Erasmus University Rotterdam

MSc in Maritime Economics and Logistics

2012/2013

Determinants of Bilateral Liner Shipping Connectivity

By

[Jan-Willem van Hoogenhuizen]

Acknowledgement

Here before you lies the cumulative result of studying the master Maritime Economics and Logistics at the Erasmus University Rotterdam in the form of this thesis. I will use this opportunity to thank those whose help and assistance where invaluable for me in achieving this result.

First of all, my gratitude goes towards my supervisor Dr. Jan Hoffmann for his assistance and advice in writing this thesis. Without his valuable comments this thesis would not be what it is now. I would also like to express my special thanks to those at UNCTAD for their valuable help and support during my internship with an extra show of gratitude for Bismark.

Next to this, I also feel thankful for the whole academic body of Maritime Economics and Logistics for the many things they have learned me over the last year. A special mention goes to the late Dr. Eelco van Asperen, who stimulated my enthusiasm for the maritime sector and whose feedback on my thesis proposal was highly valued. My thanks are also extended to those in the MEL office for their valuable support and assistance over the year.

And as no acknowledgement can pass without using the opportunity to show ones gratitude to his family, I would like to use this opportunity to thank them (in particular my parents) for all their support they gave me in few years as a student.

Furthermore, if I have omitted to mention anyone here who should have been mentioned, I would like to thank him or her now as well for all the assistance, advice and support given to me.

Abstract

In this paper, the relations between bilateral liner shipping connectivity and its determinants are investigated. It is also the first paper ever in the knowledge of this author that links the structure of the global liner shipping network to the concept of connectivity and where the metrics that define the structure were used as predictors for bilateral trade.

In order to address the above stated research purpose of this paper a discussion and description of the global liner shipping network structure is given first and combined with the concept of connectivity. Following this combination, a set of possible predictors of trade have been identified that can potentially act as a proxy of connectivity. Subsequently, these predictors have been tested with the gravity equation of trade using the BV-OLS model by Baier and Bergstrand. After this analysis, a set of potential components for a index that can indicate the bilateral connectivity in liner shipping (the LSBCI) have been selected in order to generate a few possible LSBCI's. These indices have than been tested against bilateral trade in order to draw some conclusions about them.

By defining the structure of the liner shipping network in terms of nodes and edges, where the nodes represent the countries and the edges the direct connections between these countries it becomes possible to determine a set of metrics that represent the structure of the liner shipping network (such as the degree centrality and cluster coefficients). Furthermore it can be assumed the location of all the nodes in a network is the result of all the forces acting upon them. When applied to the liner shipping network, these forces can vary from the size of the nodes economy to a nodes institutional differences with another node.

Liner shipping connectivity itself can be considered to be the measure of access to regular liner shipping services. When this is applied to the concept of the network structure, it can be concluded that bilateral connectivity is this an indication of the relative distance between two nodes in a network. Therefore, in order to get an indication of connectivity. Thus, the components that can form a BLCI need to be a proxy of the forces acting between the nodes in the network. The potential components that have been identified in this paper are the betweenness centrality, degree centrality, nodal strength, cluster coefficient, the K-core, the quality of ports and the LSCI, an index for liner shipping connectivity at country level.

However, in the small set of LSBCI's generated with the above components there were no significant differences, indicating that only one component that represents the liner shipping structure should be sufficient and that other components that capture other aspects of liner shipping should be included.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Background

For international trade, connectivity matters. This is for instance illustrated by the Director-General of the WTO, Pascal Lamy in July 2013. Here, he mentioned in a speech during the opening of the fourth global review of Aid for Trade the importance of connectivity: By ‘connecting’ the least connected countries to the international value chains, these countries should become connected to the global trading network (WTO, 2013). The first question that arises after hearing this statement, is what being connected actually encompasses and how connectivity can be created. Lamy mentioned in his speech that as our trading system is increasingly based on trade in tasks and not in final goods, one can take part in the global system by excelling at one task, and then by slowly adding more to ones repertoire (Ibid.). According to Farole (2013), integration with regional and international markets is critical as peripheral regions may have significant unexploited trade related opportunities to build economic agglomeration that may be thwarted by trade barriers.

Another aspect mentioned by Lamy on being connected is that you need a good functioning window in order to be able to have access to a system if you intend to make use of it. If your customs or other border services are inefficient, you’re easy of accessing the global network is decreased, leading to a lower connectedness with said network (Ibid.).

One of the main reasons for Lamy to talk about having access to the international trading system is that it might help countries and regions in the world increase their participation in international trade: the idea is that by improving the access to the network or a system, more people and firms are inclined to make use of it as they will have less barriers to enter to this network or system, which in turn reduces the costs (both financial and non-financial) of making use of it.

However, in order to know how to facilitate the access to a system, first, it has to be known precisely what factors determine having or not having access to the system. In order be able to find these determinants, the system itself and its structure must in turn be analyzed and described.

If the example is taken of the maritime sector of the international trade network, the first characteristic that is relevant to know is that demand for transport is increasing yearly with remarkable rates despite a few ‘hiccups’ in the growth as a result of the global and financial crisis starting in 2008 (UNCTAD, 2012). As a result of ever increasingly lower (maritime) transport costs, in part from the huge overcapacity in the world merchant fleet, it is nowadays much more economically attractive to produce products overseas. This has to led interesting situations, where specific parts of a finished product can be produced (almost) anywhere in the world (and thus usually where the production is cheapest) before being shipped to the final destination.

As a result of this, firms that produce lower value manufacturing goods that require lower labour skills and lower capital investments keep relocating and moving towards low wage countries (Ibid.). More developed countries have thus the tendency to move up the value chain, presenting the opportunity for lesser developed countries to become a part of the global value chain. However, in order to be part of it, a country must have access to the global transport and trade network, making the issue of connectivity again a relevant one.

As a result of the above described mobility in goods and firms, the demand for transportation of (semi) manufactured goods has increased substantially over the last decade, resulting in its high share of total international seaborne trade. One of the main facilitators of the increase in trade and transportation of manufactured goods has been the container. By placing the products in a container and then transporting the container instead of the individual goods itself, the transport costs decreased. It is even claimed that containers have been more important for the process of globalization that freer trade (The Economist, 2013).

The structure of the maritime transport in containers is such that it can be compared to a regular bus schedule: vessels leave and arrive from and at ports at a fixed frequency. Therefore, the maritime transportation network of containers (and thus of (semi) manufactured goods is described as a liner service.

Conclusively, it can thus be stated that with the current high share of international seaborne trade for manufactured goods, access to the global market for manufactured goods depends on the access to the international liner shipping services. The concept of the amount of access to these international shipping services can be defined as liner shipping connectivity (Hoffmann et al., 2013).

An issue arising from this concept of connectivity is how to determine differences in the connectivity and how these differences can be explained; what are the determinants for these differences in access to the international liner shipping services? Once they have been identified and quantified, they can be used together as the components of a connectivity index that can then be used to compare the connectivity of different regions and can also be used to help policymakers generating access to the global liner shipping network.

Such a connectivity index is created annually by UNCTAD since 2004 for the country level. This index, the Liner Shipping Connectivity Index (LSCI) consists out of several components that together form a representation of the connectivity of a country to the global liner shipping network: the higher the index, the easier it becomes to participate in international maritime containerised trade from that country. Previous research has shown a significant and strong relation between the components of this index and the transport costs (Wilmsmeier& Hoffmann., 2008).

