Building a new export network: the case of Lithuania

Vaiva Petrylė[a†] Marco J. van der Leij[b‡]

13 November 2020

[a] *Vilnius University, Department of Economics and Business Administration*
[b] *European Novitiate, Congregation of the Blessed Sacrament; University of Amsterdam, CeNDEF*

Abstract

In this paper we empirically analyze the formation of the export network of Lithuania. The analysis builds on a natural experiment of the collapse of a Soviet bloc when Lithuania among a number of other countries entered the international trade market. The main research questions are which factors were the most important in the spread of the Lithuania's export network and how the influence of these factors changed throughout the years. Our results suggest that network effects, especially those of the distance, economic development of the destinations and common spoken languages, are important for the development of the extensive trade margin. We also find that during the first years of opening to the foreign trade cultural and language factors were the most important. In later years we see the domination of the economic factors as well as the increase of the importance of friends-of-friend search and clustering of the destination markets.

Keywords: trade networks, network formation, extensive trade margin.

JEL Classification: F14, D85

* Vaiva Petrylė is especially grateful to Franc Klaassen for his valuable advices, University of Amsterdam, especially CeNDEF, for hosting and showing how a real research is being made, and all the seminar participants at University of Amsterdam for helpful discussions.
† vaiva.petryle@evaf.vu.lt
‡ mvanderleij@gmail.com

I Introduction

In the small world we have today countries have formed complex trade relations with each other. We see elaborated trade networks where every country exports different goods to almost every other country. Analyzing the existing trade networks gives us information about the resources that different countries lack or that are abundant, countries' competitiveness as well as their growth limits and perspectives. According to Hidalgo and Hausmann (2009), the level of a country's economic development and its growth perspectives highly depends on its export structure: the products it sells and the markets it serves. The structure of its trade network and its intersectoral relations also affects a country's volatility to idiosyncratic shocks (Acemoglu et al., 2012).

Although both extensive and intensive trade margins have always interested economists, the empirics of extensive network effects for the expansion of international trade has not been thoroughly analysed. We are still hardly able to determine the factors that played their parts in the formation of such a trade network. Hence, the aim of the paper is to empirically examine the importance of network effects for the development of the extensive trade margin.

The notion of the network effects in the paper comes from the concept of networks in trade (e.g. Chaney, 2016). Following Chaney, 2014 and Chaney, 2016 we assume that the country's export network develops gradually. In the first stage the country of origin finds a couple of export destinations directly. In the second and other stages the country of origin finds more export destinations in two ways: directly from the origin, and indirectly – from its current export destinations which had already been found in the previous stages. We call the factors that affect this indirect search procedure "network effects".

In this paper we take Lithuania – a newly opening market which is as close to the beginning of its network formation process as possible, and empirically analyze which factors are the most important for the development of its' extensive trade margin. The analysis builds on a natural experiment of the collapse of a Soviet bloc when a number of countries suddenly entered the international trade market and started to form their own trade network. The interest of this paper is: which factors were the most important in the spread of the export network of such a country, and how did the influence of these factors change throughout the years? As all the ex-Soviet bloc countries are similar in their trade patterns, due to the computational feasibility we decided to focus on just one small open economy – Lithuania.

Lithuania was chosen because of a couple of reasons. First, before 1990 Lithuania was not a separate country, but incorporated in the USSR. Hence, being a completely new market after declaring its independence, Lithuania stands closer to the headwaters of its trade network, than those countries that, although belonging to the Soviet bloc and controlled by the USSR, always had been separate entities. Second, Lithuania is a developed and democratic Eastern European country which has reliable statistical data since 1995.

In 1990 when the Soviet Union collapsed, Lithuania once again became a player in the international market and started creating its own trade relations. Before 1990, Lithuania was a part of the Soviet Union and no free trade was possible. The country's economy was centrally planned and the government institutions directed which goods and what quantities should be manufactured, as well as for what prices and were they should be exported. After 1990 Lithuania's market was liberalised and regular foreign trade became possible. These facts allow the researchers to build on a natural experiment of trade network formation and to analyse

the formation process of Lithuania's export's network almost from the very beginning of its development. The paper empirically examines the factors that influence the formation of Lithuania's export's network.

In particular, we aim to answer the following research questions:

1. Which factors were the most important in the spread of the export's network of Lithuania?
2. How did the influence of these factors change throughout the years?

This paper contributes to the literature on international trade networks. Current trade literature could be classified in research on gravity modelling, trade networks and research based on some specific aspects of international trade.

Gravity modelling. The main contemporary articles on gravity modelling are Anderson (2011) and Anderson and Winkoop (2003). Rauch (1999) constructs gravity models to estimate which factors matter more for the trade of different types of commodities. He finds that country's proximity, common language and common colony are more important for differentiated goods; also, search barriers affect trade of differentiated goods more than that of homogeneous goods. Helpman et al. (2008) propose a two-stage procedure of gravity modelling. In the first stage a Probit model for extensive trade margin is estimated, and in the second stage predicted components of the first stage equation are used to estimate a log-linear gravity model.

International Trade. Research based on specific aspects of international trade is also extensive. Felbermayer and Kohler (2006) examine the effect of distance on trade. Melitz (2018) analyses the determinants of exporters' mark-ups and finds that mark-ups are higher if the market size of an export destination is larger and has stronger competition. Allen (2014) claims that trade reduction with distance could be explained not only by transport costs, but also by search costs due to information frictions. Broda and Weinstein (2006) show that the greater variety in imported products the greater the welfare of the importing country.

Trade networks. Finally, there is some literature on networks in trade. Silva et al. (2014) analyse the extensive trade margin by examining the determinants of the number of sectors exporting to other countries. Albornoz et al. (2012) examine Argentinian exporters and show that trade liberalisation in a country can promote entry not only to this country, but also to the other countries because of the spread of network relations. Chaney (2014) analyses the spread of French firm-level exports. Rauch and Trindade (2002) use gravity modelling to estimate the effects on Chinese social networks on international trade. Konrad et al. (2013) focus on the effect of social relations between East and West Germany on economic growth of separate German regions after the fall of the Berlin Wall.

This paper attempts to combine gravity models with network formation theory. It relates mostly to the articles of Chaney (2014) and Helpman et al. (2008), but will differ from the above literature in 3 main aspects:

1. It will contribute to the trade and gravity literature by examining a large number of network effects (distance, common culture, language, trade unions, emigration network, etc.) on the extensive margin of foreign trade.
2. It will focus on the new entrants to the international trade market and analyse the very origins of countries' export network formation (current empirical research is based mainly on matured markets (e. g. the USA, France, China, etc.)).
3. We use a much larger dataset than has typically been used in the literature of social networks.

The idea of the paper is to apply the gravity and networks theory and empirically test a number of factors (not only the geographical distance) that could influence trade expansion, to determine the most important ones and to disclose the dynamics of their effects over time.

Our results show that network variables are important for the development of an extensive trade margin. We find that the most important factors for the spread of the export network of Lithuania are: previous exporting to the destination market, exporting to more markets, closer distance and ability to communicate with the destination market (either directly or through the intermediate markets). Having emigrants either in the destination countries or in the other intermediate markets also seemed to have driven exports. Lithuania was also more tempted to start exporting to economically stronger countries. Moreover, economic factors not only showed to be more and more important as the years passed, but could be one of the main reasons to redirect Lithuania's exports from the former USSR countries to the Western Europe in the early stages of the trade creation.

Surprisingly, indirect distance, that is the sum of the distances between the current export destinations of an origin and their export partners, seemed to be more important than direct distance between the origin and its destination. Another surprise is that the membership in various trade organisations (i. e. the EU, the Soviet bloc, the free trade) appeared to have been insignificant for the spread of an extensive trade margin.

The setup of the paper is as follows: in Section 2 we present the methodology and the main model of our analysis, Section 3 gives an insight to the data and the variables, Section 4 is intended for the descriptive statistics, Section 5 presents the results of our main model, in Section 6 we present yearly changes of the direct and network effects, Section 7 gives robustness checks and finally Section 8 concludes.

2 Methodology

The background network model is the one developed by Chaney (2014). In the model there is a discrete set S of locations x (countries) each having a finite set of firms. The number of firms grows (by birth) at each location at a constant rate γ, they never change their initial location and never die. Firms sell their production to firms at other locations. Each period a firm meets new consumers in 2 ways (based on the model of Jackson and Rogers (2007)):

1 Directly at random. The firm searches for $\gamma\mu$ new consumers from its original location, choosing the destination (location) of every new customer randomly according to the probability function $g(0,x)$. The probability function depends on the distance between the destination and the firm's origin (0) and on the size of the destination (x).

2 By friends-of-friends method. For each existing consumer in each location the firm searches for $\gamma\mu\pi$ new consumers in other locations (here π measures the relative importance of friends-of-friends versus direct search). The destinations (locations) of firm's new consumers are chosen randomly according to the probability function $g(y,x)$. This probability depends on the distance between the previously found destination of the origin y and a newly found destination x, and the size of a new destination x.

The model claims that over time the distribution of served locations of two initially identical firms diverges and any firm's consumers become more and more geographically dispersed and clustered.

This paper applies the elaborated version of the model of Chaney (2014) to the analysis of the formation of Lithuanian export's network. The baseline Chaney (2014) model assumes that the probability function, which defines the search process of the new consumers, depends only on the distance between the origin and the destination and the size of the latter. However, this paper seeks to include more factors that may be important for the consumer search process and consequently affect the export's network, i.e. the factors that are commonly used in gravity models: the similarity and/or knowledge of languages, cultural similarities, belonging to the same economic union, free trade agreements, emigrants, etc.

In the paper we attempt to test which direct and undirect effects are the most important for the development of the extensive margin of international trade.

We also expect to find that during the first years of the opening to foreign trade, the cultural, language and historical factors should matter more to the formation of Lithuania's export's network than geographical and economic growth factors. In later years geography and economic factors are expected to dominate over the first ones. It is based on the assumption of inertia, the fear of changes and the need of time to learn new things. It should be easier to export to the same markets as before, which are already known and culturally similar. However, after some time the exporters, seeking larger profits which could be gained by exporting to richer countries, would learn new languages, gather information about other markets and form links with the other, more economically developed countries.

Moreover, we expect to find the growing importance of friends-to-friends search of the foreign markets and clustering with every year. This expectation comes from the network theory which relates higher clustering to lower costs of the search of new partners that are close by (Jackson, 2008). With every time period the network becomes larger, and the importance of indirect search becomes higher. According to the "preferential attachment" link formation process, new relations, formed by the current export partner of an origin should form with the other markets that are close by (either by distance or by other kind of similarities (in our case: cultural, economic, language, etc.)) to the existing export partners of an origin. Hence, clustering should increase.

To examine the factors that are important for the export's network formation the following model was estimated:

$$\mathbf{1}\left[Export^{s}_{i,c,t+1} > 0 \right] = \alpha Markets^{s}_{i,t} + \boldsymbol{\beta}_1 \mathbf{X}^{s}_{\underset{\text{(direct)}}{\text{i,c,t}}} + \boldsymbol{\beta}_2 \mathbf{X}^{s}_{\underset{\text{(f-o-f)}}{\text{i,c,t}}} + \delta \mathbf{1}\left[Export^{s}_{i,c,t} > 0 \right] + Controls^{s}_{c,t} + \varepsilon^{s}_{i,c,t} + u^{s}_{i,c}$$

Here $\mathbf{1}\left[Export^{s}_{i,c,t} \right]$ is an identity function indicating if product *i* from the country of origin *s* (i. e. Lithuania) was exported to country *c* at time *t*. $Markets^{s}_{i,t}$ shows the number of country's *s* export markets at time *t*. Vector $\mathbf{X}^{s}_{\underset{\text{(direct)}}{\text{i,c,t}}}$ stands for all the factors influencing the direct search procedure (e.g. average growth of imports of country *c*, distance, language, cultural similarities, organisational dependency, free trade agreements, etc. between the country of origin *s* and country *c*). Vector $\mathbf{X}^{s}_{\underset{\text{(f-o-f)}}{\text{i,c,t}}}$ stands for all the factors influencing the friends-of-friends search procedure (e.g. average growth of imports of country *c* from the current export markets of country *s*, distance, language, cultural similarities, organisational dependency, free trade agreements, etc. between country *c* and the current country's *s* export markets at time *t*). Controls stand

for any additional controls needed (e.g. country size, industry dummy, etc.). Finally, ε is an error term and u is the individual effect for the destination and product.

