.258)	(0.097)
English Dummy	0.969***	0.439*	1.100***
	(0.176)	(0.237)	(0.218)
Regional Trade	-0.869***	-0.582***	-0.902***
Agreement	(0.139)	(0.173)	(0.170)
Value Added	0.048	-0.035	0.076
(log)(lag)	(0.088)	(0.061)	(0.115)
Domestic Corporation	-0.701***	-0.325***	-0.789***
	(0.049)	(0.052)	(0.062)
Multinationals	-0.986***	-0.397***	-1.126***
	(0.057)	(0.072)	(0.070)
Import Dummy	-0.309***	-0.234***	-0.286***
(lag)	(0.055)	(0.068)	(0.068)
Country Experience	0.018	0.140**	-0.026
	(0.063)	(0.060)	(0.082)
Observations	355,612	85,284	270,328

Note: The number in the table represents the marginal effects. Both the coefficients and standard errors are multiplied by 10,000 for convenience. To make the calculation of the marginal effects possible, year and industry dummies are excluded. The exclusion should impact only slightly in the results, since the difference in the estimated raw coefficients between regressions with and without the dummies is at the third digit after zero at most. Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

while regional trade agreement shows the greatest effect among intangible barrier variables (in terms of coefficient size, although not reported). Furthermore, value added does not have any impact on export, while corporate group affiliation and import experience have a strong negative impact, contradicting with OLS results.

However, when we evaluate the economic impact in table 3, the differentiated products have a greater distance impact on exports, which is in line with the network/search hypothesis. A 1% increase in distance reduces the export/total sales ratio by 1.09 basis point or almost twice compared to 0.63 basis point for homogeneous products. The magnitude of impact is also similar in regional trade agreement variable and corporate affiliations, i.e. 0.3 - 0.6 basis points for homogeneous and 0.8 - 1.1 basis points for differentiated products.

Joint decisions of export participation and export intensity

Instead of treating zero exports the same as export amount, they can be generated from a self-selection process. Hence, an alternative estimation method is required to deal with these zeros. I present the marginal effects from zero-inflated beta (ZOIB) regression first in table 4 and follow by Heckman selection results in table 5. For ZOIB regression, the dependent variable is export value per total sales and it is total export value for Heckman.

The inflation of zeros is estimated in columns 10-12 in table 4 and the intuition is the opposite to the interpretation for the rest of this paper. This means that a bigger GDP will have a negative impact on the firm to have zero export (instead of stating that it has a positive impact on the firm to participate in export).

Regarding the statistical inference, all the variables show the same significance but with opposite signs as the probit results with multinationals affiliation as the only one exception (see columns labelled "Probit" and "ZOIB - Export" in table 10). Referring to the estimated coefficient of the distance variable in table 11, the coefficient for homogeneous products is greater, but in table 3 we can see again that the economic impact is greater for differentiated products (0.024 compared to 0.022 percentage point), meaning an increase in distance will lead to a slightly more probability of zero export for differentiated products. Similar to probit results earlier, this difference between product groups is very small. In terms of other variables, the marginal effects calculated from ZOIB in general are slightly smaller in absolute value than probit.

For export intensity, the statistical signs and significance differ from QMLE in GDP per capita, contiguity, value added, and country experience (table 10). In terms of magnitude, QMLE results tend to have a larger coefficient size in all variables, both statistically and economically. The estimated coefficient for distance is higher for homogeneous products but the marginal effect is higher for differentiated products, similar to QMLE.

When we turn to Heckman results in table 5, the selection equation exhibits fundamentally similar results as the probit with only an addition of human capital variable.

For the export intensity, we can see from table 10 (labeled with "Heckman - Export Value") that the signs are similar to ZOIB in the country-specific (GDP and GDP per

Table 4:: Zero-inflated beta results

Variables	(10)	(11)	(12)	(13)	(14)	(15)
	Zero Inflate			ZOIB (Export/Total Sales)		
	All	Homog.	Diff.	All	Homog.	Diff.
GDP	-0.004***	-0.001*	-0.005***	0.176***	0.027**	0.232***
(log)	(0.000)	(0.001)	(0.000)	(0.012)	(0.011)	(0.016)
GDP Per Capita	-0.002***	-0.002**	-0.002***	0.082***	0.033*	0.100***
(log)	(0.001)	(0.001)	(0.001)	(0.018)	(0.020)	(0.023)
Distance	0.024***	0.022***	0.024***	-0.773***	-0.443***	-0.869***
(log)	(0.001)	(0.001)	(0.001)	(0.038)	(0.038)	(0.048)
Contiguity	-0.023***	-0.023***	-0.023***	0.277***	0.254***	0.250***
	(0.001)	(0.002)	(0.001)	(0.038)	(0.041)	(0.048)
Landlocked	0.010***	0.005*	0.012***	-0.335***	-0.121**	-0.413***
	(0.001)	(0.003)	(0.002)	(0.042)	(0.047)	(0.054)
English Dummy	-0.006***	0.001	-0.008***	0.350***	0.059	0.415***
	(0.002)	(0.003)	(0.002)	(0.062)	(0.066)	(0.079)
Regional Trade	0.019***	0.017***	0.019***	-0.587***	-0.331***	-0.613***
Agreement	(0.002)	(0.003)	(0.002)	(0.062)	(0.068)	(0.077)
Value Added	-0.005***	-0.002**	-0.006***	-0.068***	-0.026	-0.074**
(log)(lag)	(0.001)	(0.001)	(0.001)	(0.026)	(0.025)	(0.035)
Domestic Corporation	0.009***	0.004***	0.010***	-0.555***	-0.176***	-0.666***
	(0.001)	(0.001)	(0.001)	(0.028)	(0.025)	(0.037)
Multinationals	0.013***	0.007***	0.013***	-0.788***	-0.317***	-0.901***
	(0.001)	(0.001)	(0.001)	(0.035)	(0.031)	(0.045)
Import Dummy	0.003***	-0.001	0.003***	-0.201***	-0.098***	-0.193***
(lag)	(0.001)	(0.001)	(0.001)	(0.024)	(0.025)	(0.031)
Country Experience	-0.004***	-0.009***	-0.002**	0.062***	0.127***	0.025
	(0.001)	(0.001)	(0.001)	(0.022)	(0.023)	(0.028)
Observations	355,612	85,284	270,328	19,021	3,688	15,333