The next step in improving our understanding of liner shipping connectivity is by looking at the country pair connectivity: a measure of how well two countries are connected with each other in the global liner shipping network. In other words: what is the measure of accessibility of any country j from any country i in the global liner shipping network. This is what UNCTAD has been working on in the recent period.

The intention is there to expand the LSCI to bilateral or country pair index: a Liner Shipping Bilateral Connectivity Index (LSBCI). In order to be able to create such an index, potential determinants of this bilateral connectivity have to be determined first. And in order to be able to understand the underlying concepts of bilateral liner shipping connectivity, these potential determinants have to be further analysed. In order to do this, potential bilateral relations between countries that can represent their connectivity have to be tested.

1.2 Problem Statement

Following from the above, the purpose of this thesis is to identify and quantity potential determinants of bilateral liner shipping connectivity by looking at bilateral relations and factors that either affect, or at least have a relation with, the supply of liner shipping services. This results if the following main research question:

<u>What are the relations between bilateral liner shipping connectivity and its determinants?</u>

The scope of this research is as follows: It is focused on bilateral maritime liner shipping for container trade. This means that the determinants for the connectivity of liner shipping for pairs-of-countries (or routes between 2 countries) have to be identified and quantified.

First of all, in order to be able to answer the main research question, several sub questions and issues have to be addressed. As described earlier in this introduction, in order to determine what defines having access to a system or a network, the network or system itself has to be analysed. Next to this, the concept of connectivity needs to be further clarified in order to identify parameters that might act and work as its proxy's and can thus be used as its determinants. When the above two issues have been addressed, they can be used to help determine possible proxies of connectivity of bilateral international seaborne trade to be tested for relations with maritime trade in manufactured goods. From this list of potential determinants, it can then be determined which of these parameters can represent or determine the measure of accessibility to the global network.

Therefore, the following sub research questions will be addressed in this thesis:

1. What is the structure of the liner shipping network?
2. What is liner shipping connectivity
3. What are potential proxies or determinants of liner shipping connectivity?
4. What are the relations of the possible determinants with international maritime trade between two countries?

On a final note, this research will help determine how the access to the international liner shipping services can be improved (this follows from the identification of the determinants) and it will also improve the existing scientific knowledge on the drivers of international trade (and thus helps improve the gravity model of trade). Thus, this research will help countries improve their access to the global market, which will be especially beneficial for developing countries.

1.3 Structure.

The structure of this thesis is as follows: in chapter two a theoretical framework is given that discusses the existing scientific literature on the relevant topics. Using this framework, the pillars of trade and resulting potential determinants of bilateral connectivity can be derived.

In chapter three the methodology is discussed. Here, the models that are used are given and discussed, and the data is described together with its sources. Furthermore, in this chapter the validity, reliability and limitations of this research are discussed.

In chapter four, the liner shipping itself and its network structure is described in order to answer the first sub research question. Here, some numbers are given on the existing bilateral maritime trade flows that together form a representation of the global liner shipping network.

Following this chapter, the connectivity is discussed in chapter five. Here, the concept of connectivity is addressed in more detail and the LSCI and its components and implications are discussed. This chapter ends with a further discussion on the LSBCI and a description what the index should represent. of the currently proposed components of this index. Furthermore, the results of the index using these proposed components are given by showing and discussing the index itself and its relation with seaborne trade in manufactured goods.

In chapter six, the identified pillars of international seaborne trade containerized goods are discussed and further investigated for their relation to trade volumes and measures of bilateral connectivity. These four pillars are: 1) the network structure, 2) ports and infrastructure, 3) policy 4) institutions and setting. In chapter seven are the variables that were identified as potential components of the LSBCI used to generate a set of possible indices. In the last chapter, eight, the conclusions, where the main research question is answered and the resulting recommendations are given and discussed. Following this, the thesis ends with the references and appendices.

Chapter 2 Theoretical Framework

In this chapter, the existing scientific literature is discussed that is relevant for the topic of this thesis. This framework can then be used to help determine the possible components for bilateral liner shipping connectivity.

2.1 Trade Costs

A decline in maritime transport costs has a substantial and positive impact on trade and particularly on the intensive margin of trade through an increase of the average price of traded goods (Bensassi et al., 2011). Therefore, trade volumes should be expected to increase for decreasing transport costs.

Several authors have discussed and argued that the remoteness to the main international liner shipping networks has a higher effect on maritime transport costs than the actual geographical distance (Figueiredo De Oliveira, 2013). Furthermore, Wilmsmeier and Hoffmann (2008) have shown that variables related to port infrastructure, service quality and LSCI together explain almost sixty percent of freight rate variance (sourced from Figueiredo De Oliveira, 2013). Next to these factors other authors have shown that the number of carriers also influences the freight rates as it affects the competition on a specific route (Ibid.).

Another variable that is often used to explain trade costs is distance. It is assumed that freight rates will increase as the distance between two markets increases. However, as Figueiredo De Oliveira (2013) states, in the literature it is mostly only used assuming that maritime transport between two countries is made only by direct services. As the liner shipping network is constructed in such a way that for a majority of routes at least one transshipment is required, it implies that the number of transshipments should also be taken into account in maritime trade models as this affects the effective maritime distance between the origin and destination.

Next to this, the transport costs also tend to vary per vessel as their characteristic also influence their fuel use. Because of this, vessels that are employed on the major routes will have an advantage in terms of costs as they are usually the biggest and newest vessels (Ibid.), whilst the older vessels are cascaded down to other (however, physical restrictions on certain trade routes in sea canals or port which will limit the possibility of using ever larger ships. These physical restrictions cannot be removed in the short run, which means that the cascade effect is necessarily limited. (Maatsch, 2010)), less heavily employed shipping routes. Therefore, it is likely that trade costs will also depend on the use of each particular shipping route: the higher the demand for a route, the higher the likelihood that new and efficient vessels are employed on said route, and the lower the actual transport costs.

In other research it is shown that variance of freight rates across routes results from the characteristics of the routes in terms of fuel cost, amount of transshipment, trade volume, connectivity and port infrastructure. All these factors mean thus indeed that the freight rates applied on the major routes tend to be lower than those on secondary ones. (Figueiredo de Oliveira, 2010).

Willsmeier et al. (2006) investigate several port characteristics that can act as determinants of international transport costs. They found that port efficiency, port infrastructure, private sector participation and inter port connectivity have significant impacts on maritime transport costs. In their article, they also state that earlier research has shown a strong and significant impact of port efficiency on bilateral trade flows.

Bensassi et al. (2011) found that political actions in order to protect maritime sea lines to spur competition in the maritime transport industry, innovations helping increasing the size of merchants ships, and limiting the consumption of fuels by ships all have an impact on the volume and the composition of international trade that is together more strongly related to transport cost than approximations by use of geographical distances between countries.

From all of the above it follows that not only the actual geographical distance matters as a determinant of trade and trade costs, but the 'economic distance' ; the closer a route is to the 'center' of the network structure and the better the access to the center of the structure, the lower the transport costs, and thus the trade costs.

2.1 Pillars

It becomes clear from reviewing the literature on both seaborne and general international trade that the focus of the research on the relation between trade and its potential and possible determinants is split up in several main topics. Those that are relevant for this paper are the shipping network structure, policy, ports and infrastructure and finally a pillar named institutions and international that encompasses a multitude of factors such as culture, geographical location (such as borders) and institutions. Below, the literature is discussed for each of these four pillars.