The model is specified under Linear Probability (OLS) and Probit specifications.

3 A Description of the Data and Definition of Variables

The paper uses detailed product and country level data from a number of different databases. The main product level bilateral trade data (between Lithuania and the other countries) comes from the BACI World trade database developed by CEPII. The period used for the analysis is from 1995 to 2015, because 1995 is the first year that any statistics for Lithuania were started to be collected. Commodities are classified by hs92 6-digits classification which is the lowest possible aggregation distinguishing about 5000 commodities (Gaulier and Zignago, 2010). Export is accounted for if its value is at least 1 000 USD per year.

Aggregated bilateral import data by country is taken from Robert Feenstra's database (www.robertfeenstra.info). Nominal GDP and immigration data come from the World Bank's database. Immigration data is represented by the international migrants' stocks, i. e. "*the total number of international migrants present in a given country at a particular point in time*" (Handbook on Measuring..., 2017), in the destination country by each country of origin as of the year 1990. Finally, the data for weighted average distances between countries, the measure of common spoken languages between a pair of countries as of the year around 2012 and free trade agreements (FTA) between different countries (according to the information from the WTO) are retrieved from www.cepii.org. FTA is specified by a dummy variable, which is equal to 1 if WTO reported a free trade agreement between the two countries in year t and 0 otherwise.

European Union membership is accounted for each year. Soviet bloc membership is marked if a country was considered to belong to the Soviet bloc before 1990s. The countries included in the Soviet bloc in the paper are: Lithuania, Latvia, Estonia, Moldova, Ukraine, Belorussia, Hungary, Poland, Romania, Slovakia, Check Republic, Bulgaria, Russia, Armenia, Azerbaijan, Georgia, Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, Uzbekistan, Afghanistan, Vietnam, Mongolia, Laos, Mozambique, Yemen, Ethiopia, Cuba, North Korea, Benin, Angola and Republic of Congo. Both EU membership and Soviet bloc membership are accounted by dummy variables. The EU dummy is equal to 1 if both countries were the members of the EU in year t and 0 otherwise. The Soviet bloc dummy is equal to 1 if both countries belonged to the Soviet bloc any time before the 1990s and 0 otherwise.

This paper analyses one country of origin s, i. e. Lithuania, and its exports to the other countries of destination c. We have a total of 183 countries of destination. Exported products are identified by an index i. The total number of Lithuania's exported products is 4,863 and the number of products exported from Lithuania to any of these 183 countries in any year between 1995 and 2015 is 2,099. As this paper deals with an extensive margin of international trade, we are interested only in the fact if Lithuania exported product i to country c, or not. The exact volume of such exports is irrelevant for the analysis. However, our data records yearly exports of a product only if it is not less than 1,000 USD. Therefore, our identity function $Exporting^{s}_{i,c,t} = \mathbf{1}\left[Export^{s}_{i,c,t} > 0 \right]$ gives 1 for any exports of commodity i from the country of origin s to the country of destination c if it was at least 1,000 USD in year t, and 0 otherwise.

Our **endogenous variable** is *Exporting*, which shows if product *i* was exported from the country of origin *s* to the country of destination *c* at time *t+1*. Its value can be either 0 (not exported) or 1 (exported).

$$Exporting = Exporting^{s}_{i,c,t+1} = \mathbf{1}\left[Export^{s}_{i,c,t+1} > 0 \right].$$

We have two kinds of **exogenous variables**: direct effects and indirect, or friends-of-friends, effects. The direct variables give general information about a particular destination, e.g. its' overall import growth, GDP level, the measure of its' ability to communicate with the origin, etc. On the other hand, the friends-of-friends (i.e. f-o-f) variables give network-specific information about a particular destination, e.g. its' import growth from the current export destinations of the origin, the sum of the distances between this particular destination and all the current export partners of the origin, the average measure of its' possibility to communicate with the current export partners of the origin, etc. All the exogenous variables are presented below.

- Previous exporting ($Exporting^{s}_{i,c,t}$) shows if a product *i* was exported from the country of origin *s* to the country of destination *c* in the previous year (i. e. exporting with a lag of one year):

  $$Exporting^{s}_{i,c,t} = \mathbf{1}\left[Export^{s}_{i,c,t} > 0 \right].$$

  It reveals if the history of exporting of the product *i* to the country of destination is important for future exports.

- Export markets ($Markets^{s}_{i,t}$) gives the number of product's *i* export markets that Lithuania had in year *t*.

  $$Markets^{s}_{i,t} = \sum_{c} \mathbf{1}\left[Export^{s}_{i,c,t} > 0 \right]$$

  We expect that higher number of existing export markets helps to speed up the spread of Lithuania's export to the other markets by the indirect search procedure, and, hence, results in an increase of Lithuania's exports markets in the next year.

- Logarithm of nominal GDP ($lngdp_{c,t+1}$) of the destination country in millions USD at time *t+1*. It is expected that higher nominal GDP as an indication of country's economic development leads to greater and more diverse imports, as the exporting countries see larger exports market and greater potential there.

- Import growth ($grimp_{c,t}$) is calculated as the growth of total imports (from all the other countries *c'*) to destination *c* between the years *t* and *t+1*, i. e. during the whole year *t*.

  $$grimp_{c,t} = \ln\left[\sum_{c'} Imports^{c'}_{c,t+1} \right] - \ln\left[\sum_{c'} Imports^{c'}_{c,t} \right]$$

  It is expected that higher overall growth of imports to country *c* attracts more imports to the same country from the country of origin *s*. The reason for that is the growth of potential in this market.

- Import growth f-o-f ($impgrow\_fof_{i,c,t}$) gives the total growth of imports during the whole year *t* to the destination country *c* from all the countries in which the country of origin *s* exported a product *i* at year *t*. Here, *c* and *c'* both stands for any destination country.

  $$impgrow\_fof_{i,c,t} = \ln\left[\sum_{c'} Exporting^{s}_{i,c',t} \cdot Imports^{c'}_{c,t} \right] - \ln\left[\sum_{c'} Exporting^{s}_{i,c',t-1} \cdot Imports^{c'}_{c,t-1} \right]$$

We expect the country of origin to have higher probability of starting exports to the country in which the exports of its other export partners is growing because of the imitating, i. e. if my neighbour thinks that it is good exporting there, it seems good for me to start exporting there as well.

- Distance from the origin ($distance\_or_c$) is calculated as the log of distance (in thousands) between the origin country *s* (Lithuania) and the destination country *c*.

$$distance\_or_c = \ln\left(distance_{s,c}/1000\right)$$

Distances between countries are taken from the CEPII database. We use the measure of a weighted distance *distw* which is "*based on bilateral distances between the biggest cities of the two countries, those inter-city distances being weighted by the share of the city in the overall country's population*" (Mayer and Zignago, 2011). Such a distance between the countries *i* and *j*($distw_{ij}$) was calculated by the CEPII according to the formula below. Here pop_k stands for the population of agglomeration *k* belonging to country *i* and the parameter θ is set equal to 1. For the 10 countries having only one city counted in the dataset, CEPII replaced the weighted distances by the simple distances between the largest cities of the countries.

$$distw_{ij} = \left(\sum_{k \in i} \left(pop_k / pop_i \right) \sum_{l \in j} \left(pop_l / pop_j \right) d_{kl}^{\theta} \right)^{1/\theta}$$

According to the gravity theory, trade is expected to be higher with less distant countries, i.e. the larger distance from the origin, the lower the probability of starting exports to a particular country.

- Distance f-o-f ($distance\_fof_{i,c,t}$) indicates the sum of the distances between any destination country *c* and all the other countries to which the origin exported the product *i* in year *t*. It is calculated as follows:

$$distance\_fof_{i,c,t} = \sum_{c'} \left[\ln\left(distance_{c,c'}/1000\right) \cdot Exporting^{s}_{i,c',t} \right]$$

We expect to have a higher probability that the origin starts exporting to the country which is closer to the current export partners of the origin, i.e. the larger $distance\_fof_{i,c,t}$ the lower the probability to start exporting to country *c*.

- Soviet bloc membership ($sb_{s,c}$) is a dummy variable which can be either 0 or 1. $sb_{s,c} = 1$ if both country *s* and country *c* belonged to the Soviet bloc before 1990, and 0 otherwise. This stands for cultural similarities. It is expected for Lithuania, which belonged to the Soviet bloc, in the first years to have higher probability to start exporting to the countries that also belonged to the Soviet bloc, i. e. $sb_{s,c}$ should positively affect the endogenous variable.

$$sb_{s,c} = sb_s \cdot sb_c$$

- Soviet bloc f-o-f membership ($sb\_fof_{i,c,t}$) is a dummy variable indicating indirect effect of belonging to the Soviet bloc. It is calculated as following:

$$sb\_fof_{i,c,t} = \frac{\sum_{c'} \left[sb_{c,c'} \cdot Exporting^{s}_{i,c',t} \right]}{\sum_{c'} Exporting^{s}_{i,c',t}}$$

For the first years we expect higher number assigned for $sb\_fof_{i,c,t}$ to show that the country *c* is more likely to become Lithuania's export partner for the good *i*, because it is more similar (in terms of belonging to the Soviet bloc) to more current Lithuania's export partners for the good *i*.

- EU membership ($eu_{s,c,t+1}$) is a dummy variable which can be either 0 ir 1. $eu_{s,c,t+1} = 1$ if both country *s* and country *c* belonged to the European Union at time *t+1*, and 0 otherwise. It is expected that the EU as an economic alliance promotes trade among its member states; hence, EU membership dummy should positively affect exporting.

$$eu_{s,c,t+1} = eu_{s,t+1} \cdot eu_{c,t+1}$$

- EU f-o-f membership ($eu\_fof_{i,c,t+1}$) is a dummy variable indicating indirect effect of belonging to the EU. It shows that country *c*, as well as a larger number of current Lithuania's export partners of a good *i*, is also an EU member, and is calculated as following:

$$eu\_fof_{i,c,t+1} = \frac{\sum_{c'} \left[eu_{c,c',t+1} \cdot Exporting^{s}_{i,c',t} \right]}{\sum_{c'} Exporting^{s}_{i,c',t}}$$

We expect that higher value of $eu\_fof_{i,c,t+1}$ positively affects Lithuania's exporting to country *c*, due to the benefits of the same trade bloc.

- Free trade agreements (FTA) ($fta_{s,c,t+1}$) is a dummy variable, indicating that there was a free trade agreement between country of origin *s* and the destination country *c* at time t+1. If both countries had a free trade agreement at time t+1 $fta_{s,c,t+1} = 1$, otherwise it is 0. We expect that the countries having a free trade agreement are more likely to have trade relations with each other.

$$fta_{s,c,t+1} = fta_{s,t+1} \cdot fta_{c,t+1}$$

- FTA f-o-f ($fta\_fof\_a_{i,c,t+1}$) is a dummy variable, indicating indirect effect of belonging to the same trade union at time t+1, calculated as follows:

$$fta\_fof\_a_{i,c,t+1} = \frac{\sum_{c'} \left[fta_{c,c',t+1} \cdot Exporting^{s}_{i,c',t} \right]}{\sum_{c'} Exporting^{s}_{i,c',t}}$$

Positive relationship is expected due to the same reasons as for the EU f-o-f membership variable.

- Language ($lang_{s,c}$) shows countries' ability to communicate with each other. It is measured by "*cls*" measure evaluated by CEPII and can be between 0 (nobody in the two countries can understand each other) and 1 (every two people taken from the two countries will be able to communicate with each other) (Melitz and Toubal, 2014). This variable should positively affect the probability of exporting.

- Language f-o-f ($lang\_fof_{i,c,t}$) stands for indirect countries' ability to communicate and is calculated as following:

$$lang\_fof\_a_{i,c,t} = \frac{\sum_{c'}\left[cls_{c,c'} \cdot Exporting^{s}_{i,c',t}\right]}{\sum_{c'} Exporting^{s}_{i,c',t}}$$

  It is an average sum of all "cls" measures between country c and all Lithuania's export partners of a good i. Higher value of $lang\_fof_{i,c,t}$ shows that country c is more capable to communicate with current Lithuania's export partners. It shows that Lithuania is more likely to find this market by the indirect search procedure. Hence, it should be positively related to the exporting variable.