Note: The number in the table represents the marginal effects. Both the coefficients and standard errors in the ZOIB columns (13-15) are multiplied by 10,000 for convenience. To make the calculation of the marginal effects possible, year and industry dummies are excluded. The exclusion should impact only slightly in the results, since the difference in the estimated raw coefficients between regressions with and without the dummies is at the third digit after zero at most. Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5:: Heckman selection results

Variables	(16)	(17)	(18)	(19)	(20)	(21)
	Selection (Export Dummy)			Outcome (Export Value)		
	All	Homog.	Diff.	All	Homog.	Diff.
GDP	0.004***	0.001	0.005***	0.195***	0.084***	0.210***
(log)	(0.000)	(0.001)	(0.000)	(0.013)	(0.029)	(0.015)
GDP Per Capita	0.002***	0.003**	0.002***	0.177***	0.113**	0.191***
(log)	(0.001)	(0.001)	(0.001)	(0.023)	(0.053)	(0.026)
Distance	-0.023***	-0.022***	-0.023***	-0.270***	-0.347***	-0.266***
(log)	(0.001)	(0.001)	(0.001)	(0.031)	(0.076)	(0.034)
Contiguity	0.026***	0.027***	0.025***	-1.024***	-0.729***	-1.076***
	(0.001)	(0.002)	(0.001)	(0.042)	(0.103)	(0.046)
Landlocked	-0.011***	-0.006**	-0.012***	-0.483***	-0.800***	-0.411***
	(0.001)	(0.003)	(0.002)	(0.068)	(0.180)	(0.072)
English Dummy	0.006***	-0.001	0.007***	0.363***	0.640***	0.299***
	(0.001)	(0.003)	(0.002)	(0.060)	(0.171)	(0.063)
Regional Trade	-0.018***	-0.017***	-0.017***	0.127**	0.050	0.167**
Agreement	(0.002)	(0.003)	(0.002)	(0.060)	(0.139)	(0.066)
Value Added	0.005***	0.003**	0.006***	0.152***	0.157**	0.158***
(log)(lag)	(0.001)	(0.001)	(0.001)	(0.031)	(0.074)	(0.034)
Domestic Corporation	-0.010***	-0.005***	-0.011***	-0.035	0.275***	-0.079**
	(0.001)	(0.001)	(0.001)	(0.031)	(0.069)	(0.035)
Multinationals	-0.014***	-0.008***	-0.014***	0.462***	0.700***	0.438***
	(0.001)	(0.001)	(0.001)	(0.038)	(0.088)	(0.042)
Import Dummy	-0.002***	0.002	-0.003***	0.176***	0.097	0.229***
(lag)	(0.001)	(0.001)	(0.001)	(0.027)	(0.066)	(0.030)
Country Experience	0.004***	0.009***	0.002***	-0.020	0.066	-0.014
	(0.001)	(0.001)	(0.001)	(0.027)	(0.065)	(0.030)
Human Capital	-0.008**	-0.011	-0.010***	0.012	-0.015	0.019*
(lag)	(0.003)	(0.008)	(0.004)	(0.008)	(0.030)	(0.011)
Observations	355,596	85,277	270,319	19,021	3,688	15,333

Note: The number in the table represents the marginal effects. To make the calculation of the marginal effects possible, year and industry dummies are excluded. The exclusion should impact only slightly in the results, since the difference in the estimated raw coefficients between regressions with and without the dummies is at the third digit after zero at most. Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

capita), distance and barrier variables (distance, landlockedness, and so on) but not for firm-specific variables (corporate affiliation, import experience). Furthermore, the results suggest that differentiated products have a greater distance sensitivity in the intensity decision and it is positive and insignificant for homogeneous products. But economically, homogeneous products have a greater impact from distance than differentiated products in the intensity decision. According to table 5, a 1% change in distance will decrease export of homogeneous products by 34%, compared to 27% for differentiated products. The results for the other variables are similar to probit and OLS results.