2.1.1 The network

The global liner shipping network is not evenly distributed, resulting in a centrality of the network in certain geographical areas (Hu & Zhu, 2008). As a result of this imbalance in the distribution of the network, it can be expected that a multiplication of service network types is most likely to develop as this will provide the best value attributes in dealing with global supply chains (Notteboom & Rodrigue, 2008).

According to Nazemzadehet al., (2013) firms have changed their strategies with regards to their supply chains and transportation as a result of the changes in the global environment, such as the increased oil prices,the increasing threat of pirates in certain regions in the world but also the changing of the structure of the global economic hierarchies. One can think of the emergence of China and India as main industrial centers combined with the decline in the industrial strength of Europe (Ibid).

Due to the high concentration of the shipping market, the market structure is not consistent with perfect competition on the majority of routes. This leads to the increasing (bargaining) power of shipping alliances resulting in an increased importance and influence of the decisions made by these alliances (Ibid.). However, according to Luo et al (2009), it is still very much debated in the academic world whether the level of concentration in the market is sufficient for it to enter into a monopoly, with some authors claiming that the market is already operating in an oligopoly market (see for instance Sys (2008)), whilst others argue that the accessibility of the market to new or potential competitors that determines its contestability, leading them to the conclusion that the shipping market is still highly contestable (see for instance Haralambides (2004)).

The issue of the high concentration also affects the structure of the global liner network whose main determinants are the ship operators who are the main decision makers on port maritime connections, shipping route formation and network configurations. However, shippers and freight forwarders also influence this decision (Nazemzadehet al., 2013). According to Cullinane and Wang (2006), the ship operators are constantly using their greater bargaining power by demanding ports to have higher performance levels, lower costs and better service quality.

The vessel operator is mainly interested in selecting the route which represents their least total transportation costs between their origin and destination (Ungo& Ducruex,2013). However, shippers do not only look at the transportation costs, but also at the total landed costs which are the total costs of a product incurred on moving it from origin to its final destination. Therefore, for shippers, costs such as inventory carryingcosts, handling costs, value of frequency are also relevant next to the freight rates and direct transport costs only as is usually only the case for the vessel operators (Ibid.).One example that illustrates the above is the introduction of slow steaming, where the vessels sail at lower speeds in order to deal with the increased bunker prices. One of the other results of slow steaming however is that the rotation time of the vessels is increased, leading to a reduction in the average number of miles performed in a year and in a decrease of port of calls (Cariou, 2010) This in turn leads thus to increased inventory costs as each container spends more time at sea.

When this is combined with the statement that the operators are the main determiners of the routes it follows that as they choose the network configuration such that their transport costs are minimized, this does not mean that the structure of the network is optimal from an overall logistical perspective: it is in fact more likely that that is the opposite.

Yap and Notteboom (2011) show that the outcome for container port competition was mainly influenced by transshipment and hubbing decisions of shipping lines in Asia. They attribute this to the central location of several of its ports with respect to the containerized cargo flows between East Asia, South East Asia, Australia, the Middle East and Europe. They also state that the dominant position of Singapore could be threatened with the emergence of two nearby ports who are trying to facilitate transshipments in order to attract the big shipping lines and alliances.

Trade imbalance between regions also affects trade freights. However, the impact of trade imbalance is difficult to measure and generally this effect is underestimated. Thus, one should wonder whether the import-to-export ratio is the best way to measure imbalances. (Figueiredo De Oliveira, 2013).

The increase in complexity in the system resulting from issues such as trade imbalances, rising oil prices and security issues will also add risks to the supply chain and will lead to a reduction in reliability of the network (Ibid.). The resulting dynamics of the system will lead to shipping lines to base their decisions to call at ports on a large variety of commercial and operational determinants. An example of these are the cargo generating potential of a port, the distribution and pattern of the cargo flows over the ports hinterland and the port's nautical access (Yap &Notteboom, 2011).

Yap and Notteboom (2011) also state that ports that are able to adopt to the process of integration with the liner shipping industry and that add value to the commercial, operational and strategic pursuits of shipping lines will be rated as being more competitive and become more attractive as a port of call. In their paper, they also identify a number of factor that determine port choice for shippers, including distance to the market, port chargers, cargo volume, level of port efficiency, quality of infrastructure, schedule reliability, connectivity and several others. Notteboom (2011) states that all main shipping hubs have several common characteristics, such as proximity to important hinterlands and that they are locating along main navigation routes or at the crossing points of main North-South and East-West routes. This is confirmed by the research of Caldeirinha et al. (2013).

2.2.2 Policy

There are also many authors that address the effects of policy and trade facilitation on trade volumes and trade cost itself. Whilst many authors have studied the effects of transport infrastructure on transport costs and trade in developed and developing countries, the number of studies that focus on trade facilitation itself is limited. (Martinez Zarzoso et al. 2011).

Martinez Zarzoso et al. (2011) give several definitions that exist of trade facilitation in the literature. They range from 'the simplification and harmonization of international trade procedures to definitions that focus on the effects of single measures that might facilitate trade. The areas that trade facilitation should focus on also differ per definition. Some only include port efficiency or for instance trade procedures, whilst others focus on a wide variety of measures (Ibid.).

In general, trade facilitation tends to focus on the reduction of trade barriers. Research has shown that trade facilitation variables are in relative terms more important than tariffs. Therefore, increasing trade facilitation would lead to an increase in world trade (Ibid.). The extent of this effect however, will likely vary per region and country as a result of regional institutional differences.

Lengthy procedures for exports and imports reduce the probability that firms will enter export markets for time sensitive products (Ungo & Ducreux, 2013). Therefore, in order to attract trade, the number of procedures required for import and export should decrease. This can be explained by the fact that container transport is one of the least captive cargo types resulting in the need for ports and governments to make the necessary (both monetary and non monetary) investment in order to create the potential for a market share in their ports (Zondag et al., 2010).

Fageda& Gonzalez-Aregall (2013) investigated the effect of port charges on trade flows. They found that ports that charge higher port dues attract less traffic and that ports operating more international regular lines have higher revenues per ton. They also found that if there is a significant number of nearby ports, the revenues per port tonne decreases, implying that the availability of alternatives reduces market power.

Rodriguez and Rodrik (2001) come to the conclusion that countries with lower policy induced barriers to international trade have a faster growing economy by using cross country growth regressions. The reason they found for this faster economic growth is that lower barriers means more trade and thus access to a bigger market.

2.2.3 Ports and infrastructure

There exists an immense competition between shipping lines to attain global coverage of destinations and to expand their networks beyond the seas towards the mainland. As a result of this, the choice of port is essential and affects the service level and service quality. This in turn, leads to the increasing importance of port choice by shipping lines, as this decision determines whether the firm can realize its operational, service and performance goals (Panayides and Polyviou, 2010).

As mentioned before, the actors that have the most influence on port choice are the vessel operators. However, according to Nazemzadeh et al (2013), the bulk of existing literature on port choice models is dedicated to port choice by shippers. Ship operators' port choice behavior has been recognized by few authors. In a survey performed by Nazemzadehet al (2013) the authors found that shippers and shipping companies have similar port selection criteria, a similarity that could be explained due to the fact that many shippers apply carrier haulage system contracts, which lead them to have the same view as shipping companies.

Another determinant in port choice is found by Garcia-Alonso & Sanchez-Soriano (2009). They state that physical proximity of the port to clients is less important for the attraction of traffic to a certain port than achieving scale economies and reducing the time necessary to offer door to door services. Thus, as mentioned before, the geographical distance seems to be less important that the ‘economic distance’.