- Migration ($migr\_ln_{s,c}$) is the log of migration stock from the country of origin s to the destination country c in 1990. We believe that higher numbers of immigrants living in country c promotes exports from their country of origin to the country c. Hence, we expect to find positive relationship between migration and exporting.

- Migration f-o-f ($migrln\_fof\_a_{i,c,t}$) is the log of average indirect migration calculated for the destination country c:

$$migrln\_fof\_a_{i,c,t} = \ln\left[\frac{\sum_{c'}\left[migr_{c,c',1990} \cdot Exporting^{s}_{i,c',t}\right]}{\sum_{c'} Exporting^{s}_{i,c',t}}\right]$$

  Higher number of $migrln\_fof\_a_{i,c,t}$ represents higher average stock of immigrants in the country c from all Lithuania's exports partners of a good i. We believe that it positively affects Lithuania's start of exporting to country c because of a chain reaction.

For the regressions, all the variables were normalised simply by subtracting the means and dividing by standard deviations of each variable.

4 Descriptive Statistics

In Table 1. we provide descriptive statistics of all the variables before normalisation. The number of observations differs due to the differences in the variables (e. g. GDP is given for each country and year, while export is given for each country, year and product) and unbalanced panel (e. g. GDP of Afghanistan is known only since 2001).

Table 1. shows that 94 % of exports data is either zero or missing, i.e. Lithuania's exports basket varies across it's exports partners. The situation is similar for all the countries of the world.

16 % of all 183 countries belonged to the Soviet bloc. 15 % of all the countries belonged to the EU in the year 2014, however only 8 % of the countries belonged to the EU in 1995.

Table 1. Descriptive statistics

Variable	Obs	Mean	Stand. Dev.	Min	Max
Bilateral exports, thou USD	7 338 104	30 118	1 980 000	0	1.28e+09
Zero / missing values	*94 %*				
GDP, mill USD (current prices)	3 511	278 642	1 155 826	11	1.74e+07
No of migrants (stock)	711 333	3 932	67 507	0	5 211 922
Common spoken languages	707 112	0.117	0.227	0	1
Distance (weighted)	3 680	6 234	3 872	132	17 226
Import growth	3 496	0.080	0.271	-3.365	4.063
Free Trade Agreements (all years)	707 112	0.091	0.288	0	1
Existing bilateral FTAs, % obs	*9 %*				
No bilateral FTAs, % obs	*91 %*				
Soviet bloc membership (2000)	33 672	0.026	0.159	0	1
Ex-Soviet bloc countries, % countries	*16 %*				
Non-Soviet bloc countries, % countries	*84 %*				
EU membership (2014)	33 672	0.021	0.1429	0	1
EU member countries, % countries	*15 %*				
Non-EU member countries, % countries	*85 %*				

9 % of all observations have a positive bilateral free trade agreement. The number of countries having bilateral FTAs increased: in 1995 only 4 % of the analyzed 183 countries had bilateral FTAs and in 2014 such countries were almost 14 %. The number of countries having bilateral FTAs with Lithuania (22 % of the countries taking all the years) also grew over the period: in the year 1995 Lithuania had a free trade agreement with only 8 % of the analyzed 183 countries and in the year 2014 this number increased to 39 %. We can conclude that Lithuania is a small open economy.

In Figure 1 one can notice a sharp increase in Lithuania's exports both in volumes and in the number of markets during the whole period. The upper panel presents the change of the intensive margin of Lithuania's exports, i. e. log of the export volume, during the years 1995-2014 (time is shown on the x-axes). It increased from 21.9 in the first given year to 24.1 in the last given year. The middle panel shows how the number of Lithuania's export markets increased over the period. The increase was by nearly 70 %: from 109 export markets in 1995 to 184 destinations in 2014. Finally, the lower panel presents the growth of the average number of export markets of each product type. Different Lithuanian products in different years have between 0 and 121 export markets. In 1995 one Lithuanian export product on average had 4.6 export markets and in 2014 one product on average was exported to 12.1 destinations.

Figure 1. The spread of Lithuanian Exports, 1995-2014.

The greatest increase and the general formation of Lithuania's trade pattern is seen between the years 1999 and 2008. After 2008 exports shrank due to the effect of the world economic crises and slowly recovered afterwards. The same export dynamics is seen for the other Eastern European countries.

Due to a large number of variables we also have a high possibility of multicollinearity problems. According to the cross-correlations table, given in Table 2, it is highly likely that direct variables are correlated with the corresponding indirect variables. It is especially true for the Soviet bloc, EU, FTA and language variables. Direct migration variable is also correlated with direct language variable. Distance from the origin variable, being correlated with direct migration as well as indirect EU and FTA variables, seems to have the highest number of cross-correlations.

Table 2. Cross-correlations among the variables

	Exporting	Dist. Or.	Dist. f-o-f	Sov. Bl.	SB f-o-f	EU	EU f-o-f	FTA	FTA f-o-f	Lang.	Lang.f-o-f	Migr.	Migr.f-o-f
Exporting	**1,00**												
Distance origin	**-0,37**	1,00											
Distance f-o-f	**-0,03**	0,39	1,00										
Soviet block	**0,21**	-0,42	-0,14	1,00									
Sov. Block f-o-f	**0,18**	-0,38	-0,18	0,90	1,00								
EU membership	**0,28**	-0,45	-0,16	0,11	0,09	1,00							
EU f-o-f	**0,28**	-0,50	-0,19	0,09	0,07	0,85	1,00						
FTA	**0,21**	-0,41	-0,10	-0,07	-0,07	0,55	0,57	1,00					
FTA f-o-f	**0,25**	-0,53	-0,17	0,13	0,11	0,47	0,58	0,82	1,00				
Languages	**0,41**	-0,43	-0,15	0,38	0,34	0,26	0,25	0,20	0,30	1,00			
Language f-o-f	**0,25**	-0,29	-0,07	0,10	0,15	0,25	0,32	0,25	0,37	0,75	1,00		
Migrants	**0,40**	-0,50	-0,19	0,46	0,41	0,32	0,31	0,23	0,35	0,57	0,32	1,00	
Migrants f-o-f	**0,29**	-0,42	0,08	0,34	0,28	0,21	0,25	0,16	0,32	0,27	0,25	0,47	1

5 Results

5.1. General results

First of all, we estimate a couple of simple benchmark regressions, which include only the most important gravity model's variables: GDP and distance from the origin in the first one and adding previous exporting and a number of markets in the second one. We estimate the following equations:

$$\mathbf{1}\left[Export^{s}_{i,c,t+1}>0\right]=\alpha+\beta_1\ln GDP_{c,t+1}+\beta_2 distance^{s}_{c}+\varepsilon^{s}_{i,c,t}+u^{s}_{i,c}$$

$$\mathbf{1}\left[Export^{s}_{i,c,t+1}>0\right]=\alpha+\beta_1\ln GDP_{c,t+1}+\beta_2 distance^{s}_{c}+\beta_3 Markets^{s}_{i,t}+\beta_4\mathbf{1}\left[Export^{s}_{i,c,t}>0\right]+\varepsilon^{s}_{i,c,t}+u^{s}_{i,c}$$

Here $\mathbf{1}\left[Export^{s}_{i,c,t}\right]$ is an identity function indicating if product i from the country of origin s (Lithuania) was exported to the country of destination c at time t. $\ln GDP_{c,t+1}$ gives logarithm of the nominal GDP of a destination country, $distance^{s}_{c}$ shows the distance between Lithuania and a destination country c, and $Markets^{s}_{i,t}$ shows the number of country's s export markets at time t. Finally, ε is an error term and u is the individual effect for the destination and product. Each model we estimated twice: using simple OLS and Probit methods with random effects. Product fixed effects are not applicable because distance from the origin changes only by destination, hence, this variable would not have been included in the regressions. We also control for the product sector, as products from different sectors are more or less likely to be exported to any particular country. Removing the sector controls does not have any noticeable impact for the results.

The results of the benchmark models are given in Table 3. It shows that classical gravity model's variables GDP and distance behave according to the trade theory (see e.g.: Anderson, 2011). However, having such a short model leaves huge amount of information unexplained, because adding previous exporting and a

number of markets improves the model significantly. Results show that GDP, the number of markets and previous exporting affects future exports positively, while larger distance affects future exports negatively as expected.

Table 3. Results of the benchmark model regressions

Variables, normalized	PROBIT marginal effects		OLS	
Dependent variable	Exporting, t+1	Exporting, t+1	Exporting, t+1	Exporting, t+1
GDP (log mill)	0.0674***	0.03661***	0.05088***	0.01307***
	(0.00106)	(0.0004)	(0.00101)	(0.00029)
Distance (origin)	-0.0509***	-0.03973***	-0.07534***	-0.03097***
	(0.000460)	(0.00023)	(0.00078)	(0.00029)
Markets, t		0.01969***		0.01804***
		(0.00023)		(0.00017)
Exporting, t		0.01189***		0.15208***
		(0.0001)		(0.00054)
Constant	yes	yes	yes	yes
Individual effects	random	random	random	random
R-squared	-	-	0.1575	0.5133
Observations	7,338,104	7,338,104	7,338,104	7,338,104
Number of ID	386,216	386,216	386,216	386,216

* Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

** All the variables in the regressions were normalized. For the OLS models regressions coefficients are given and for the Probit models marginal effects are presented. Random effects are by the destination and the product, standard errors are clustered by product. Robust standard errors are given in parenthesis.

Next we investigate the effect of our 17 exogenous variables to the probability of starting exporting. We estimate 3 types of regressions: simple OLS method introducing random effects by destination and product, and 2 kinds of Probit regressions introducing either random or correlated random effects by destination and product (see: Mundlak, 1976). Due to the evident cross-correlations of corresponding direct and indirect (f-o-f) variables (see Table 2), we constructed 3 types of regressions for each method (see Tables 4, 5 and 6 respectively): the first one includes all the exogenous variables, the second one contains just the direct factors and the third one has only the indirect variables.

As we can see from the Table 4, almost all the factors are reported to be highly significant (except for the Soviet bloc and the free trade agreements variables in some of the models). This is mainly due to the very large number of observations (7.3 mill). Most of the signs are similar among the specifications and as expected, except for the distance from the origin in the OLS model which unexpectedly becomes positive. Most probably it is because of the more significant cross-correlations between distance from the origin and the other variables (see: Table A1 and Figure A1 in the Appendix for the estimation results excluding distance from the origin).

Other interesting differences by specification include indirect migration and Soviet bloc factors. The first one, although in all the specifications significant and positive, has much higher value for both Probit specifications than for the OLS; and the latter is significantly negative for the Probit with random effects and the OLS specifications, however, insignificant for the Probit with correlated random effects specification. The greatest difference in the results are between OLS and the other two models. The reason for this might

be because the coefficients are very small numbers, therefore, only the very beginning of the non-linear distribution function is taken into account and linear smoothing in this part changes results severely. Overall, estimation results are robust and similar for all the specifications.

Table 4. Results of the regressions including both direct and undirect variables

Variables, normalized	PROBIT marginal effects		OLS
Dependent variable	Exporting, t+1	Exporting, t+1	Exporting, t+1
Markets, t	0.0289***	0.016264***	0.06839***
	(0.000192)	(0.000135)	(0.00019)
Exporting, t	0.0108***	0.007901***	0.13340***
	(0.000057)	(0.000036)	(0.00008)
GDP (log mill)	0.0257***	0.029128***	0.00905***
	(0.000195)	(0.000300)	(0.00008)
Import growth	0.00168***	0.001088***	0.00080***
	(0.000106)	(0.0000815)	(0.00006)
Distance (origin)	-0.0105***	-0.000380***	0.00315***
	(0.000157)	(0.0000946)	(0.00011)
Soviet block	-0.00311***	-0.000096	-0.00935***
	(0.000210)	(0.000133)	(0.00016)
EU membership	-0.000751***	0.000868***	0.00091***
	(0.000088)	(0.000073)	(0.00013)
FTA	-0.000058	-0.000240**	-0.00237***
	(0.000127)	(0.000105)	(0.00013)
Languages	0.0185***	0.001864***	0.03610***
	(0.000156)	(0.0000874)	(0.00013)
Migrants (log)	0.00908***	0.001489***	0.01418***
	(0.000153)	(0.000079)	(0.00010)
Imp.growth f-o-f	0.00109***	0.000587***	0.00017**
	(0.000166)	(0.000121)	(0.00007)
Distance f-o-f	-0.0123***	-0.003468***	-0.05299***
	(0.000221)	(0.000119)	(0.00020)
Sov. Block f-o-f	0.00359***	0.001293***	0.00674***
	(0.000158)	(0.000104)	(0.00015)
EU f-o-f	0.00138***	0.001605***	0.00909***
	(0.000105)	(0.000096)	(0.00013)
FTA f-o-f	0.00336***	0.002376***	-0.00223***
	(0.000139)	(0.000119)	(0.00013)
Language f-o-f	0.000242*	0.000944***	-0.00869***
	(0.000141)	(0.000089)	(0.00011)
Migrants f-o-f	0.00542***	0.003641***	0.00094***
	(0.000131)	(0.000088)	(0.00008)
Constant	yes	yes	yes
Individual effects	random	correlated random	random
R-squared	-	-	0.5332
Observations	7,338,104	7,338,104	7,338,104
Number of ID	386,216	386,216	386,216

* Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

** All the variables in the regressions were normalized. For the OLS models regressions coefficients are given and for the Probit models marginal effects are presented. Random effects are by the destination and the product. Robust standard errors are given in parenthesis.