In summary, GDP, distance, landlockedness, and English language dummy appear to be statistically robust across all the estimators, showing the expected in nearly all the regressions and samples. This implies that, regardless of product, a great value of trade will likely happen with a big and nearby destination country with access to sea and has English as an official language.

For the distance variable, homogeneous products appear to assert greater effect on exports in both the participation and intensity decisions when we use logged total export value as the dependent variable. This finding is similar to G.-J. Linders et al. (2005); Möhlmann et al. (2010) and Rauch (1999) before his adjustment on the transportability of differentiated products. However, we find the results that support the network/search hypothesis when we evaluate the economic significance of the regressions and when we use the ratio of export and total sales as the dependent variable.

The reason for this counterintuitive finding could be that homogeneous products are more standardised and competition from rivals close to destination market is fiercer than the more unique differentiated products. Once entered, producers of homogeneous products are more likely to ship in large quantity, so distance has a significant impact in determining how much to export. Having a look at the dataset, the average unit price, simply taken as value divided by weight, of differentiated products is 255.86% more expensive than those of homogeneous products but the weight of homogeneous products are heavier by 136.39%. This means that for homogeneous products, they are bulkier and producers are more likely to compete in price, whereas differentiated products are charged with a monopolistic price.

Trade experiences appear to have a pattern among all the estimators. Import experience shows a positive impact on export participation, with roughly the same magnitude as GDP, which suggests that an experience in a country is as important as the potential market size. Similarly, past export experience in a country is also a factor in the decision to participate. However, it is only for homogeneous products. This is mainly because of the characteristics of differentiated products. So that previous knowledge is not applicable for an introduction of a new differentiated product even to the same market.

Robustness check

For robustness check, I run several specifications of the gravity model on the full and sub-samples in all of the estimators used in this paper. First, I run all the estimators using the full sample but with only one distance variable while excluding

other intangible barrier variables. This is to see the overall effect of distance prior to accounting for the intangible barriers. Second, the sample is reduced to include only small- and medium-sized firms (SMEs), which has 1-50 employees. Lastly, I rerun the estimation on non-affiliated firms. Because small and non-affiliated firms often have limited resources, excluding large firms away will give a more precise look into the effects from distance. The distance coefficients from all regressions are summarised in Table 11 in the Appendix.

The conclusions of greater distance effects for homogeneous products hold for all model specifications in both export decisions with only an exception of Heckman results for export intensity, in which homogeneous products do not have a significant impact. When the intangible barrier variables are excluded, distance coefficient is lower in all of the regressions, except a positive and significant estimate for homogeneous products in Heckman.

Because the majority of firms is small, the SMEs sample only slightly decrease coefficient estimates in nearly all regressions. Similar pattern is also observed in non-affiliated firms sample. This suggests that distance effects are exaggerated by the inclusion of large firms. One possible explanation is that large corporated firms require a more careful consideration among the managers and directors within the companies before a decisive action can be executed. A greater extent of committed resources can then follow such a decision.

Conclusion

Although distance plays an important role in firm's export decisions, we still have not yet fully understood the mechanism behind its impact. In this paper I look at the distance sensitivity on firm's export participation and export intensity by different product groups, namely homogeneous versus differentiated products. Using the detailed firm-product level data of Sweden, the investigation deals with actual decisions made by each individual firm for each exported product. The findings are in contrast with the network/search view. Homogeneous products exhibit greater distance sensitivity in both export participation and export intensity. This finding can be attributed to the competition of standardised products from different producers nearby the destination market. Only when the economic impact is evaluated that we see in some model specifications the results that support the hypothesis.

Past experiences are more important for homogeneous products in terms of participation. But, for export intensity, only import experience matters. This suggests a learning effect.

References

Aitken, N. D. (1973). The effect of the eec and efta on european trade: A temporal cross-section analysis. *The American Economic Review*, *63*(5), pp. 881-892. Retrieved from `http://www.jstor.org/stable/1813911`

Anderson, J. E. (1979). A theoretical foundation for the gravity equation. *The American Economic Review*, *69*(1), 106-116.

Anderson, J. E., & Marcouiller, D. (2002). Insecurity and the pattern of trade: An empirical investigation. *Review of Economics and Statistics*, *84*(2), 342-352.

Anderson, J. E., & Wincoop, E. v. (2003). Gravity with gravitas: A solution to the border puzzle. *The American Economic Review*, *93*(1), 170-192.

Andersson, M. (2007, February). *Entry costs and adjustments on the extensive margin - an analysis of how familiarity breeds exports* (Working Paper Series in Economics and Institutions of Innovation No. 81). Royal Institute of Technology, CESIS - Centre of Excellence for Science and Innovation Studies. Retrieved from http://ideas.repec.org/p/hhs/cesisp/0081.html

Arkolakis, C. (2008, August). *Market penetration costs and the new consumers margin in international trade* (Working Paper No. 14214). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w14214

Baum, C. F. (2008). Stata tip 63: Modeling proportions. *Stata Journal*, *8*(2), 299.

Bergstrand, J. H. (1985). The gravity equation in international trade: Some microeconomic foundations and empirical evidence. *The Review of Economics and Statistics*, *67*(3), 474-481.