The same authors also found that some ports are more connected to their hinterland than other ports. According to them, the evolution of a ports activity often matches the strategic position of the ports in terms of its centrality in the network: it depends on the balance between the dynamics of the ports hinterland and on their matter of inclusion in the route of shipping lines (Ibid.).

Wang and Cullinane (2008) found that the level of port throughput at any container port is significantly and positively correlated to its accessibility. They therefore state that this accessibility can be treated as one of the indicators that directly reflect ports competitiveness.

An issue that results from the increased container traffic growth for container ports is congestion,demanding more investments in new terminals and also leading to increased inter and intra port competition as almost all ports try to become a transshipment hub (Caldeirinha et al., 2013). Inland access is also import for a ports performance (Ibid.) as it increases the size of the ports uncontested market, and thus allows it to ‘tap into’ a greater market. Therefore, the port requires a good and high quality integration with the main land transport networks, and thus also requires a good quality infrastructure in its hinterland itself (Ibid.)

Furthermore, as a result of the high oil prices and low freight rates, international shipping lines have increased their focus on the efficiency of their ship operations. One of their measures is slow steaming where the vessels sail at lower speeds in order to reduce bunker usage and to decrease the overcapacity issue in the market. However, as a result of this, in order to maintain similar vessel productivity levels, the port stay time has to decrease, leading to an increasing pressure on ports to perform faster and more efficient.

Castillo-Manzano and Gonzales-Laxe, (2011) found in their research into the evolution of commodity flows in Spanish ports and their hinterlands that the share of the traffic that these ports manage to capture depends on their share in the disputed hinterlands. The capture of this traffic is closely linked to each port's particular characteristics, including location, and also depends on how the local economy is evolving. However, there is no correlation with either the port having good intermodal port-train connections, or with the dynamics of port traffic that does not originate in the hinterland. This implies that a ports share of its hinterland is only determined by factors directly influenced by or influencing its hinterland and not by exogenous factors such as the amount of transshipment of feeder traffic.

As a final note, Figueiredo de Oliveira (2010) also found that a higher standard of infrastructure for both ports and the country itself in a country leads to reduced freight rates.

From this section, it can thus be concluded that multiple port selection criteria stem from studying the existing literature. In a research on the existing literature on this topic, Panayides and Polyviou (2010) come to the conclusion that port facilities, port Efficiency, costs, service, intermodal links, value-added services and information availability affect port selection as they all determine the ports share of the overall maritime transport costs.

2.2.4 Institutions and Setting

Trade policy barriers are only one element of the overall trade costs. Due to the decrease in the influence of tariff barriers on trade, institutional barriers may be increasing in relative importance. Results indicate that institutional trade barriershave a greater impact on trade flows than tariff barriers. According to these findings, trade policy negotiation efforts should focus on facilitating trade processes and should be at the forefront of multilateral negotiations (Ramos, 2006).

"Institutions are the humanly devised or socially constructed sets of rules that constrain and enable human interaction". These rules can be both formal, such as the law or certain conventions, and informal such as habits or taboos. In the literature three pillars of institutions are often distinguished: the regulative, the normative and the cultural-cognitive, where each has its own unique sets of logic, mechanisms of enforcement, and bases of compliance and legitimacy. These three pillars should not be understood as mutual exclusive; they might interact and to large extent influence and form each other (Jacobs et al., 2011). Institutions are not static but develop over time. Some claim that this will eventually result in an optimal equilibrium where it is in nobody's interest to change it whilst others claim that institutional change is less predictable. (Ibid.).

The institutional environment refers to all kinds of informal conventions, customs, routines and norms on the one hand, and formal sets of legally enforced rules and regulations on the other. The institutional arrangements refer to particular organizational forms (firms, state bureaucracies, cooperative networks, or, more generally: governance systems) that arise as a consequence from, and whose constitution and operations are governed by, the institutional environment (Jabos et al. 2011).

According to Sou et al. (2013), many authors have over the past few decades researched the impact of the establishment of free trade agreements between countries. Although these studies have allowed for the modeling of these effects, they do not provide any insight on how policies such as these agreements can affect maritime trade. This is the same for many potential predictors of trade: although relations can be proven, this does not however mean that they actually *cause* trade (Baier&Bergstrand, 2009a)

Several authors also show the effect of international economic alliances on bilateral trade flows. Cashelli and Medda (2012) indicate the changes in trade flows after the introduction of China into the WTO. Another possible effect on trade that has yet to be investigated substantially is the effect of alliances and the admission of new members in existing 'agreements' on international trade (Ibid.).

Bensassi and Martinez-Zarzoso (2010) investigated the effect of acts of violence against vessels and its relation to international maritime trade. They found that only the most harmful form of violence, hijackings, lower the amount of trade between nations.

Recent research into gravity models has also shown that bilateral trade is determined by membership of an economic union (UNESCAP, 2008). This research on gravity models has also shown that bilateral trade is determined by relative trading costs. Interesting to note however is that in almost all the gravity models that have been developed, the bilateral tariff barriers are missing (Ibid.).

According to Bergmans (1998), gravity models should also include institutional characteristics of the countries, such as language barriers, common policy and common culture. A large body of authors has done research to the effect of these drivers on international trade volumes and trade costs. A small selection of these papers are Dutt and Traca (2010) on the effects of corruption on bilateral trade flows, Glick and Rose (2002) on the effect of a currency union on trade, Hutchinson (2001) on the effect of linguistic barriers, and Eichengreen and Irwin (1998) on the role of history in bilateral trade flows.

Fratianni and Marchionne (2009) give an extensive overview of the relations and parameters of bilateral trade that have been found in the scientific literature. First they state that distance and national borders are a big barrier to the expansion of cross-border trade. They also claim that seas and oceans can be considered as a 'natural border' where countries sharing the same ocean trade a half more than those that do not have a common water border (under the condition that the two countries do not share a common land border). This implies that the effective, or economic, distance between two countries is not so much determined by the geographical distance alone, but also by the characteristics of the barriers between countries. It can thus also be expected that countries that share common oceans or seas are better connected to each other, and are therefore also more likely to have a direct connection in the liner shipping network.

Furthermore, they state that while heterogeneity in culture and institutions pose limits tofurther integration, the forces of geography are the most potent. Continents act as ‘natural’integrators and oceans as common water border. Other drivers they have found in the literature are language, religion and colonial relations whilst noting that they all represent a similar cultural affinity. (Ibid.). Similar international situations and common institutions tend to increase bilateral trade as people and regions that share these similarities then to group together in both business and personal life. This is also stated in De Groot et al. (2004). Here the authors say that a low quality of governance increases the transaction costs that are incurred in exchange. They also state the the quality of formal rules affects the informal norms and procedures of doing business resulting in the possibility that countries with similar levels of institutional quality may be familiar with each other's business practices and thus further reducing the transaction costs leading to increased trade.

De Groot et al (2004) further conclude that both governance and institutional quality have a significant, positive and substantial impact on bilateral trade flows. Therefore, countries that have a low quality of institutions might trade less with countries with high quality of institutions but will trade more with other countries that also have low level quality of institutions. The same goes for the quality of governance.

Many of the papers described above are either focused on a specific region in the world. Also, they might have become dated as over time society evolves and with it, international relations and (both country level and international) institutions resulting in shifting trade pattern over time (Eichengreen and Irwin, 1998). This leads to the need of almost a consistent monitoring the relation between the above described determinants and trade over time.