According to these results the most important factors, influencing Lithuanian exports are:

- the history of exports (positive),
- the development of a trade network and openness of an origin (positive),
- development of the economy of the destination market (positive),
- ability to communicate (positive),
- direct and indirect stock of migrants in the destination countries (positive),
- direct and indirect geographic distances (negative).

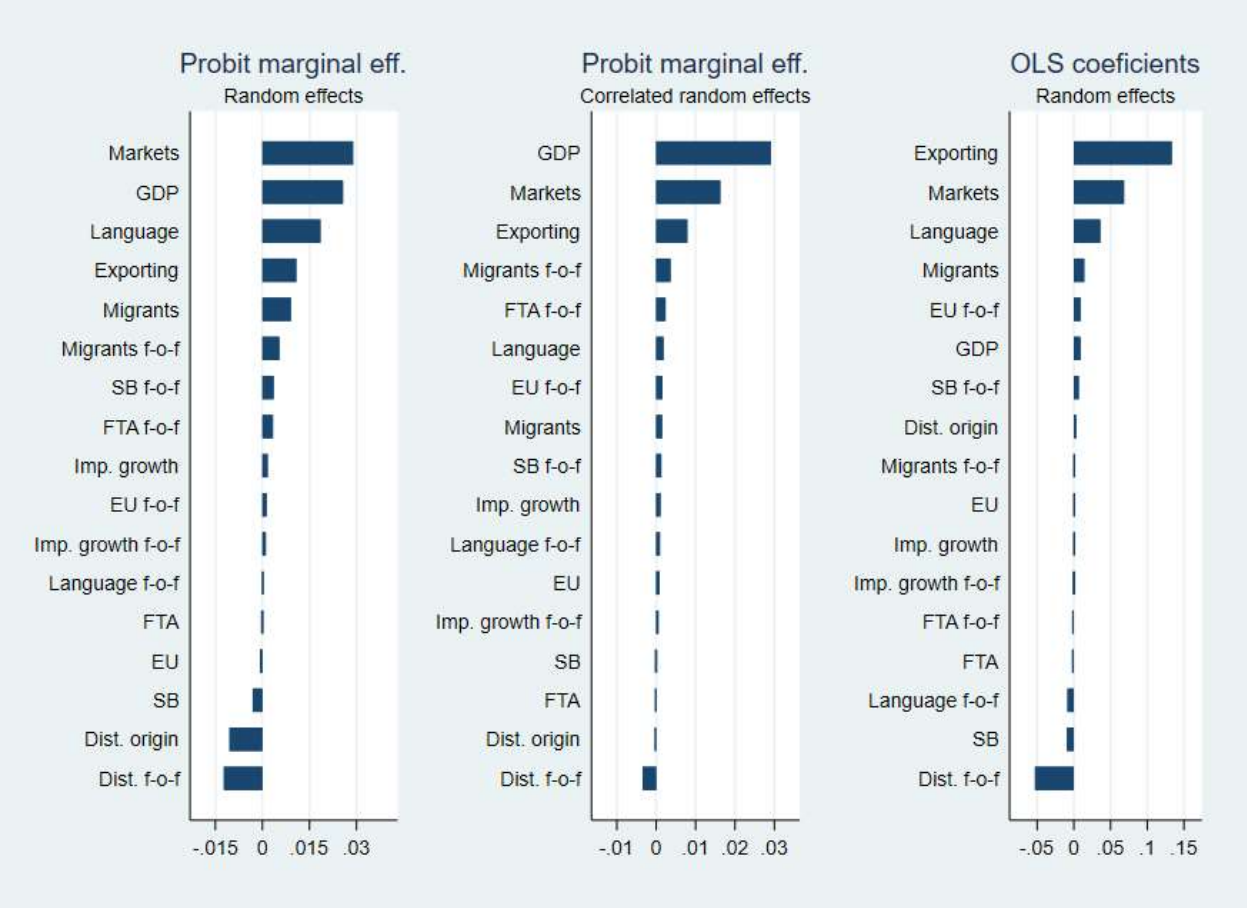
Other robust factors are: indirect effects of belonging to the EU and the Soviet bloc and both direct and indirect effects of the import growth of the export partners (all positively). The effects of the other factors seem to be ambiguous, as estimated regressions' coefficients differ by specification.

Graphic comparison of the general regression results of the four models is presented in Figure 2. In the first 2 columns marginal effects of both Probit models are given, and in the third one regression coefficients of OLS random effects model are presented. Results are displayed starting from the most influential positive (GDP, the number of markets or exporting) and finishing with the most influential negative (indirect distance). Note that we can compare the coefficients directly as all exogenous variables have been normalized.

The results confirm the hypothesis of the importance of network effects for the development of the extensive margin of international trade. The most important network effect is indirect distance. The results are in line with the economic theory by showing negative effects of both direct and indirect distances. The other important network effects are: migration, import growth of the destination market and the membership in the EU and the Soviet bloc.

However, the direct effects of the Soviet bloc and the EU variables, indirect effect of the common spoken languages, as well as both effects of the free trade agreements are controversial. The lack of influence of participation in any trade union could be explained by the rather late starting year of the data. As Lithuania became independent from the Soviet Union in 1990 and was located in the middle of Europe, i.e. having good opportunities to create trade relations with its neighboring countries due to the small distances, in 1995 Lithuania could have already created strong relations with the other EU and the Soviet bloc countries and made successful free trade agreements. Hence, examining just the extensive margin, we were not able to see any strong effects of these factors for the creation of Lithuania's trade relations, as Lithuania already had trade relations with most of these countries. However, these 3 factors could still be influential for the development of Lithuania's trade relations in terms of an intensive trade margin. The lack of evidence of an indirect effect of the common spoken languages may be due to the possible multicollinearity with the direct language variable.

Figure 2. Comparison of the general regression results.



To check if these results are affected by the possibility of a multicollinearity, we introduced separate models for direct and indirect variables

5.2. Analysis of the Direct Factors

Table 5 shows evaluation results of the regressions that contain only the direct factors, i.e. the number of markets, previous exporting, GDP, growth of import, distance from the origin, memberships to the Soviet bloc, the EU and a free trade union, common spoken languages and the stock of migrants. Omitting the network variables helps to avoid the problems that may arise because of the multicollinearity between similar direct and undirect variables seen in the Table 2. In spite of the positive and significant coefficients for the membership in the EU, the Soviet bloc and the FTA in the Probit models, the "direct" regression gives very similar results which are in line with our expectations. We may conclude that multicollinearity problem does not have a significant effect on the other results.

Table 5. Results of the regressions analyzing just the direct effects

Variables, normalized	PROBIT marginal effects		OLS
Dependent variable:	Exporting, t+1	Exporting, t+1	Exporting, t+1
Markets, t	0.01974***	0.01255***	0.02042***
	(0.00009)	(0.00009)	(0.00018)
Exporting, t	0.01128***	0.00808***	0.14017***
	(0.00006)	(0.00003)	(0.00061)
GDP (log mill)	0.02748***	0.02814***	0.00968***
	(0.00019)	(0.00030)	(0.00024)
Import growth	0.00202***	0.00123***	0.00068***
	(0.00009)	(0.00007)	(0.00005)
Distance (origin)	-0.01788***	-0.00282***	-0.01617***
	(0.00014)	(0.00008)	(0.00019)
Soviet block	0.00094***	0.00309***	-0.00232***
	(0.00014)	(0.00008)	(0.00019)
EU membership	0.00063***	0.00219***	0.00978***
	(0.00007)	(0.00005)	(0.00026)
FTA	0.00228***	0.00147***	-0.00300***
	(0.0001)	(0.00009)	(0.00014)
Languages	0.01824***	0.00228***	0.02572***
	(0.00013)	(0.00007)	(0.00024)
Migrants (log)	0.01002***	0.00216***	0.01395***
	(0.00016)	(0.00008)	(0.00022)
Constant	yes	yes	yes
Individual effects	random	correlated random	random
R-squared	-	-	0.52811
Observations	7,338,104	7,338,104	7,338,104
Number of ID	386,216	386,216	386,216

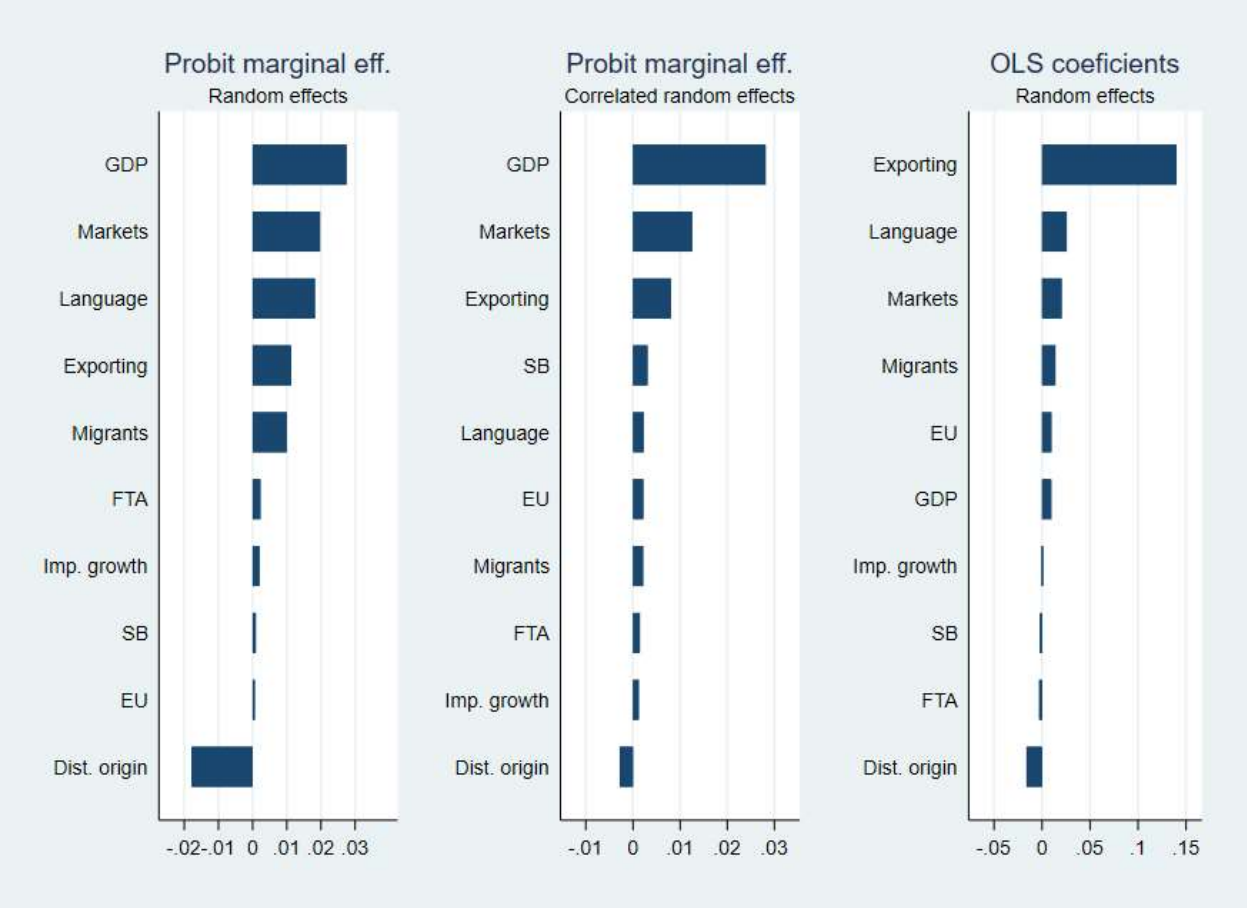
* Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

** All the variables in the regressions were normalized. For the OLS models regressions coefficients are given and for the Probit models marginal effects are presented. Random effects are by the destination and the product. Robust standard errors are given in parenthesis.