Bernard, A. B., & Jensen, J. B. (2004). Why some firms export. *The Review of Economics and Statistics*, *86*(2), 561-569.

Carrère, C. (2006). Revisiting the effects of regional trade agreements on trade flows with proper specification of the gravity model. *European Economic Review*, *50*(2), 223 - 247. Retrieved from http://www.sciencedirect.com/science/article/pii/S0014292104000583 doi: 10.1016/j.euroecorev.2004.06.001

Clerides, S. K., Lach, S., & Tybout, J. R. (1998). Is learning by exporting important? micro-dynamic evidence from colombia, mexico, and morocco. *The Quarterly Journal of Economics*, *113*(3), 903-947. doi: 10.1162/003355398555784

Cook, D. O., Kieschnick, R., & McCullough, B. (2008). Regression analysis of proportions in finance with self selection. *Journal of Empirical Finance*, *15*(5), 860 - 867. Retrieved from http://www.sciencedirect.com/science/article/pii/S0927539808000145 doi: http://dx.doi.org/10.1016/j.jempfin.2008.02.001

Disdier, A.-C., & Head, K. (2008). The puzzling persistence of the distance effect on bilateral trade. *Review of Economics and Statistics*, *90*(1), 37-48.

Egger, P., & Pfaffermayr, M. (2011, Jun). *Structural estimation of gravity models with path-dependent market entry* (Tech. Rep.). C.E.P.R. Discussion Papers. (CEPR Discussion Papers)

Feenstra, R. C. (2002). Border effects and the gravity equation: Consistent methods for estimation. *Scottish Journal of Political Economy*, *49*(5), 491–506. Retrieved from http://dx.doi.org/10.1111/1467-9485.00244 doi: 10.1111/1467-9485.00244

Grossman, G., & Helpman, E. (1993). *Innovation and growth in the global economy.* MIT Press. Retrieved from http://www.google.se/books?id=4ikgmM2vLJ0C

Grossman, G. M. (1998). Determinants of bilateral trade: Comment. In J. A. Frankel (Ed.), *The regionalization of the world economy* (p. 29-31). Chicago: University of Chicago Press for the NBER.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, *47*(1), pp. 153-161. Retrieved from http://www.jstor.org/stable/1912352

Helpman, E. (1984). Chapter 7 increasing returns, imperfect markets, and trade theory. In W. J. Ronald & B. K. Peter (Eds.), *Handbook of international economics* (Vol. Volume 1, p. 325-365). Elsevier.

Huang, R. R. (2007). Distance and trade: Disentangling unfamiliarity effects and transport cost effects. *European Economic Review*, *51*(1), 161 - 181. Retrieved from http://

www.sciencedirect.com/science/article/pii/S0014292105001455 doi: 10.1016/j.euroecorev.2005.11.004

Kasahara, H., & Rodrigue, J. (2008). Does the use of imported intermediates increase productivity? plant-level evidence. *Journal of Development Economics*, *87*(1), 106 - 118. Retrieved from http://www.sciencedirect.com/science/article/pii/S0304387808000035 doi: 10.1016/j.jdeveco.2007.12.008

Koenig, P., Mayneris, F., & Poncet, S. (2010). Local export spillovers in france. *European Economic Review*, *54*(4), 622-641.

Lankhuizen, M., de Graaff, T., & de Groot, H. L. F. (2012, July). *Product heterogeneity, intangible barriers and distance decay: The effect of multiple dimensions of distance on trade across different product categories* (Working Paper). Tinbergen Institute.

Lawless, K., Martina; Whelan. (2008). *Where do firms export, how much, and why?* (Tech. Rep.). University College Dublin. (UCD Centre for Economic Research Working Paper Series)

Lawless, M. (2010). Deconstructing gravity: trade costs and extensive and intensive margins. *Canadian Journal of Economics/Revue canadienne d'conomique*, *43*(4), 1149-1172. Retrieved from http://dx.doi.org/10.1111/j.1540-5982.2010.01609.x doi: 10.1111/j.1540-5982.2010.01609.x

Leamer, E. E., & Levinsohn, J. (1995). International trade theory: The evidence. In M. G. Gene & R. Kenneth (Eds.), *Handbook of international economics* (Vol. 3, p. 1339-1394). Elsevier.

Linders, G.-J., De Groot, H., & Rietveld, P. (2005). Institutional determinants of bilateral trade: an analysis according to product type.

Linders, G.-J., & De Groot, H. L. (2006). Estimation of the gravity equation in the presence of zero flows. *SSRN eLibrary*.

Linders, G.-J. M. (2006). *Intangible barriers to trade: the impact of institutions, culture and distance on patterns of trade*. Unpublished doctoral dissertation, Vrije Universiteit Amsterdam.