Chapter 3 Methodology

In this chapter, the methodology of this thesis is discussed. The used models are explained together with several other methodologies used in the course of this thesis and the reasons for choosing them are justified. Next, the used data is briefly described and its sources are given. In the third part a small description of the reliability and validity of this research is given and discussed and this chapter ends with the main limitations of the research.

3.1 Components of the LSBCI

One of the most important parameters for the development of an index is its weighing method. According to a content analysis (Karamperidis *et al.*, 2013) most of the indices used in the maritime transport sector use weights which are extracted from expert opinion Within that category, the most frequently mentioned techniques for weights generation are the following: the Budget Allocation (BA) method, the Public Opinion approach, the Analytic Hierarchy Process and Conjoint Analysis approach (Ibid.). In this thesis however, the focus does not lie on generating the index itself, but on finding potential determinants for it and how they could be related to a LSBCI.

One way of generating potential constructs of a bilateral liner shipping connectivity index is by looking at the variation in the trade flows in country pairs and how this variation can be explained. It is then assumed that those bilateral determinants who have a significant positive relation to trade flows can act as a proxy of connectivity as it follows from the theoretical framework that trade is usually stimulated after the reduction of barriers that reduce the accessibility to the global trade networks. In other words, by reducing the economic distance the transaction costs for participating in the global liner trade network are reduced resulting in a higher accessibility of the network itself resulting in an increase in trade.

Therefore, the method used to generate the potential determinants for a bilateral liner shipping connectivity index the bilateral parameters that represent an indication of bilateral trade barriers and that thus increase the economic distance between any two countries are first identified using the gravity model of trade. Next, these parameters are used to generate a set of possible indices and are subsequently tested to infer whether their chosen components actually represent the measure of bilateral liner shipping connectivity.

3.2 The gravity model of trade

The gravity equation is typically used to explain cross-sectional variation in country pairs' trade flows in terms of the countries' incomes, bilateral distance, and dummy variables for common languages, for common land borders, and for the presence or absence of an FTA. It was first published in the early sixties by Jan Tinbergen although he had no theoretical foundation for his gravity equation. In the scientific literature usefulness of the gravity model applied to bilateral trade flows widely recognized resulting in some authors referring to this model as the workhorse of empirical trade studies (Ramos, 2006). It is based on the gravity equation in mechanics that can be used to show the forces acting between two bodies with a certain mass that are at a certain distance from each other. The general gravity equation takes the following form as shown in equation (1):

$$F_{ij} = G\left[\frac{M_i M_j}{D_{ij}}\right] \quad (1)$$

Here, F is the force acting between i and j, M is the mass for respectively i and j, G is a constant and D represents the distance between i and j. When applied to international or inter regional trade, the trade flows are considered to be the forces acting between i and j and the masses M are a representation of the size of the economy of i and j, as illustrated nominal GDP.

Traditionally, the relations are estimated by transforming the above non linear relation into a linear relation by means of a log linear model in order for it to be investigated by using a simple ordinary least square regression. This log linear equation takes the general form as given in equation (2):

$$ln(F_{ij}) = \beta_0 + \beta_1 \ln(M_i) + \beta_2 \ln(M_2) + \beta_3 \ln(D_{ij}) + \varepsilon_{ij} \quad (2)$$

Most authors include a similar set of variables when researching trade related relations using gravity models. They are often related to the potential size of markets and to the societal structure. Therefore, variables such as GDP, common borders and cultural parameters are often used in the literature.

In most studies, the researchers separate the import and export data when performing gravity equations. However, this is not always done leading to biased results as the equation is a modified expenditure function with a market-clearing condition (Fratianni and Marchionne, 2009) resulting in that it only serves to explains force directed trade flows, meaning trade flows from region i to region j.

Another issue with using the gravity equation became clear from research into economic integration agreements. In the literature on international trade, the gravity equation has dominated as being the main econometric approach toward estimating the partial or direct effects of economic integration agreements and other natural and policy-based bilateral trade costs on aggregate bilateral trade flows. Usually these economic integration agreements (EIA) refer to preferential trade agreements, free trade agreements, customs unions, common markets, and economic unions (Baier et al. 2013).

As a result of self selection of country pairs governments into agreements, the effects of bilateral relations often suffer from endogeneity bias. When accounted for this bias, Baier and Bergstrand (2007) showed that EIA's tend to have a larger and more precise effect on trade flows compared to earlier results using gravity equations. On a side note however, they also showed that the full impact of EIA on trade flows takes between ten and 15 years.

Following from the above, as is the same with all non exogenous variables, the multilateral effects need to be taken into account when using a gravity equation that involves bilateral relations. Baier and Bergstrand (2004) showed that when accounting econometrically for the FTA variable's endogeneity that the effect of FTA's on trade flow's is quintupled. This shows the potential relevance of accounting for the endogeniety of variables.

Most studies assumed an exogenous dummy variable to represent the multilateral FTA treatment effect (where the dummy would be 1 if both countries i and j in the country pair also have a FTA with a common country k). In reality however, FTA dummies are not exogenous random variables; countries select endogenously into FTAs, for reasons possibly correlated with the level of trade: the presence itself of an FTA is not exogenous. (Baier and Bergstrand, 2004).

Thus the fact that typical cross-section ordinary least squares (OLS) estimation using the gravity equation often yields small positive and often even negative average “treatment” effects on trade flows, is likely attributable to selection bias. By correcting the variables using econometric methods, this problem can be solved. (Baier&Bergstrand, 2008).

The paper that has been most influential in addressing variable bias in gravity models for trade is "Gravity with Gravitas" by Anderson and Van Wincoop (2004). In their paper they show that the traditional gravity equation (eq. 2) is misspecified and that coefficient estimates of the variables are very likely to be biased as a result of omitting the nonlinear multilateral price resistance terms of the countries i and j. This followed from the fact that bilateral trade flows are determined in a world with more than two regions resulting in an effect of the economic distance of the countries i and j from the rest of the world's regions on the bilateral trade volumes between i and j. They showed that the estimation of unbiased coefficients require the minimizing of the sum of squared residuals.

The other main outcome of the Anderson and Van Wincoop paper is that they emphasize the dependence of trade on the bilateral and multilateral resistance factors. They showed that the multilateral resistances are price indices that depend on all bilateral resistances, including those that do not directly affect or involve the exporting country.

However, as Baier and Bergstrand show (2009a), the Anderson and Van Wincoop method cannot fully ensure the unbiased estimation of the effect of bilateral relations on the trade flows for the reasons mentioned earlier (the self selection of country pairs) and as a result of the dependency of the variables on the typically used gravity model variables. Another issue related to the use of the method developed by Anderson and Van Wincoop is that it requires a nonlinear least squares (NLS) approach, which is computationally complex relative to a ordinary least squares approach (OLS) and is subject to measurement errors associated with internal distance measures (Baier&Bergstrand, 2009b).

An alternative to the NLS approach by Anderson and Van Wincoop is the use of country or region specific fixed effects (Baier&Bergstrand, 2006). Feenstra (2004) showed that unbiased estimates of the average effects of the bilateral relations on trade can be generated through the gravity equation by means of the region of pair fixed effects. Thus, the problem that arises from applying country fixed effects is that using this method, the multilateral price terms that are required to generate quantitative comparative static effects cannot be retrieved (Baier&Bergstrand, 2009b) resulting in the fact that many explanatory variables that are region or country specific such as internal policy measures in a country cannot be compared when using the fixed effects (Baier&Bergstrand, 2006).