Figure 3 presents visual comparison of the regressions' coefficients for the OLS model and marginal effects for the Probit models.

The most influential direct factors for the expansion of Lithuanian trade are: the number of current export markets, exporting to the same destination in the previous year, common spoken languages and GDP growth (all of these affecting positively) as well as distance from the origin affecting negatively. Other robust, however, less influential factors include: the number of migrants and import growth in the destination country.

Figure 3. Comparison of the direct regression results



Similarly, as in the case of a general regression, the EU, the Soviet bloc and the FTA membership tend to have very low values and to change signs depending on the specification. The reason for it might be the same as in the general case.

We have almost identical results if we use indirect distance instead of the distance from the origin in these regressions (see: Table A2 and Figure A2 in the Appendix). The major change comparing these two models is that indirect distance appears to be more influential for the development of the extensive margin of Lithuania's foreign trade than the direct distance.

5.3. Analysis of the Indirect Factors

Next, we examine the "indirect" regressions including only the control variables (i.e. the number of markets, previous exporting and GDP) and the network variables (i.e. indirect import growth, distance, membership of the Soviet bloc, the EU and a free trade union, common spoken languages and the stock of migrants). The results of these regressions are presented in Table 6. The greatest change of the results of the "indirect" regressions comparing with the general models is in the effect of the indirect common spoken languages. In the reduced model it becomes clearly positive and significant for all the specifications. It shows that this

factor is influential for the development of the extensive trade margin of Lithuania, but possibly suffers from a significant multicollinearity in the general model. Similarly as in the previous models, the result for belonging to the FTA is unstable. As other results do not change, we may conclude that multicollinearity problem in the general regression is not severe.

Table 6. Results of the regressions analyzing just the indirect effects

Variables, normalized	**PROBIT marginal effects**		**OLS**
Dependent variable:	**Exporting, t+1**	**Exporting, t+1**	**Exporting, t+1**
Markets, t	0.0391***	0.0173***	0.06589***
	(0.000194)	(0.000122)	(0.00047)
Exporting, t	0.0124***	0.00792***	0.14348***
	(0.000064)	(0.000035)	(0.00052)
GDP (log mill)	0.0328***	0.0293***	0.01442***
	(0.000172)	(0.000286)	(0.00029)
Imp.growth f-o-f	0.00255***	0.00142***	0.00051***
	(0.000118)	(0.000098)	(0.00003)
Distance f-o-f	-0.0274***	-0.00443***	-0.05381***
	(0.000221)	(0.000101)	(0.00051)
Sov. Block f-o-f	0.0114***	0.00290***	0.00917***
	(0.000096)	(0.000056)	(0.00018)
EU f-o-f	0.000169**	0.00246***	0.00706***
	(0.000081)	(0.000068)	(0.00022)
FTA f-o-f	0.00345***	0.00153***	-0.00441***
	(0.000113)	(0.000099)	(0.00013)
Language f-o-f	0.0125***	0.00228***	0.01879***
	(0.000106)	(0.00006)	(0.00019)
Migrants f-o-f	0.00853***	0.00396***	0.00220***
	(0.000142)	(0.000083)	(0.00010)
Constant	yes	yes	yes
Individual effects	random	correlated random	random
R-squared	-	-	0.5223
Observations	7,338,104	7,338,104	7,338,104
Number of ID	386,216	386,216	386,216

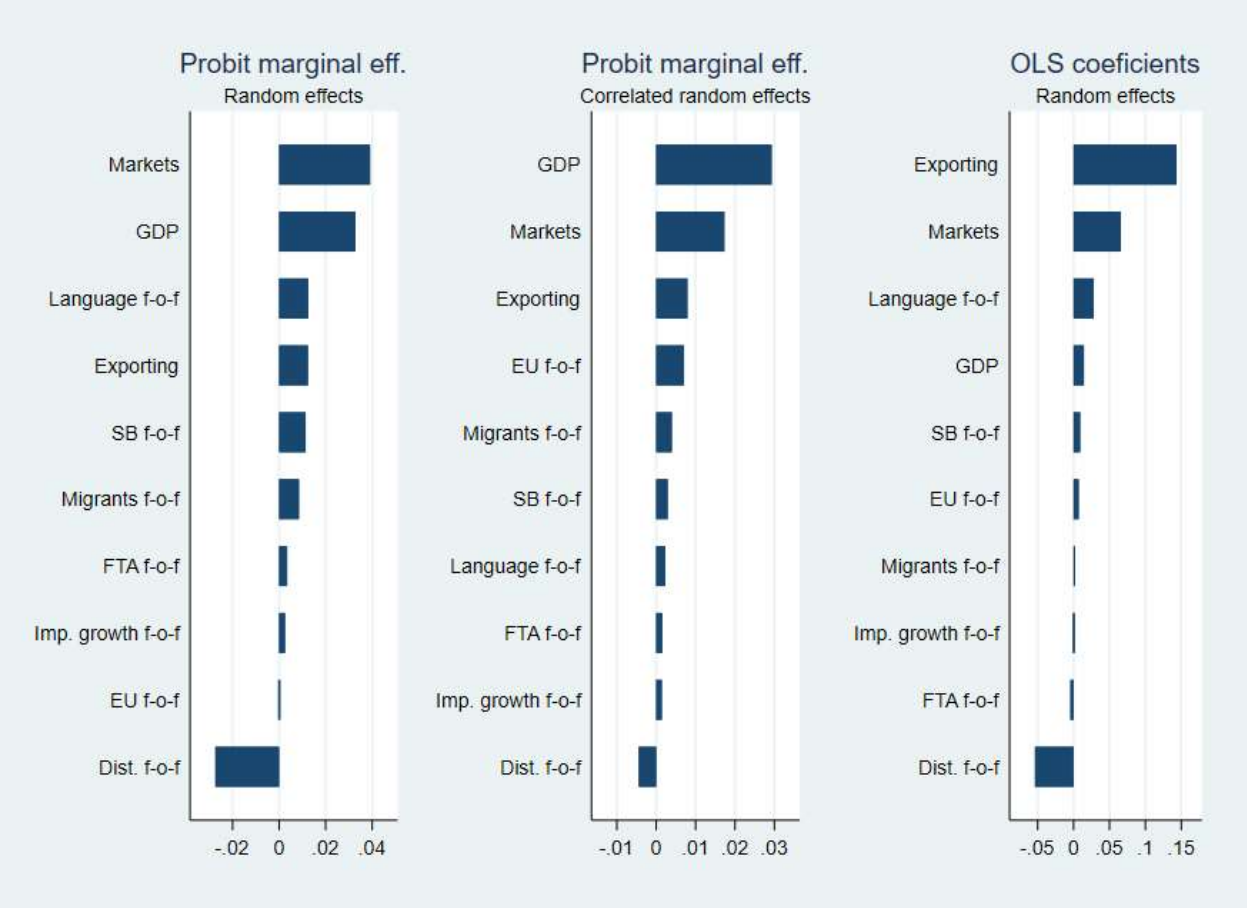
* Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

** All the variables in the regressions were normalized. For the OLS models regressions coefficients are given and for the Probit models marginal effects are presented. Random effects are by the destination and the product. Robust standard errors are given in parenthesis.

Figure 4. presents visual comparison of the indirect regressions' coefficients for the OLS model and marginal effects for the Probit models. It shows that the most influential robust factors are the number of export markets, previous exporting, GDP growth, indirect common spoken languages (all of which affects probability to export positively) and indirect distance (affecting negatively). We also have stable positive results for indirect migration and import growth.

We may conclude that Lithuania uses intermediate countries during the procedure of the search of the new export markets.

Figure 4. Comparison of the indirect regression results.



5.4. Summary of the Results

The results above indicate that:

- Network variables are important for the development of an extensive trade margin.
- Lithuania's extensive trade margin is mostly affected by previous exporting in a destination market, the number of current Lithuania's export markets, GDP of the destination market, both direct and indirect measure of common spoken languages, both direct and indirect migration (all positively) and both direct and indirect distance (negatively). Direct and indirect import growth to the destination country and the EU membership also has a small positive impact for Lithuania's trade creation.
- Indirect distance matters more than direct distance.
- There is no clear effect of the Soviet bloc membership and the free trade agreements. The reason for the lack of the Soviet bloc influence could be that trade was already created with the Soviet bloc

countries, while the motivation for the lack of the influence of the FTA could be that the most important of them were made during the first 2-4 years of Lithuania's independence. Hence, since 1995 there were not much fluctuations in Lithuania's extensive trade margin with these markets.

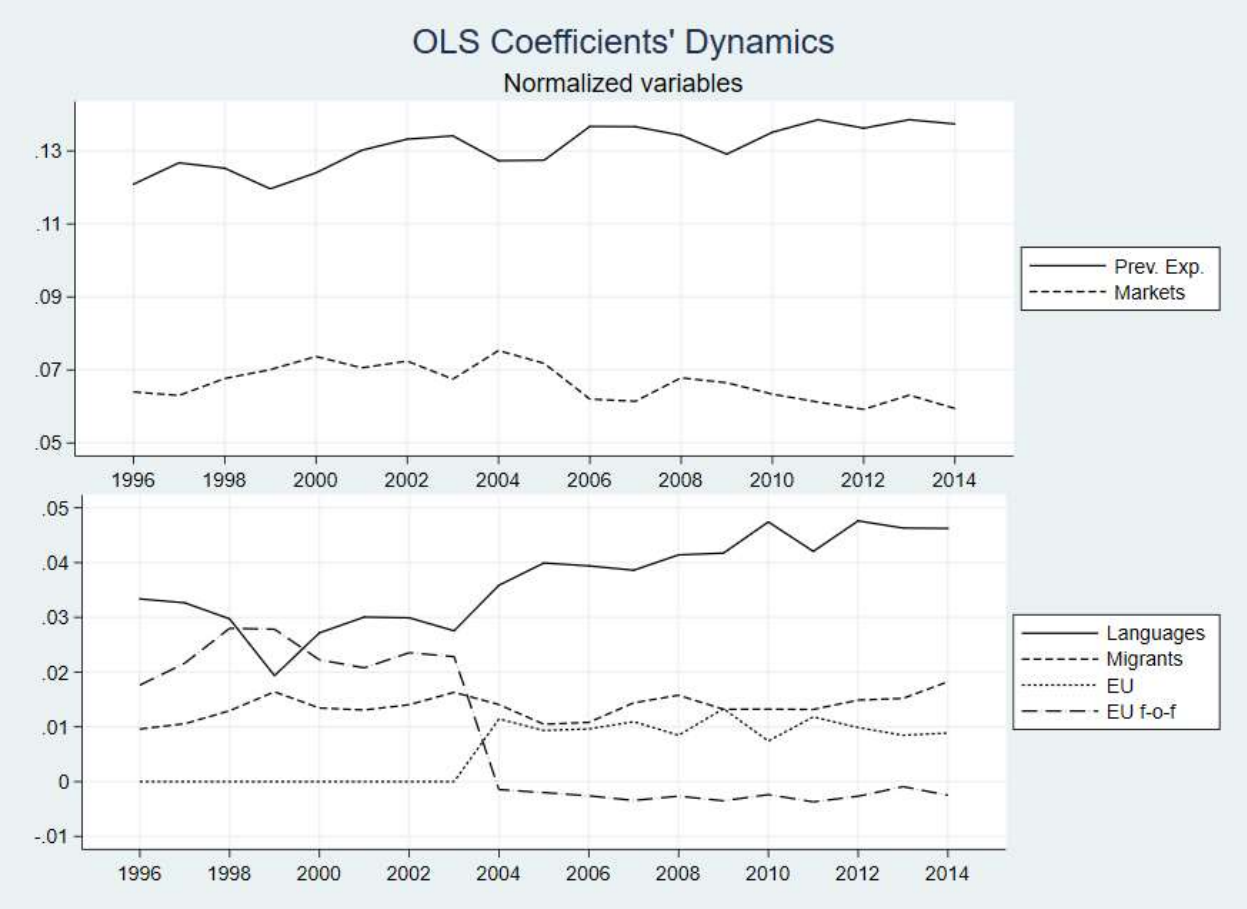
These findings show that the probability to export any product to any destination market mostly depends on the export's history of an origin (i.e. the number of current export markets of an origin) and of the economic factors of both markets (i.e. destination's GDP level). Lithuania's export network develops in an accelerating pace due to the friends-of-friends search procedure and this conclusion is in line with the networks theory. As already shown in a number of previous papers (Rauch, 1999; Felbermayer & Kohler, 2006; Allen, 2014; Chaney, 2014; etc.), even in the globalized world we have today, the larger is the distance between the countries the lower is the probability to start trade relations. Results remain the same for indirect distance. As was already shown by Rauch (1999), the ability to communicate with each other facilitates exports. We find that the ability to communicate also plays an important role during the friend-of-friend search procedure. We also find strong evidence that emigration boosts export from the country of origin. The reason for it could be that in their new country the emigrants like to consume the same goods that they were used to consume at home.