Martin, W., & Pham, C. S. (2008). *Estimating the gravity equation when zero trade flows are frequent* (MPRA Working Paper). Munich RePEc Personal Archive [http://mpra.ub.uni-muenchen.de/perl/oai2] (Germany). Retrieved from http://mpra.ub.uni-muenchen.de/9453/1/gravity_june_08.pdf

Martínez-Zarzoso, I. (2013). The log of gravity revisited. *Applied Economics*, *45*(3), 311-327. doi: 10.1080/00036846.2011.599786

Mátyás, L. (1997). Proper econometric specification of the gravity model. *World Economy*, *20*(3), 363–368. Retrieved from http://dx.doi.org/10.1111/1467-9701.00074 doi: 10.1111/1467-9701.00074

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, *71*(6), pp. 1695-1725. Retrieved from http://www.jstor.org/stable/1555536

Möhlmann, L. J., Ederveen, S., de Groot, H. L. F., & Linders, G.-J. M. (2010). Intangible barriers to international trade. In P. van Bergeijk & S. Brakman (Eds.), *The gravity model in international trade: Advances and applications* (p. 224-252). Cambridge University Press. Retrieved from http://books.google.se/books?id=vQSwRPuMoZMC

Papke, L. E., & Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics*, *11*(6), 619–632. Retrieved from http://dx.doi.org/10.1002/(SICI)1099-1255(199611)11:6<619::AID-JAE418>3.0.CO;2-1 doi: 10.1002/(SICI)1099-1255(199611)11:6⟨619::AID-JAE418⟩3.0.CO;2-1

Persson, M. (2013). Trade facilitation and the extensive margin. *The Journal of International Trade & Economic Development*, *22*(5), 658-693. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/09638199.2011.587019 doi: 10.1080/09638199.2011.587019

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of International Economics*, *48*(1), 7 - 35. Retrieved from http://www.sciencedirect.com/science/article/pii/S0022199698000099 doi: 10.1016/S0022-1996(98)00009-9

Redding, S., & Venables, A. J. (2004). Economic geography and international inequality. *Journal of International Economics*, *62*(1), 53 - 82.

Santos Silva, J., & Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, *88*, pp. 641-658.

Segura-Cayuela, R., & Vilarrubia, J. (2008). *Uncertainty and entry into export markets* (Working Papers No. 0811). Banco de España. Retrieved from http://ssrn.com/abstract=1144564

Tinbergen, J. (1962). *Shaping the world economy; suggestions for an international economic policy*. Twentieth Century Fund, New York. Retrieved from http://hdl.handle.net/1765/16826

Verbeek, M. (2008). *A guide to modern econometrics*. John Wiley & Sons. Retrieved from http://books.google.se/books?id=uEFm6pAJZhoC

Verhoogen, E. A. (2008). Trade, quality upgrading, and wage inequality in the mexican manufacturing sector. *The Quarterly Journal of Economics*, *123*(2), 489-530. Retrieved from http://qje.oxfordjournals.org/content/123/2/489.abstract doi: 10.1162/qjec.2008.123.2.489

Vogel, A., & Wagner, J. (2010). Higher productivity in importing german manufacturing firms: self-selection, learning from importing, or both? *Review of World Economics*, *145*, 641-665. Retrieved from http://dx.doi.org/10.1007/s10290-009-0031-4 doi: 10.1007/s10290-009-0031-4

Wagner, J. (2001). A note on the firm size - export relationship. *Small Business Economics*, *17*(4), 229-237. doi: 10.1023/A:1012202405889

Appendix

Variable description

Variable	Description	Source	Exp. Sign
Export Dummy	Dummy taking value of 1 if the firm has positive export of a product to destination country at current year, 0 otherwise	Author-generated	
Export Value	Total amount exported in current Swedish krona	Statistics Sweden	
Export Per Sales	A ratio of firm's export value and total sales	Statistics Sweden	
GDP	Gross Domestic Product of destination country in current USD (in log).	CEPII	+
GDP per capita	GDP per capita of destination country (in log).	CEPII	+
Distance	Weighted distance as measured in km. from Sweden, calculated using great circle distance between major cities as weight (in log).	CEPII	-
Contiguity	Dummy taking value of 1 if the destination country shares border with Sweden and 0 otherwise.	CEPII	+
Landlocked	Dummy taking value of 1 if the destination country does not have coastal line.	CEPII	-
English Dummy	Dummy taking value of 1 if one of the official languages in the destination country is English.	CEPII	+
Regional Trade Agreement	Dummy taking value of 1 if the regional trade agreement is in effect.	CEPII	+
Value-Added	Firm's value-added per employee in SEK (in log and lagged 1 year).	Statistics Sweden	+
Human Capital	Fraction of employees graduated at university level (lagged 1 year).	Statistics Sweden	+/-
Domestic Corporation	Dummy taking value of 1 if the firm belongs to a domestic corporation group	Statistics Sweden	+
Multinationals	Dummy taking value of 1 if the firm belongs to a multinational enterprise	Statistics Sweden	+
Import Dummy	Dummy taking value of 1 if the firm import from destination country a year before and 0 otherwise.	Author-generated	+
Export Country Experience	Dummy taking value of 1 if the firm already exported to the country previously.	Author-generated	+

Descriptive statistics

Variables	Obs.	Mean	Std. Dev.	Min	Max
Export Dummy	395,900	0.05	0.22	0	1
Export Value	20,814	675,424.80	2,479,948	1	125,036,227
Export Per Sales	395,900	0.0001	0.002	0	0.80
GDP	395,361	910,098.90	2,149,115	367.2	13,201,819
GDP per capita	395,202	24,864.0	17,153.82	84.56	89,563.63
Distance	395,900	2,531.94	3,196.99	450.08	17,389.62
Contiguity Dummy	395,900	0.25	0.43	0	1
Landlocked Dummy	395,900	0.09	0.28	0	1
English Dummy	395,900	0.15	0.36	0	1
RTA Dummy	395,900	0.73	0.45	0	1
Value-Added	395,900	30,120.62	151,299.30	3	5,593,307
Human Capital	395,900	0.06	0.11	0	1
Domestic Corp.	395,900	0.31	0.46	0	1
Multinational Corp.	395,900	0.32	0.47	0	1
Import Dummy	395,900	0.63	0.48	0	1
Country Experience	395,900	0.56	0.50	0	1

The mean value for dummy variables indicates the percentage of 1's.