In order to deal with the above described issues of these two often used methods for gravity equations, Baier and Bergstrand (2006) developed a simple OLS technique for estimating both the average and comparative statics from a gravity equation that include the theoretically motivated exogenous multilateral resistance terms. According to the two authors, the advantage of this method of the Anderson and van Wincoop method is that it is computationally simpler and the advantage over the fixed effects method is that it can provide quantitative estimates of the comparative statics using the estimated coefficients.

Therefore, the model used in this thesis for the gravity equation is one by Baier and Bergstrand(2006) where they developed a linear approximation to all influences of multilateral trade resistance. After this approximation, a regular OLS regression can be performed on the gravity equation. In contrast to the traditional log linear OLS gravity approach, the method proposed by Baier and Bergstrand is thus able to model and break down the influence of the multilateral resistances on the trade flows.

In the model of Baier and Bergstrand (Ibid.), a first order log linear Taylor expansion series is used to derive an OLS equation to the explanatory variables that include the multilateral resistance terms. The authors also showed that this model is applicable in the context of both the world as in the context of the region.

The model is based on the first order log linear Taylor series expansion to the system of price equations as developed by Anderson and Van Wincoop (2004) to generate a reduced form gravity equation that can be estimated using OLS. The method itself is henceforth referred to as the BV-OLS as it was called by Baier and Bergstrand. One issue that however needs to be taken into consideration when using the BV-OLS is that it is a linear approximations of a nonlinear system resulting in a reduction in precision of the estimates of the region specific or region pair specific comparative statics. However, Baier and Bergstrand (2006) show that their method results in almost identical results compared to the NLS approach by Anderson and Van Wincoop.

The Taylor expansion transformation of the independent variables to be put in the BV-OLS is performed as follows:

$$\left(X_{ijk}\right)P_i p_j = \frac{1}{N}\sum_{i=1}^{N} x_{irk} + \frac{1}{N}\sum_{j=1}^{N} x_{jrk} - \frac{1}{N}\sum_{i=1}^{N} x_{irk}\,\frac{1}{N}\sum_{j=1}^{N} x_{jrk} \qquad (3)$$

$$\left(X_{ik}X_{jk}\right)P_i p_j = \frac{1}{N}\sum_{i=1}^{N} x_{ik} + \frac{1}{N}\sum_{j=1}^{N} x_{jk} - \frac{1}{N}\sum_{i=1}^{N} x_{ik}\,\frac{1}{N}\sum_{j=1}^{N} x_{jk} \qquad (4)$$

Here, the X refers to the independent variables to be put in the gravity equation and r is an index of the country partners of respectively i and j. Equation (3) refers to variables with a bilateral variability, such as distance, and equation (4) refers to variables with country or sectoral variability, but which are common for all the trading partners (such as country specific policies). The subscript PiPj refers to it that the variables are transformed in order to deal with the multilateral price resistances.

After the independent variables have been transformed, a log linear OLS regression has to be performed that is in the following form:

$$ln\left(X_{ijk}\right) = \propto_0 + \propto_1 \ln(M) + \propto_2 \left[\ln\left(Dist_{ij}\right) - ln(Dist_{PiPj})\right]$$

$$+\propto_3 \left[\ln(ET_i)\ln(ET_i)\right]_{PiPj} + \varepsilon_{ijk} \qquad (5)$$

In Equation (5):

$$X_{ijk} = valueofexportsfromcountryitojforcommodityk$$

$$M = the\ product\ of\ the\ nominal\ GDP\ of\ countries\ i\ and\ j$$

$$Dist_{ij}\ isthegeographicalmaritimedistance$$

$$Dist_{PiPj} = thegeographicalmaritimedistancethathasbeentransformed$$

$$by\ means\ of\ the\ Taylor\ expansion$$

Furthermore, $ET_i\ and\ ET_j$ are the country or region specific variables for respectively the exporting and importing countries.For bilateral variables that have been transformed as with equation (3), they substituted in equation (5) in the same manner as the geographical maritime distance.

Although there exist clear economic differences both inter and intra countries leading to differences in the way the determinants of bilateral trade flows behave, for this study only country level data is analyzed assuming each country to be internally homogeneous. In order to deal with the inter country differences; country-heterogeneity is taken into account when analyzing the bilateral trade flows. Furthermore, sector homogeneity is assumed in the following analysis despite research suggesting differences in trade determinants between sectors (Ramos, 2006). The reasons for this decision are a lack of sectoral data of sufficient quality and because this analysis focuses on the behavior and effects of liner shipping in general, and not on its specific sub sectors.

One issue concerning using a gravity equation in aggregate bilateral flows is the existence of zeroes and zero trade flows as they will distort the results of the regression (Baier et al., 2013). Usually this problem is dealt with by excluding the bilateral relations that include zeroes or have a zero trade flow or by transforming the zero valued trade flows to a small positive number (Burger et al., 2009). Although some authors show that both of these measures cause a bias in the results (Baier et al., 2013), the majority of researchers still uses one of these methods. Santos Silva and Tenreyo (2006) showed that the zero trade flows can also be accounted for by estimating a multiplicative version of the gravity equation using a Poisson Quasi Maximum Likelihood Estimator. However, as the bias in the results that is created by transforming the zero trade flows in a arbitrary small number (into a 0.5) is very limited, and as several authors have severe doubts about the Santos Silva and Tenreyo method (see for instance Anderson (2013) and Ramos (2006) it has been chosen to transform the zero trade flows into a 0.5. Furthermore, the reason as to why it was chosen to transform the zeroes instead of deleting them is that deleting them will likely lead to a larger bias in the results (De Groot et al., 2009).

Fratianni and Marchionne (2009) discuss in their paper also the mention of the three common estimation errors that are frequently made when performing empirical gravity equations that usually arise from the issue of the multilateral trade factors. The first one is that researchers sometimes use the real GDP instead of the nominal GDP (where the multilateral factors are not well identified and the model errors fail to be orthogonal to the regressors, resulting that the OLS estimator is asymptotically downwards biased), the second one is employing two-way bilateral trade (as the gravity equation is a modified expenditure function with a market-clearing condition, the theory explains only one-way bilateral trade and not two-way trade) and finally the overall omittance of the multilateral resistances as the largest mistake commonly made.

In this research, a particular care has been taken to prevent these errors from happening. In the analysis, the nominal GDP is used and the gravity equation is used for the force directed flows (meaning trade flows going from a country i to a country j). The final and main common mistake is prevented by using the BV-OLS model by Baier and Bergstrand where the variables are transformed in order to deal with the multilateral resistances.

3.3 The network structure

Other concepts that need to be addressed in this chapter is that of network centrality and network structure in order to help sufficiently describe the structure of the global liner shipping network. The global liner shipping network can be considered to a so called complex network. This is a network or graph with certain non trivial topological features. These features include high degree distributions (a high number of connections or edges between the nodes), high clustering and hierarchical structures that all together determine the structure of the network. By analyzing the hierarchical structure of a network one can determine the centrality of each node within said network. In the next chapter, these concepts will be explored further and analyzed for the liner shipping network in 2013.

Network theory has been used and applied for several decades in the social sciences starting with Freeman (1978) with his influential paper on the conceptual clarification on centrality in (social) networks. Here, Freeman gives a definition of three types of centrality (degree centrality, betweenness centrality and concentration centrality) that can define the overall centrality of each node in a network. For maritime networks however, network theory is hardly used. There have only been a few papers published starting from 2010 where some measures of network theory are applied to the maritime sector and most of these have been written by Ducruet (see for instance Ducruet and Notteboom, 2010; Ducruet and Lugo, 2011; Ducruet et al., 2012 and Wei et al., 2012). Almost all these papers also only focus on port to port level network structure, and not on the bilateral structure of the global liner shipping network.