6 Analysis of the Dynamics of the Coefficients'

Having analyzed the overall effects of a number of direct and indirect variables it is interesting to examine the dynamics of their effects over time. Our hypothesis is that during the first years of the opening to foreign trade the influence of cultural, language and historical factors should dominate, i.e. Lithuania should trade with the other former Soviet bloc countries. In later years geographical and economic factors should become more important, i.e. trade should expand further from the former USSR countries.

To analyse the dynamics we apply a simple OLS model. In Figure 5 and Figure 6 we graph coefficients for each year since 1996 till 2014. The OLS model was chosen because of computational feasibility. To avoid at least some of the multicollinearity effects we will omit distance from the origin and have only indirect distance in the model.

Figure 5. Dynamics of the coefficients of an OLS regression with the most influential positive regressors

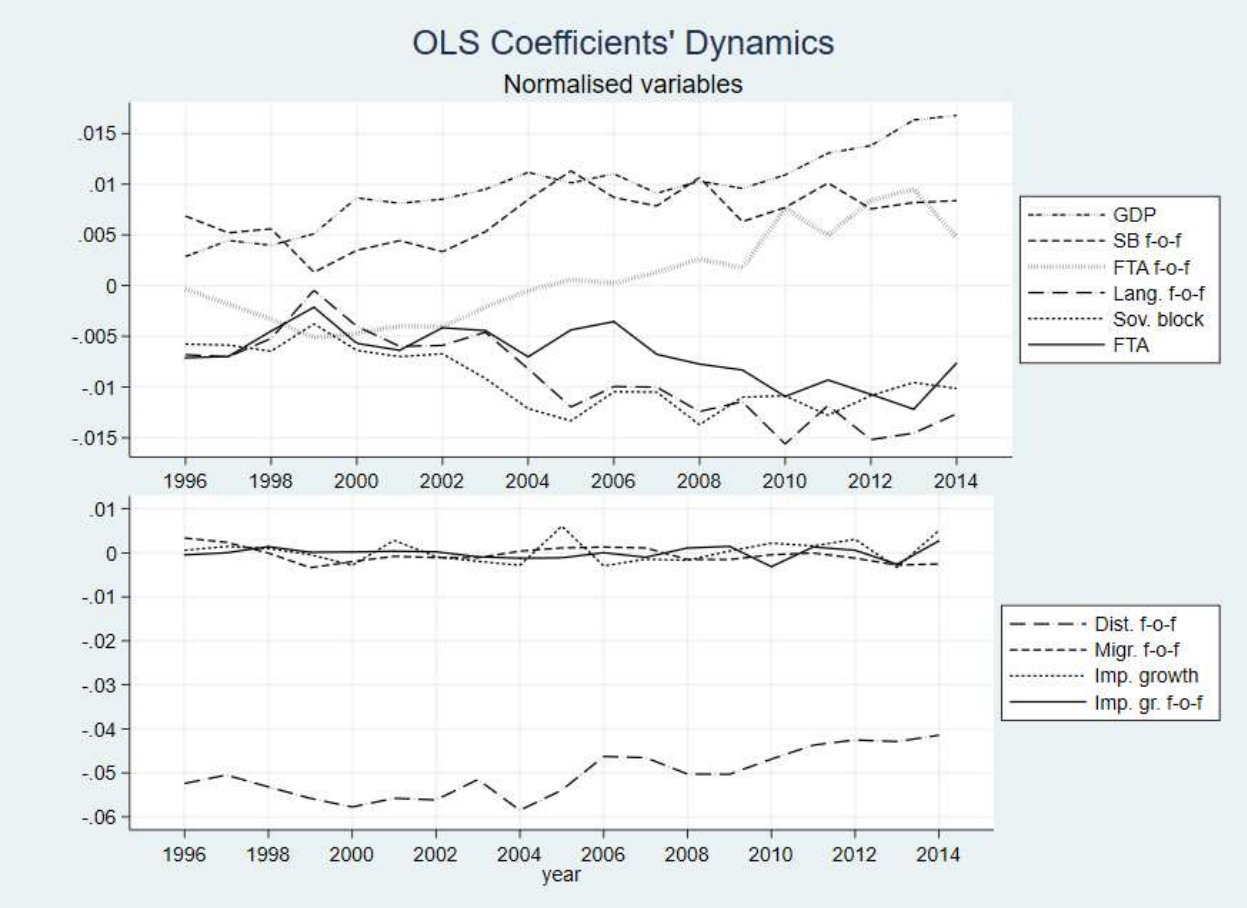


The upper panel of Figure 5 depicts the two most influential factors for the probability of exporting. The most influential factor is previous exporting to the destination market. According to the OLS model, the probability to export a particular product to any destination in year 1996 increased by ~12 % with one standard deviation increase in exporting of this product to the same destination in year 1995. The influence of this factor gradually increased throughout the years. The second most influential factor is a number of current export markets. The influence of this factor is stable throughout the period and does not have a clear trend.

The lower panel of Figure 5 shows an interesting trend of the coefficients of a direct and indirect membership of the EU. It could be explained by the fact that Lithuania joined the EU in 2004. Based on the construction of the direct EU dummy, it equals 0 in all the years when Lithuania was not an EU member. On the other hand, indirect EU membership was much more influential before the admission. It could show that before joining the EU indirect contacts with the EU members were crucial for Lithuania to get access to the EU market. However, both these coefficients are very small, and overall effects of the EU membership on the endogenous variable are negligible. Direct migration has a stable significant effect on the endogenous variable. The importance of common spoken languages is increasing gradually since 2003.

The other increasingly important positive factors, as shown in the upper panel of Figure 6, are destination country's GDP (which had a negligible effect on starting to export in 1996, however, increased sharply since 2009), indirect Soviet bloc membership and indirect free trade agreements.

Figure 6. Dynamics of the coefficients of an OLS regression with the other regressors



Increasingly important negative factors affecting the probability to start exporting are: Soviet block membership, free trade agreements and indirect effects of the common spoken languages. Their effects were close to 0 (that of the indirect FTAs even negative) before the 2002. However, later these factors have become more influential. Still, overall effects of these variables are negligible.

Indirect distance, given in the lower panel of Figure 6, is one of the most influential factors of the endogenous variable. It's effect was stable for the first decade and started to decrease slightly in recent years. The reason for it could be that Lithuania has already started to export to all the close-by countries, and other factors, like economic strength and ability to communicate, have become more important for the trade expansion to more distant regions.

The other variables (indirect migration and both direct and indirect import growth) has no clear trend and make a stable and significant, although negligible, influence to the probability of starting to export.

We draw the following conclusions from the analysis:

- The influence of the economic factors, i. e. GDP, increases throughout the years.
- The indirect effect of common spoken languages becomes more negative and the direct one increases. It can be explained by the exporters' attempts to learn foreign languages and to communicate with their export partners directly.
- The direct effect of the Soviet bloc membership becomes more negative and is in line with the hypotheses. The indirect effect of belonging to the Soviet bloc increases and that shows the expected increase of clustering of the countries.
- The indirect effect of free trade agreements increased in recent years. It may signalize the increase of clustering and of the importance of indirect search of the export markets.
- Indirect EU membership was significant only before Lithuania became the member of the EU. On the other hand, direct EU membership was important after entering the EU. These findings show that Lithuania's indirect contacts with the EU members before joining the EU could be crucial for getting the access to the EU market.

Overall the results confirm our hypothesis that cultural, language and historical factors (i.e. common spoken languages and Soviet bloc membership) were more important during the first years of the formation of Lithuanian exports network and the influence of the economic factors (i.e. GDP) increased gradually. Results also confirm the hypotheses about the increase of clustering, as the effects of the Soviet bloc and the free trade membership grow.

7 Robustness Checks

7.1. Taking the 11 most important products

To check the robustness of our results we have evaluated the same regression containing the data for only 11 products (HS 6-digits classification) which are the most important for Lithuania's exports. We assumed that the most important products are the ones that were exported to the largest number of markets.

The products used for the evaluation of the regressions are (sorted in the descending order by their importance) given in Table 7.

Table 7. The widest exported products of Lithuania in the period of 1995-2014.

No	HS92 6-digits code	Meaning
1	270300	Peat, incl. peat litter, whether or not agglomerated
2	901890	Instruments and appliances used in medical, surgical or veterinary sciences, n.e.s.
3	271000	Petroleum oils, etc, (excl. crude); preparation
4	382200	Diagnostic or laboratory reagents on a backing, prepared diagnostic or laboratory reagents whether or not on a backing, and certified reference materials (excl. compound diagnostic reagents designed to be administered to the patient, blood-grouping reagents, animal blood prepared for therapeutic, prophylactic or diagnostic uses and vaccines, toxins, cultures of micro-organisms and similar products)
5	940360	Wooden furniture (excl. for offices, kitchens and bedrooms, and seats)
6	350790	Enzymes and prepared enzymes, n.e.s.
7	392690	Articles of plastics and articles of other materials of heading 3901 to 3914, n.e.s (excl. goods of 9619)
8	852990	Parts suitable for use solely or principally with transmission and reception apparatus for radio-broadcasting or television, television cameras, digital cameras, video camera recorders, radar apparatus, radio navigational aid apparatus or radio remote control apparatus, monitors and projectors, n.e.s.
9	732690	Articles of iron or steel, n.e.s. (excl. cast articles or articles of iron or steel wire)
10	940390	Parts of furniture, n.e.s. (excl. of seats and medical, surgical, dental or veterinary furniture)
11	840999	Parts suitable for use solely or principally with compression-ignition internal combustion piston engine "diesel or semi-diesel engine", n.e.s.

All the regressions were evaluated using the same models as for the main regressions.

As the results (see: Appendix A3) are similar to the ones containing all the products, we can conclude that our results are robust.

As the third most important Lithuanian exports product, petroleum oil, is in it's nature not an exports product, but mainly re-exports, the same model was evaluated taking only the other 10 top products (without petroleum oil). The results were almost identical.

7.2. Taking the 10 most important Lithuania's export partners

We have also evaluated the same regression containing the data for only 10 major Lithuania's export partners. We assumed that the major export partners are the countries in which the largest number of products were exported (calculating the number of exported products according to HS 6-digits classification in all the years). These countries are (sorted in descending order): Latvia, Russia, Estonia, Belorussia, Germany, Poland, Ukraine, Denmark, Sweden and the UK. All the regressions were evaluated using the same models as for the main regressions.

As the signs of the results (see: Appendix A4) are similar to the ones containing all the countries, we can conclude that our results are robust. The differences in the numbers could be associated with the non-linearity of the distribution function and the results being distributed mostly in the tail.

8 Conclusions

Motivated by the recent findings that country's competitiveness and it's growth perspectives depend on the country's export structure, we investigated the factors that influence country's export structure which in turn influence country's economic situation and growth. As country's export network takes a long time to form and spreads gradually, we employed the theories of gravity modeling and network economics to examine the importance of network effects for the development of an extensive trade margin. In order to examine the spread of the export network from the very beginning we looked at a newly opened market which recently started to create its trade relations with the other countries of the world, and empirically analyzed which factors were the most important for the development of its extensive trade margin.