Participation of Swedish Exports

SNI	Industry	Total Producers	Exporters* (%)	Exported** (%)	Average Products	Average Destinations
15	Food products; beverages and tobacco	1296	18.9	17.57	11.08	6.97
16	Tobacco products	3	33.33	3.58	23.82	29.82
17	Textiles and textile products	380	41.84	18.58	12.16	8.10
18	Wearing apparel; dressing and dyeing of fur	102	51.96	26.6	32.83	7.57
19	Leather; luggage, handbags, and footwear	65	58.46	19.98	8.23	5.71
20	Wood and wood products except furniture	1540	31.75	25.32 4.94	5.23	
21	Pulp, paper and paper products	218	78.44	31.96	11.55	17.86
22	Publishing, printing and reproduction of recorded media	1958	18.74	5.03	4.75	4.36
23	Coke, refined petroleum products and nuclear fuel	16	56.25	49.21	14.70	11.29
24	Chemicals, chemical products and man-made fibres	308	75	32.24	20.74	16.37
25	Rubber and plastic products	718	58.91	23.15	9.17	9.05
26	Other non-metallic mineral products	401	39.9	18.14	10.07	8.34
27	Basic metals	226	64.6	35.07	15.98	15.92
28	Fabricated metal products except machinery	4272	27.88	16.04	6.70	5.98
29	Machinery and equipment n.e.c.	2069	48.53	29.55	13.79	13.41
30	Office machinery and computers	90	36.67	34.4	13.31	14.15
31	Electrical machinery and apparatus n.e.c.	527	49.91	21.03	12.62	10.27
32	Radio, television and communication equipment and apparatus	192	45.83	31.05	15.47	10.29
33	Medical, precision and optical instruments, watches and clocks	747	37.88	36.77	15.53	16.97
34	Motor vehicles, trailers and semi-trailers	366	59.56	26.15	20.76	9.42
35	Other transport equipment	353	36.26	28.54	15.17	7.35
36	Furniture; manufacturing n.e.c.	859	45.52	18.27	8.52	6.88
	Average	759	46.19	24.92		

* Exporters' share of total number of producers.
** Average share of exports per total firm's sales.

Country list

ISO2	Country Name	Distance*	ISO2	Country Name	Distance*
AE	United Arab Emirates	4,859.49	DK	Denmark	450.08
AF	Afghanistan	4,644.21	DO	Dominican Republic	8,006.54
AL	Albania	1,995.41	DZ	Algeria	2,709.28
AM	Armenia	2,899.19	EC	Ecuador	10,457.59
AN	Netherland Antilles	8,441.07	EE	Estonia	595.36
AO	Angola	7,644.17	EG	Egypt	3,412.79
AR	Argentina	12,404.68	ER	Eritrea	5,250.37
AT	Austria	1,228.47	ES	Spain	2,486.55
AU	Australia	15,385.40	ET	Ethiopia	5,847.94
AW	Aruba	8,587.53	FI	Finland	604.91
BA	Bosnia & Herzegovina	1,644.60	FJ	Fiji	15,252.19
BB	Barbados	7,930.83	FO	Faroe Islands	1,303.04
BD	Bangladesh	6,912.31	FR	France	1,616.32
BE	Belgium	1,151.50	GA	Gabon	6,577.58
BF	Burkina Faso	5,408.34	GB	United Kingdom	1,292.80
BG	Bulgaria	1,912.32	GE	Georgia	2,708.50
BH	Bahrain	4,526.21	GH	Ghana	6,005.78
BI	Burundi	7,027.18	GI	Gibraltar	2,956.84
BJ	Benin	5,803.46	GL	Greenland	3,368.65
BM	Bermuda	6,456.30	GM	Gambia	5,712.82
BN	Brunei Darussalam	10,069.25	GN	Guinea	5,966.61
BO	Bolivia	11,201.18	GR	Greece	2,353.03
BR	Brazil	10,185.49	GT	Guatemala	9,539.39
BS	Bahamas	7,808.63	HK	Hong Kong	8,368.68
BW	Botswana	9,199.48	HN	Honduras	9,338.07
BY	Belarus	986.48	HR	Croatia	1,519.27
CA	Canada	6,347.80	HT	Haiti	8,142.33
CG	Congo	7,007.02	HU	Hungary	1,315.38
CH	Switzerland	1,422.90	ID	Indonesia	10,632.05
CI	Cte d'Ivoire	6,129.18	IE	Ireland	1,549.43
CL	Chile	12,956.19	IL	Israel	3,315.60
CM	Cameroon	5,907.75	IN	India	6,308.11
CN	China	7,276.97	IQ	Iraq	3,552.56
CO	Colombia	9,491.13	IR	Iran	3,765.08
CR	Costa Rica	9,629.91	IS	Iceland	2,047.33
CU	Cuba	8,246.69	IT	Italy	1,833.43
CV	Cape Verde	5,794.42	JM	Jamaica	8,463.56
CY	Cyprus	2,955.68	JO	Jordan	3,358.22
CZ	Czech Republic	1,009.36	JP	Japan	8,226.76
DE	Germany	929.32	KE	Kenya	6,957.80
KH	Cambodia	8,820.19	PL	Poland	848.39
KP	North Korea	7,371.20	PT	Portugal	2,821.62
KR	South Korea	7,682.77	PY	Paraguay	11,477.31
KW	Kuwait	4,107.62	QA	Qatar	4,653.14
KY	Cayman Islands	8,589.82	RW	Rwanda	6,884.48
KZ	Kazakstan	3,774.62	SA	Saudi Arabia	4,479.74
LB	Lebanon	3,148.39	SD	Sudan	5,100.44
LC	Saint Lucia	7,928.13	SG	Singapore	9,782.64
LK	Sri Lanka	7,849.86	SI	Slovenia	1,420.52