According to Ducruet (2013), most research on networks in (maritime) transportation focuses on simple graphs where only one type of link occurs among nodes. While this is largely due to the absence of a proper mathematical definition of multigraphs (or multiplex graphs), the possibility and implications of having two or more types of links were approached by sociologists since the 1960s. These multiedge relations are in practice often analyzed as simple graphs where links have multiple attributes following the application of conventional network analytical methods.

In a complex network all nodes can be considered to have certain forces working on them, some with a positive direction, and some with a negative direction (so some forces attract other nodes and some repel them). The strength of these forces also depend on the distance to the node: some will increase over the distance whilst others will decrease. Thus, when all these nodes are put together, they will end up in an equilibrium position where the sum of all forces on each node equals to zero. If edges (the connections between nodes) are predefined and if these edges have certain values attributed to them, these edges can be considered to be the forces working on the nodes.

When this is applied to the global liner shipping network, the parameters of the edges (such as the maximum ship size on that particular route) could be considered to be the bilateral forces whilst the specifications of each country (such as its quality of infrastructure) act as the forces generated by the nodes themselves. Using this concept, one can generate graphical representations where the location of each node is determined by all the forces working in the whole network. Also, when the liner shipping network is regarded in this light of forces working between nodes, the concept of connectivity can thus be considered as the sum of the absolute of all the forces acting between the two nodes (the absolute has to be taken because otherwise the sum will always be zero as the overall network tends towards an equilibrium where all sums of all forces is zero. If this sum is not zero, the structure of the network will change until it is again in equilibrium).

The concept of the multilateral resistances as described in the previous chapter can also easily be illustrated using the idea of forces working on the nodes. If one considers an network with only two nodes i and j, with strong attracting forces between them in an equilibrium state. If a third node k is added to the network with weak forces the location (and thus also the length of the edge between them) of both node i and j will shift slightly, leading to a different bilateral relation, even if there is no direct strong link between i and k or between j and k.

When the above described concept is applied to the gravity model for trade in the maritime liner network, the bilateral variables between two countries can be considered to be the forces in the network that act only between the respective countries. Furthermore, the country or region specific variables can be considered to be the forces acting only on their respective country. Therefore, as every single force acting in the overall network will affect the overall network structure it has to be concluded that also these country or region specific variables need to be applied in the gravity model of trade. However, both the country or region specific variables and the bilateral specific variables still need to be transformed using the first order Taylor expansion in order to include the multilateral resistances (or in the terms of the concepts given in this section, to incorporate the effects of the forces in the network not directly working on the countries i and j on the relation between i and j). Therefore, equation (3) has to be used for the transformation of all bilateral relations, and equation (4) for the transformation of all country or region specific variables.

The above applies to all bilateral representations or results in a network, and thus also in the case of a bilateral index that must represent a measure of connectivity between two countries in the global liner shipping network. Thus, for the LSBCI the components do not only need to exist out of the bilateral relations, but can also consist out of country or region specific variables as they will also affect the bilateral relation for any country pair.

3.4 Data description

One of the main issues involving the estimation of the value and volume of containerization freight flows between countries is the availability of complete databases. As none of the previous studies have looked into this ratio, Sou et al. (2013) chose to use a value-weight container conversion factor based on the data available to them in order to estimate maritime trade flows in ASEAN countries.

The authorsused a method to transform the value of the maritime trade into the number of containers transported between two countries. Using data from Containerization International, Eurostat and the USA Trade online they developed a mean weight container conversion factors and value-weight conversion factors for the trade flows between four ASEAN countries (Sou et al. 2013).

The problem when using this method however when looking at the global trade flows and not just at a specific region is the difficulty in obtaining the correct conversion factors for each region. Also, in order to determine the conversion factors for each country pair, a lot of time is required and considering the limited availability of time for this research combined with the difficulty of obtaining the relevant data to determine these factors is has been chosen to use for the purpose of this research bilateral trade data has been obtained from the trade database COMTRADE. The analyses performed in the next chapters have been done by valuating the trade values in billion US Dollar.

As this research focuses on liner shipping, only the manufactured goods are used in this analysis. The selected data came from the HS2012 nomenclature using 2 digit classification data. The reason this nomenclature and digit classification was used as the dependent variables is that as this research only focuses on liner shipping, the resulting relations are very likely to be biased if data on trade is used that is not containerized. By using this classification, the distinction between each classification is clear enough to select without having the trade data polluted by non-containerized data.

The countries that have been used for this analysis are only coastal countries. The reason for this is that when non-coastal countries are taken into the analysis, the results will be distorted as the research and all the models focus on maritime relations and by including them, land based relations also have to be included in order to prevent the results to be biased. Out of the 186 possible coastal countries, a selection has been made for 159 countries to be included in this research. The reason as to why several countries were dropped is either because of a unavailability of reliable data for these countries or because the countries are so small that their effects on the overall network are neglect able (or because of a combination of the two). In appendix A an overview is given of the selected countries combined with their ISO3 classification that used further in other tables to indicate the countries.

However, not all countries trade with each other in all commodities, and not all countries trade at all with each other. This is also illustrated in the dataset on the bilateral trade that is used for this research. Only 5946 out of 12562 possible country pairs trade with each other in manufactured goods. This means that in only 47.33 percent of all possible country to country trade routes there is actually any trade in manufactured goods occurring. Similar percentages are also found in other research on trade indicating that the used data set on bilateral trade values is reliable.

Another source of data is the databases by UNCTAD used in their research for liner shipping connectivity and for other maritime related research. This database includes variables such as the maritime distance between countries, several institutional dummies such as common language, currency, border but also whether a country is an island.

Other datasets have been obtained from the World Bank performance indicator database (WorldBank, 2013). From this database, indicators such as logistic performance and port infrastructure have been obtained. Another database that was used was the doing business database from the World Bank. From this database, indicators such as the number of procedures required when importing or exporting have been acquired.

The data on the structure of the network itself originates from Lloyds List Intelligence (LLI, 2013) that is annually provided to UNCTAD. This dataset gives all the existing direct country pair connections and includes the number of ships sailing in this route, the TEU capacity of all these vessels together per direct route, the TEU capacity of the largest vessel per direct route, the number of operators per route (both those who operate their own vessels and those who do not) and the number of services per direct route (both unique and non-unique). This dataset from Lloyds List Intelligence that is used for this paper is based on the global liner shipping network in May 2013.

3.4 Limitations

As the research for this thesis is quantitative, the quality of the results depends on both the quality of the analysis itself and on the quality of the original data sets. Simply put, if the data is not of sufficient quality, the results are also not of sufficient quality.

This is especially an issue with regards to the data on the bilateral trade value. As it is very difficult to accurately determine the actual and real trade volumes and values it is also difficult to make any judgment on the quality of this data. It is interesting to see however that almost every trade database has different values for specific routes and countries themselves report different trade values and volumes (there are also cases known where a country i reports a certain volume of exports to country j where country j reports a completely different import volume from country i).

Chapter 4 Liner shipping

In this chapter a description of the global liner shipping network is given and discussed. This is done by first giving the main characteristics of liner shipping and how they affect the structure of the network. Next, the structure of the network in 2013 itself is discussed and subsequently several metrics that can be used to describe the position in a network are presented and applied to the network.