We confirm that the network variables are important for the spread of a country's export network. The growth of the extensive trade margin depends not only on the usual gravity model variables (e.g. distance between the origin and the destination markets, economic standing and growth of the destination countries, destinations' belonging to the same trade unions as the origin, etc.), but also on the extensive network effects between the current export partners of a country of origin (intermediate markets) and their own export partners (e. g. distance and ability to communicate between the intermediate markets and their own export partners, economic standing and growth of the export partners of the intermediate markets, etc.).

Our analysis shows that the factor that influence the spread of an extensive trade margin the most is the current trade relations of the country of origin (the wider the country's export's network is, the faster it grows), previous exporting to the same market, better economic standing of the destination market, ability to communicate between the export partners, the stock of emigrants and the distance (the latter negatively). We prove that both direct and indirect emigration and both direct and indirect ability to communicate significantly increase the chance to start exporting to a new destination market.

We find that as the market develops the importance of economic factors grows and the importance of direct belonging to various economic and trade unions tends to become less important. The importance of knowing more languages and being able to communicate directly with the trade partners is also increasingly important. Moreover, we find a strong evidence of an increase of clustering of the trading partners.

Our empirical findings suggest further empirical studies and several extensions. Firstly, we analysed only the extensive trade margin. Examining intensive trade margin would propose a new angle for the topic and possibly interesting results. Secondly, it would be interesting to model diversity of the products (e. g. if a country exports banana, perhaps it would be easier for it to start exporting apples also, but not cars). Thirdly, we may wish to include possible intermediaries, i. e. the countries from which the exports spreads to the other markets, as well. We leave these questions for future research.

9 References

1. Albornoz F., Pardo H.F.C., Corcos G., Ornelas E., 2012. Sequential Exporting. Journal of International Economics 88, 17–31.
2. Allen T., 2014. Information Frictions in Trade. Econometrica, Vol. 82, No. 6, 2041–2083.
3. Anderson J., 2011. The Gravity Model. Annual Review of Economics 3, 133–60.
4. Anderson J., Winkoop E., 2003. Gravity with Gravitas: A Solution to the Border Puzzle. The American Economic Review 93, No. 1, 170-192.
5. Acemoglu D., Carvalho V.M., Ozdaglar A., Tahbaz-Salehi A., 2012. The Network Origins of Aggregate Fluctuations. Econometrica 80, No 5, 1977-2016.
6. Broda C., Weinstein D.E., 2006. Globalization and the Gains from Variety. The Quarterly Journal of Economics 121(2).
7. Chaney Thomas, 2014. The Network Structure of International Trade. American Economic Review 104(11), 3600-3634.
8. Chaney T., 2016. Networks in International Trade. The Oxford Handbook of the Economics of Networks.
9. Feenstra R.C., Lipsey R.E., Deng H., Ma A.C., Mo H., 2005. World Trade Flows, NBER working paper 11010, 1962-2000.
10. Felbermayer G., Kohler W., 2006. Exploring the Intensive and Extensive Margins of World Trade. Review of World Economics 142, 642-674.
11. Gaulier G., Zignago S., 2010. BACI: International Trade Database at the Product-level. The 1994-2007 Version, No 2010-23 (http://www.cepii.fr/PDF_PUB/wp/2010/wp2010-23.pdf).
12. Handbook on Measuring International Migration through Population Censuses, United Nations, Department of Economic and Social Affairs Statistics Division, 7-10, March 2017 (https://unstats.un.org/unsd/statcom/48th-session/documents/BG-4a-Migration-Handbook-E.pdf).
13. Helpman E., Melitz M., Rubinstein Y., 2008. Estimating Trade Flows: Trading Partners and Trading Volumes. Quarterly Journal of Economics 123, No. 2, 441-487.
14. Hidalgo C.A., Hausmann R., 2009. The Building Blocks of Economic Complexity. Proceedings of the National Academy of Sciences of the United States of America 106, No. 26, 10570-10575.
15. Jackson M.O., 2008. Social and Economic Networks. Princeton University Press.
16. Jackson M.O., Rogers B.W., 2007. Meeting Strangers and Friends of Friends: How Random Are Social Networks?. American Economic Review, 97(3), 890-915.
17. Konrad B.B., Hassan T.A., 2013. The Economic Impact Of Social Ties: Evidence From German Reunification. The Quarterly Journal of Economics 128(3), 1219–1271.
18. Mayer T., Zignago S., 2011. Notes on CEPII's distances measures: The GeoDist database. CEPII, Working Paper No. 2011, 25 December.
19. Melitz M.J., 2018. Competitive effects of trade: theory and measurement. Review of World Economics 154(1), 1-13.
20. Melitz J., Toubal F., 2014. Native Language, Spoken Language, Translation and Trade. Journal of International Economics 93(2), 351-363.
21. Mundlak Y., 1976. On the Pooling of Time Series and Cross-section Data. Econometrica 46(1), 69-85.
22. Rauch J.E., 1999. Networks versus Markets in International Trade. Journal of International Economics 48(1), 7-35.

23. Rauch J.E., Trindade V., 2002. Ethnic Chinese Networks in International Trade. The Review of Economics and Statistics 84, No. 1, 116-130.
24. Silva J.M.C.S., Tenreyro S., Wei K., 2014. Estimating the Extensive Margin of Trade. Journal of International Economics 93, 67–75.

Appendixes:

Table AI. General regression excluding distance from the origin

Variables, normalized	PROBIT marginal effects		OLS
Dependent variable	Exporting, t+1	Exporting, t+1	Exporting, t+1
Markets, t	0.0339***	0.0165***	0.06555***
	(0.000178)	(0.000125)	(0.00050)
Exporting, t	0.0105***	0.00789***	0.13357***
	(0.000057)	(0.000035)	(0.00061)
GDP (log mill)	0.0257***	0.0291***	0.00896***
	(0.000190)	(0.000300)	(0.00026)
Import growth	0.00174***	0.00109***	0.00076***
	(0.000099)	(0.000081)	(0.00005)
Soviet block	-0.00124***	-0.000062	-0.00957***
	(0.000205)	(0.000133)	(0.00033)
EU membership	-0.00117***	0.000848***	0.00108***
	(0.000089)	(0.000073)	(0.00035)
FTA	0.00102***	-0.000241**	-0.00254***
	(0.000128)	(0.000105)	(0.00020)
Languages	0.0205***	0.00194***	0.03559***
	(0.000158)	(0.000086)	(0.00050)
Migrants (log)	0.00956***	0.00149***	0.01404***
	(0.000149)	(0.000079)	(0.00023)
Imp.growth f-o-f	0.00115***	0.000591***	0.00015***
	(0.000148)	(0.000121)	(0.00002)
Distance f-o-f	-0.0196***	-0.00370***	-0.04961***
	(0.000193)	(0.000103)	(0.00050)
Sov. Block f-o-f	0.00395***	0.00131***	0.00642***
	(0.000159)	(0.000104)	(0.00031)
EU f-o-f	0.00194***	0.00162***	0.00863***
	(0.000106)	(0.000096)	(0.00038)
FTA f-o-f	0.00393***	0.00237***	-0.00249***
	(0.000141)	(0.000119)	(0.00024)
Language f-o-f	0.000345**	0.000932***	-0.00854***
	(0.000139)	(0.000089)	(0.00035)
Migrants f-o-f	0.00545***	0.00366***	0.00065***
	(0.000130)	(0.000088)	(0.00013)
Constant	yes	yes	yes
Individual effects	random	correlated random	random
R-squared	-	-	0.5332
Observations	7,338,104	7,338,104	7,338,104
Number of ID	386,216	386,216	386,216

* Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

** All the variables in the regressions were normalized. For the OLS models regressions coefficients are given and for the Probit models marginal effects are presented. Random effects are by the destination and the product. Robust standard errors are given in parenthesis.

Figure AI. Comparison of the general regression results excluding distance from the origin

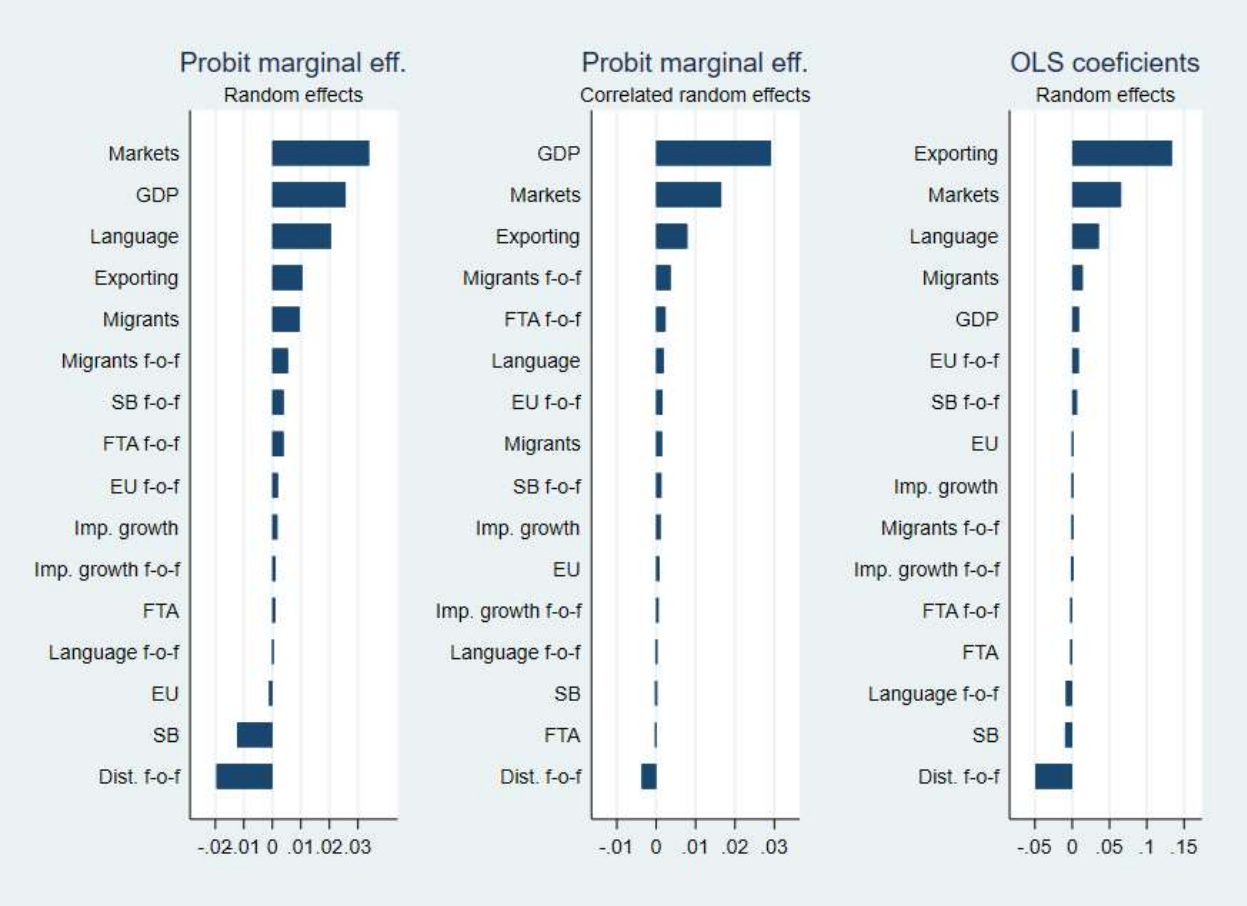


Table A2. Direct regression results excluding distance from the origin

Variables, normalized	**PROBIT marginal effects**		**OLS**
Dependent variable:	**Exporting, t+1**	**Exporting, t+1**	**Exporting, t+1**
Markets, t	0.0359***	0.0179***	0.06491***
	(0.000175)	(0.000121)	(0.00048)
Exporting, t	0.0102***	0.00776***	0.13449***
	(0.000056)	(0.000035)	(0.00058)
GDP (log mill)	0.0271***	0.0275***	0.00958***
	(0.000186)	(0.000293)	(0.00025)
Import growth	0.00224***	0.00130***	0.00080***
	(0.000082)	(0.000068)	(0.00005)
Soviet block	0.00330***	0.00303***	-0.00252***
	(0.000127)	(0.000082)	(0.00019)
EU membership	-0.000195***	0.00182***	0.00720***
	(0.000066)	(0.000053)	(0.00024)
FTA	0.00368***	0.00139***	-0.00408***
	(0.000097)	(0.000088)	(0.00013)
Languages	0.0214***	0.00286***	0.02812***
	(0.000132)	(0.000064)	(0.00025)
Migrants (log)	0.0106***	0.00223***	0.01475***
	(0.000151)	(0.000077)	(0.00023)
Distance f-o-f	-0.0222***	-0.00521***	-0.04985***
	(0.000193)	(0.000099)	(0.00048)
Constant	yes	yes	yes
Individual effects	random	correlated random	random
R-squared	-	-	0.5326
Observations	7,338,104	7,338,104	7,338,104
Number of ID	386,216	386,216	386,216

* Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

** All the variables in the regressions were normalized. For the OLS models regressions coefficients are given and for the Probit models marginal effects are presented. Random effects are by the destination and the product. Robust standard errors are given in parenthesis.