ISO2	Country Name	Distance*	ISO2	Country Name	Distance*
LT	Lithuania	676.56	SK	Slovakia	1,176.30
LU	Luxembourg	1,207.73	SL	Sierra Leone	6,101.36
LV	Latvia	591.22	SM	San Marino	1,678.00
LY	Libya	2,993.48	SN	Senegal	5,613.46
MA	Morocco	3,274.22	SO	Somalia	6,638.56
MD	Moldova, Rep.of	1,580.09	SR	Suriname	8,366.51
MG	Madagascar	9,152.54	SV	El Salvador	9,548.48
MH	Marshall Islands	12,283.25	SY	Syrian Arab Republic	3,084.28
MK	Macedonia	1,950.69	TC	Turks & Caicos Is.	7,815.33
MO	Macau (Aomen)	8,201.04	TG	Togo	5,878.81
MT	Malta	2,558.88	TH	Thailand	8,415.42
MU	Mauritius	9,593.82	TJ	Tajikistan	4,346.91
MV	Maldives	7,861.62	TK	Tokelau	14,475.37
MW	Malawi	8,326.36	TN	Tunisia	2,582.25
MX	Mexico	9,357.39	TO	Tonga	15,710.15
MY	Malaysia	9,568.98	TR	Turkey	2,453.42
MZ	Mozambique	9,058.94	TT	Trinidad & Tobago	8,286.25
NA	Namibia	8,993.66	TW	Taiwan	8,551.70
NC	New Caledonia	15,294.21	TZ	Tanzania	7,468.98
NE	Niger	5,062.04	UA	Ukraine	1,616.60
NG	Nigeria	5,721.76	UG	Uganda	6,634.94
NI	Nicaragua	9,522.18	US	U.S.A.	7,440.51
NL	Netherlands	1,009.40	UY	Uruguay	12,286.37
NO	Norway	502.69	UZ	Uzbekistan	4,141.06
NP	Nepal	6,223.75	VC	St Vincent	8,018.46
NZ	New Zealand	17,389.62	VE	Venezuela	8,692.38
OM	Oman	5,162.00	VG	British Virgin Is.	7,718.33
PA	Panama	9,511.23	VN	Viet Nam	8,727.68
PE	Peru	11,219.56	YE	Yemen	5,474.30
PF	French Polynesia	15,277.91	YU	Serbia & Montenegro	1,686.69
PH	Philippines	9,639.51	ZA	South Africa	9,838.57
PK	Pakistan	5,294.92	ZM	Zambia	8,207.19
RO	Romania	1,640.88	ZW	Zimbabwe	8,722.59
RU	Russian Federation	2,081.84		**Total countries**	**165**

* Great-circle distance measured as km. from Sweden with major cities' population as weight.

Table 10:: Summary of signs and magnitude comparison for the estimated coefficients

Estimator	Probit	OLS	QMLE		ZOIB	Heckman	
Dependent variable	Export	Export Value	Export Value/Sales	Export*	Export Value/Sales	Export	Export Value
GDP	+ / D	+ / D	+ / D	- / D	+ / D	+ / D	+ / D
GDP Per Capita	0	+ / D	- / H	0	+ / D	0	+ / D
Distance	- / H	- / H	- / H	+ / H	- / H	- / H	- / D
Contiguity	+ / H	- / D	0	- / H	- / D	+ / H	- / H
Landlocked	- / D	- / H	- / D	+ / D	- / D	- / D	- / D
English Dummy	+ / D	+ / H	+ / D	- / D	+ / H	+ / D	+ / H
Regional Trade Agreement	- / H	+ / D	- / H	+ / H	- / H	- / H	0
Value Added	+ / D	+ / D	0	- / D	- / D	+ / D	+ / D
Domestic Corporation	- / D	0	- / H	+ / D	- / H	- / D	+ / H
Multinationals	- / H	+ / D	- / D	- / H	- / D	- / H	+ / H
Import Dummy	+ / D	+ / D	- / H	- / H	- / H	+ / H	+ / D
Country Experience	+ / H	0	+ / H	- / H	0	+ / H	0
Human Capital						- / H	

Note: The coefficient signs are represented as + (positive and significant at $p < 0.1$ level or less in at least one product type), - (negative and significant), 0 (not statistically significant for all product types). The product type that has a greater magnitude in absolute value is either H for homogeneous or D for differentiated. * The interpretation of the sign is the opposite of the rest of the table.

Table 11:: Distance coefficients from all model specifications

Model Specification	(22) All	(23) Homog.	(24) Diff.	(25) All	(26) Homog.	(27) Diff.
			Main sample			
Baseline Probit & OLS	-0.270***	-0.348***	-0.265***	-0.270***	-0.348***	-0.265***
with no dummies	(0.031)	(0.076)	(0.034)	(0.031)	(0.076)	(0.034)
Probit & OLS	-0.240***	-0.304***	-0.231***	-0.249***	-0.374***	-0.251***
with Industry dummies	(0.008)	(0.019)	(0.009)	(0.031)	(0.078)	(0.034)
Probit & OLS	-0.242***	-0.304***	-0.230***	-0.276***	-0.372***	-0.269***
with Year dummies	(0.008)	(0.019)	(0.009)	(0.031)	(0.076)	(0.034)
Probit & OLS	-0.244***	-0.314***	-0.234***	-0.249***	-0.400***	-0.248***
with both dummies	(0.008)	(0.019)	(0.009)	(0.031)	(0.078)	(0.034)
QMLE				-1.028***	-1.294***	-0.997***
with both dummies				(0.060)	(0.160)	(0.065)
Zero-inflated Beta	0.560***	0.756***	0.529***	-0.121***	-0.153***	-0.125***
with both dummies	(0.019)	(0.049)	(0.021)	(0.011)	(0.028)	(0.013)
Heckman	-0.243***	-0.314***	-0.233***	-0.324***	0.031	-0.365***
with both dummies	(0.008)	(0.020)	(0.009)	(0.040)	(0.151)	(0.046)
			Main sample: one Distance variable			
Probit & OLS	-0.233***	-0.332***	-0.221***	-0.055***	-0.177***	-0.071***
with both dummies	(0.006)	(0.015)	(0.006)	(0.020)	(0.059)	(0.022)
QMLE				-0.611***	-0.884***	-0.594***
with both dummies				(0.042)	(0.122)	(0.044)
Zero-inflated Beta	0.529***	0.789***	0.493***	-0.025***	-0.058***	-0.032***
with both dummies	(0.014)	(0.036)	(0.015)	(0.008)	(0.021)	(0.009)
Heckman	-0.232***	-0.331***	-0.219***	-0.182***	0.304**	-0.239***
with both dummies	(0.006)	(0.015)	(0.006)	(0.049)	(0.146)	(0.045)
			SMEs sample: 76.4% of total observations			
Probit & OLS	-0.228***	-0.277***	-0.222***	-0.252***	-0.288***	-0.265***
with both dummies	(0.009)	(0.023)	(0.010)	(0.033)	(0.091)	(0.035)
QMLE				-1.009***	-1.135***	-0.989***
with both dummies				(0.062)	(0.161)	(0.068)
Zero-inflated Beta	0.524***	0.656***	0.505***	-0.117***	-0.118***	-0.121***
with both dummies	(0.021)	(0.056)	(0.023)	(0.013)	(0.035)	(0.014)
Heckman	-0.227***	-0.314***	-0.233***	-0.343***	0.031	-0.365***
with both dummies	(0.009)	(0.020)	(0.009)	(0.040)	(0.151)	(0.046)
			Non-affiliated firms sample: 37.4% of total observations			
Probit & OLS	-0.221***	-0.324***	-0.204***	-0.266***	-0.468***	-0.267***
with both dummies	(0.014)	(0.037)	(0.015)	(0.045)	(0.151)	(0.047)
QMLE				-0.949***	-1.171***	-0.939***
with both dummies				(0.082)	(0.273)	(0.087)
Zero-inflated Beta	0.501***	0.762***	0.462***	-0.115***	-0.213***	-0.111***
with both dummies	(0.031)	(0.091)	(0.033)	(0.018)	(0.058)	(0.019)
Heckman	-0.247***	-0.315***	-0.237***	-0.338***	0.013	-0.386***
with both dummies	(0.008)	(0.020)	(0.009)	(0.058)	(0.162)	(0.058)

Note: The independent variables for the "Main sample" and "SMEs sample" regressions include GDP, GDP per capita, Distance, Contiguity, Landlocked, English, Regional Trade Agreement, Value added, Domestic Corporation, Multinationals, Import Dummy, and Country Experience variables. The "Main sample: one Distance" regressions exclude Contiguity, Landlocked, English, and Regional Trade Agreement. The "Non-affiliated firms sample" exclude Domestic Corporation and Multinationals variables. There is no distinction between export participation and export intensity decisions in QMLE regressions but the results are listed in the export intensity columns for convenience. Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$