Figure A2. Comparison of the direct regression results excluding distance from the origin

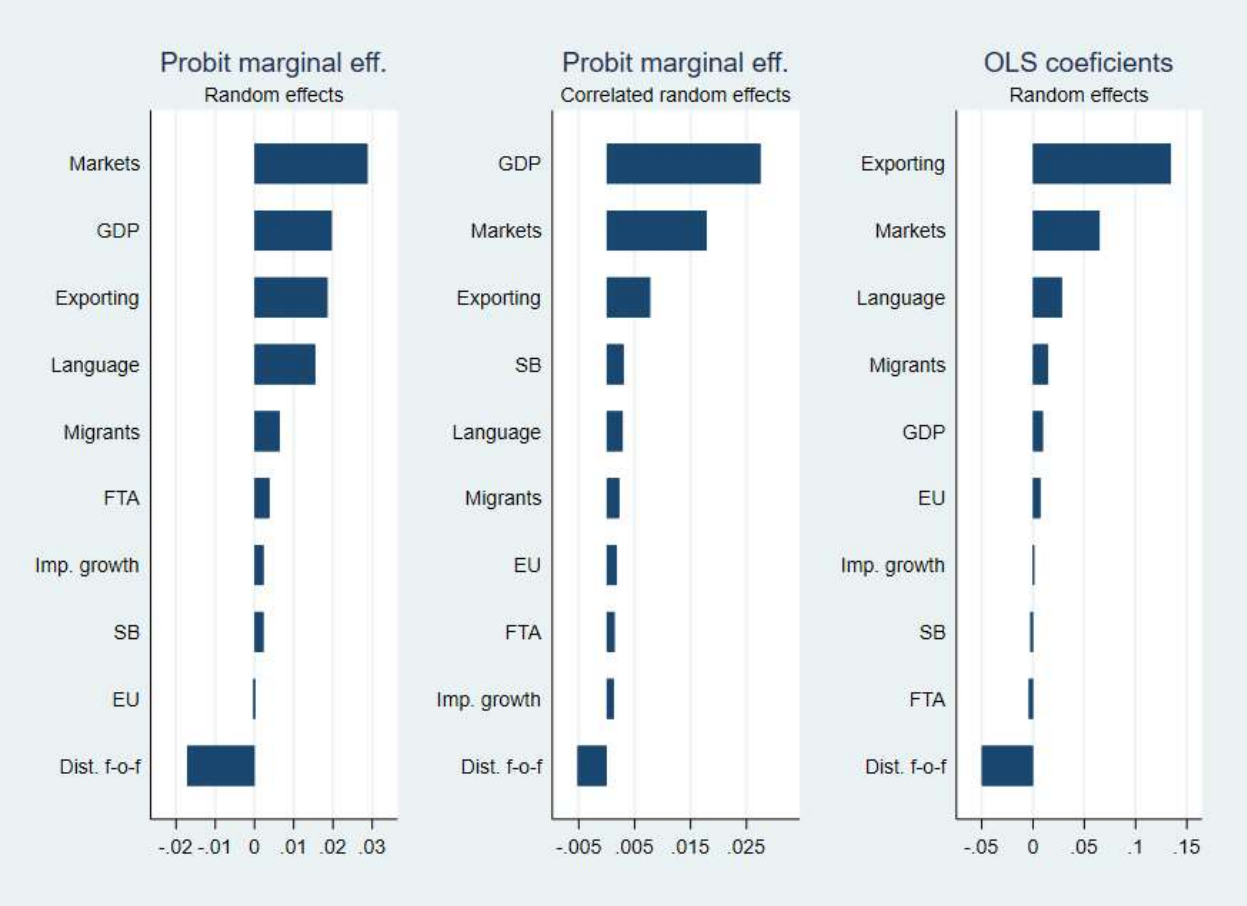


Table A3. Results of the regressions taking the 11 most important Lithuania's export products

Variables, normalized	PROBIT marginal effects			OLS		
	General reg.	Direct reg.	Indirect reg.	General reg.	Direct reg.	Indirect reg.
Dependent variables:	Exporting, t+1	Exporting, t+1	Exporting, t+1	Exporting, t+1	Exporting, t+1	Exporting, t+1
Markets, t	0.0410***	0.0236***	0.0507***	0.03405***	0.00932***	0.02571***
	(0.00421)	(0.00128)	(0.00270)	(0.00367)	(0.00088)	(0.00157)
Exporting, t	0.0279***	0.0293***	0.0305***	0.14034***	0.14386***	0.14881***
	(0.00127)	(0.00126)	(0.00135)	(0.00238)	(0.00233)	(0.00098)
GDP (log mill)	0.0921***	0.106***	0.114***	0.05544***	0.06270***	0.06310***
	(0.00437)	(0.00396)	(0.00434)	(0.00340)	(0.00303)	(0.00209)
Import growth	0.00666*	0.00556***		0.00524***	0.00315***	
	(0.00374)	(0.00172)		(0.00137)	(0.00109)	
Distance (origin)	-0.0153*	-0.0519***		0.01518**	-0.02868***	
	(0.00805)	(0.00431)		(0.00593)	(0.00335)	
Soviet block	-0.0241***	-0.00401		-0.02116***	-0.00494**	
	(0.00694)	(0.00375)		(0.00626)	(0.00230)	
EU membership	-0.00236	-0.000693		-0.00035	0.00624***	
	(0.00314)	(0.00247)		(0.00290)	(0.00202)	
FTA	-0.00925***	0.00754***		-0.01860***	0.01527***	
	(0.00333)	(0.00231)		(0.00355)	(0.00230)	
Languages	0.0336***	0.0329***		0.03199***	0.02348***	
	(0.00492)	(0.00400)		(0.00380)	(0.00252)	
Migrants (log)	0.0403***	0.0442***		0.03777***	0.03815***	
	(0.00447)	(0.00469)		(0.00377)	(0.00375)	
Imp.growth f-o-f	-0.00346		0.0133***	-0.00687**		0.00396
	(0.0105)		(0.00515)	(0.00323)		(0.00335)
Distance f-o-f	-0.0158***		-0.0270***	-0.02243***		-0.01637***
	(0.00429)		(0.00252)	(0.00365)		(0.00132)
Sov. Block f-o-f	0.0381***		0.0605***	0.02750**		0.03820***
	(0.0122)		(0.00632)	(0.01176)		(0.00307)
EU f-o-f	0.00374		-0.000345	0.01474**		0.00886***
	(0.00587)		(0.00491)	(0.00631)		(0.00287)
FTA f-o-f	0.0359***		0.0301***	0.05871***		0.03531***
	(0.00571)		(0.00427)	(0.00569)		(0.00273)
Language f-o-f	0.00214		0.0467***	-0.00808**		0.02804***
	(0.00604)		(0.00480)	(0.00386)		(0.00204)
Migrants f-o-f	0.0341***		0.0556***	0.00868**		0.02006***
	(0.00645)		(0.00698)	(0.00375)		(0.00312)
Constant	yes	yes	yes	yes	yes	yes
Individual effects	random	random	random	random	random	random
R-squared	-	-	-	0.6642	0.6616	0.6571
Observations	38,456	38,456	38,456	38,456	38,456	38,456
Number of ID	2,024	2,024	2,024	2,024	2,024	2,024

* Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

** The table presents the estimation results of the model taking the data of 11 products that Lithuania exports to the largest number of markets. All the variables in the regressions were normalized. For the OLS models regressions coefficients are given and for the Probit models marginal effects are presented. Random effects are by the destination and the product.

Table A4. Results of the regressions taking the 10 most important Lithuania's export partners

Variables, normalized	PROBIT marginal effects			OLS		
	General reg.	Direct reg.	Indirect reg.	General reg.	Direct reg.	Indirect reg.
Dependent variable:	Exporting, t+1	Exporting, t+1	Exporting, t+1	Exporting, t+1	Exporting, t+1	Exporting, t+1
Markets, t	0.204***	0.13991***	0.204***	0.17580***	0.08657***	0.14771***
	(0.00254)	(0.00190)	(0.00265)	(0.00190)	(0.00241)	(0.00152)
Exporting, t	0.0476***	0.05004***	0.0521***	0.10589***	0.11572***	0.11810***
	(0.000513)	(0.00047)	(0.000550)	(0.00062)	(0.00081)	(0.00036)
GDP (log mill)	0.143***	0.13123***	0.00799***	0.11858***	0.07501***	0.00546***
	(0.00388)	(0.00406)	(0.00213)	(0.00301)	(0.00318)	(0.00104)
Import growth	0.00578***	0.01201***		0.00840***	0.01281***	
	(0.00131)	(0.00085)		(0.00128)	(0.00081)	
Distance (origin)	-0.0871***	-0.10811***		-0.05243***	-0.06084***	
	(0.00302)	(0.00313)		(0.00238)	(0.00206)	
Soviet block	-0.0271***	-0.03628***		-0.00310	-0.02280***	
	(0.00280)	(0.00235)		(0.00234)	(0.00141)	
EU membership	-0.0113***	-0.00283***		-0.01059***	-0.00015	
	(0.000777)	(0.00069)		(0.00066)	(0.00069)	
FTA	0.0285***	0.03493***		0.02663***	0.02778***	
	(0.00153)	(0.00147)		(0.00124)	(0.00116)	
Languages	0.165***	0.17002***		0.11471***	0.11149***	
	(0.00220)	(0.00210)		(0.00176)	(0.00201)	
Migrants (log)	0.00960***	0.00738**		-0.00124	0.00859***	
	(0.00272)	(0.00293)		(0.00201)	(0.00198)	
Distance f-o-f	-0.161***		-0.104***	-0.18652***		-0.11146***
	(0.00633)		(0.00723)	(0.00356)		(0.00301)
Imp.growth f-o-f	0.0128***		0.0169***	0.00919***		0.01568***
	(0.00222)		(0.00150)	(0.00237)		(0.00138)
Sov. Block f-o-f	0.00183		0.0344***	-0.00335**		0.02154***
	(0.00132)		(0.00106)	(0.00132)		(0.00069)
EU f-o-f	0.0104***		0.0133***	0.00857***		0.00753***
	(0.000970)		(0.000775)	(0.00091)		(0.00056)
FTA f-o-f	0.00468***		0.00605***	-0.00679***		-0.01716***
	(0.00133)		(0.00137)	(0.00130)		(0.00101)
Language f-o-f	0.0295***		0.0153***	0.04513***		0.04195***
	(0.00137)		(0.00125)	(0.00142)		(0.00081)
Migrants f-o-f	0.00115		0.0188***	0.00618***		0.02256***
	(0.00145)		(0.00146)	(0.00126)		(0.00094)
Constant	yes	yes	yes	yes	yes	yes
Individual effects	random	random	random	random	random	random
R-squared	-	-	-	0.4484	0.4496	0.4300
Observations	398,810	398,810	398,810	398,810	398,810	398,810
Number of ID	20,990	20,990	20,990	20,990	20,990	20,990

* Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

** The table presents the estimation results of the model taking the data of 10 largest Lithuania's export partners. All the variables in the regressions were normalized. For the OLS models regressions coefficients are given and for the Probit models marginal effects are presented. Random effects are by the destination and the product. Robust standard errors are given in parenthesis.