4.1 Liner shipping characteristics

The global relevance of the maritime container transport sector is clear as a result of the following numbers as Karamperidis et al. (2013) shows: Although containerships have a deadweight share of 12.9 percent of the total world fleet it represents 52 percent of the global seaborne trade in US dollar terms while if the seaborne trade (in monetary terms) between containerships and dry bulk carriers is compared, each tonne of a container ship deals with 27 times more trade than each tonne of a dry bulk carrier. Furthermore, from UNCTAD (2012) it follows that the maritime sector carries around 80 percent of international trade (in volume) and 70 percent (in value).

As already mentioned in the introduction, the main characteristic of Liner shipping is that companies provide regular scheduled services and fixed routed shipping services. Their planning decisions are normally carried out every three to four months in order to catch up with annual and seasonal fluctuations in demand. These decisions are usually starting with forecasts of the expected shipment demands after which the decision maker designs the container flows in such a way that all routes can be utilized in the most optimal manner. Only after this are the container ships assigned to the routes in order to meet the capacity requirements (Liu et al., 2011).

Yap and Notteboom (2011) show that the outcome of the container port choice competition was mainly influenced by transshipment and hubbing decisions of shipping lines in Asia. They attribute this to the central location of several of its ports with respect to the containerized cargo flows between East Asia, South East Asia, Australia, the Middle East and Europe. They also state that the dominant position of Singapore could be threatened with the emergence of two nearby ports who are trying to facilitate transshipments in order to attract the big shipping lines and alliances.

Next to this, the majority of vessels that are deployed to serve the US East Coast market are routed through the Panama Canal, in which its physical restriction puts a limit of vessel size. However, the expansion of the Panama Canal will enable ships up to 12,500 TEUs to transit through it. This will provide opportunities for shipping lines to use larger ships taking advantages of economies of scale to lower unit cost. (Guan & Yahalom, 2013). It implies that after the opening of the new Panama Canal, the whole network and structure is likely to undergo substantial transformations. In the future, these changes could be even more intense if the now (again) proposed Nicaragua Canal is actually going to be built.

Another recent change in the global liner shipping network structure follows from Maersk Line, the world's biggest container shipping company, no longer using the Panama Canal but instead the Suez Canal to transport goods from Asia to the US east coast in order to make better use of economies of scales and its resulting lower operating costs per TEU compared to using the Panama route (Bloomberg, 2013). However, for other routes Maersk continues to use the Panama Canal (Ibid.).

In the last decade, a shift is occurring in the location of the main global demand centers from the Northern Hemisphere to several regions placed in the Southern Hemisphere direction. For these regions, it must be expected that vessel traffic will increase both in intensity as that it will lead to the usage of different classes of container ships on those routes (Seoane et al., 2013). This requires substantial investments for ports in these regions in order to deal and cope with this likely future further increase in vessel size and draft. The authors also state that the infrastructure in the South Atlantic ports is also currently not sufficiently developed to cope efficiently with the South Atlantic interconnections themselves (Ibid.), resulting in barriers to trade in that region.

However, in general ports and governments have responded to the overall market growth and dynamics in container transport with large investments in their infrastructure to attract these cargo flows. According to Zondag et al. (2010) these investments have focused on all components of the maritime transport chain. One example of type of investments is in improving the accessibility of ports for the increasing vessel sizes.

Thus, it can be concluded that the structure of the liner shipping network may change in the future as a result of shifts in global economic production and consumption centers and as a result of changes in the route possibilities. Furthermore, the changing economic environment also creates a constant dynamic in the structure on the liner shipping network. A question that arises from this is whether global trade patterns will change as a result of the changing global network following shifts in the global economic environment or whether the changes in the shipping network are more caused by changes in global trade. As stated before, the tactical decision planning of liner shipping companies is normally carried out every 3 to 4 months in order to deal with fluctuations in demand. This implies that the network depends more on the global trade patterns than viva versa. However, for the purpose of this research is this question not relevant as in order to determine possible proxies of bilateral liner shipping connectivity only a relation between trade and the proxy is required: causality is not required. The in this paragraph posed question could however be the interesting topic of future research.

4.2 Liner shipping network structure 2013

By using the data from Lloyds List Intelligence, the structure of the liner shipping network in 2013 can be analyzed in this subchapter. As stated in the chapter on the methodology, this dataset describes all the direct connections for coastal countries. As however, not all coastal countries are included in the analysis for this paper; the description of the liner shipping structure only takes the earlier mentioned 158 countries into consideration.

Between these 158 countries, there are 2169 direct connections out of a possible of 12562 country pairs. This implies that a substantial part of all country pairs requires at least one transshipment in order to transport goods between the two countries. This is clear from table 4.1 Here, an overview is given of the minimum amount of transshipments required to connect each country pair. In appendix B this overview is given per country. The results are based only on the direct connections between countries: If a direct connection exist between a country pair, zero transshipments are required. If this is not the case, it has been calculated whether a connection is possible with only one transshipment and subsequently with two or with three transshipments. All connections between each country pair are possible with less than four transshipments.

Table 4.1. Share of minimum number of transshipments required for all country pairs (Source: Author)

	0 transhipments	**1 transhipments**	**2 transhipments**	**3 transhipments**
Percentage of all possible connections	17.26	63.61	19.03	0.10

As is shown in table 4.1 about 17 percent of all routes is directly connected, almost 64 percent requires at least one transshipment, 19 percent of all routes require at least 2 transshipments and 0.1 percent of all possible routes requires at least 3 transshipments.

From appendix B it follows that whilst a few countries have many direct connections, others only have a handful of direct connections, thus requiring at least one, two or even three transshipments in order to reach other countries. Two clear examples of both extremes are the USA (with 98 direct connections, 59 times one transshipment required and 1 time two transshipments required) and Greenland(with only 1 direct connection, 28 possible connections with one transshipment, 116 possible connections with two transshipments and 13 connections where at least three transshipments are required).

However, it has to be taken into consideration that these numbers are only based on the existing direct connections and that these connections have not been given any weight. As a result of this, the direct connection between China and Hong Kong has the same 'value' as the direct connections between the USA and all the small Caribbean island states.

In table 4.2, an overview is given of the maximum and the average of each of the seven components in the dataset provided by Lloyds List on the existing direct country pair connections. From this table it follows that the maximum values for all the routes (except for the maximum ship size, as this ship is operational along multiple country pair routes) are all in Asia, with four out of the seven being on the route between Hong Kong and China implying a very strong link and connection between these two countries. It has to be noted however that this data set originates from May 2013, so at a time when the Emma Maersk was still the largest container vessel explaining maximum capacity of the largest vessels is 16020 TEU.

Table 4.2, Overview data on direct routes (source: author based on Lloyds List Intelligence)

	Total number of ships sailing on route	**Total transportation capacity on route in TEU**	**Largest capacity vessel on route in TEU**	**Number of Operators active on the route**	**Nu Operators providing services**	**Number of services on route**	**Number of non unique Services on route**
Average	27.585	134780.4	4446.774	4.64667	6.8671	4.8381	6.5953
Max	997	5730013	16020	52	67	173	292
Route with the max value	China Hong-Kong	China Hong-Kong	Multiple routes	Malaysia Singapore	China South Korea	China Hong Kong	China Hong Kong

Table 2.

4.3 Graphical representation of the global liner shipping network

As a result of the existing direct connections between the different nodes (countries) in the network a certain structure is created where some nodes are more centrally located in the network than others, meaning that some countries can more easily be reached from any random position in the overall network than others. This can also be shown in a graphical representation of the network.

By using the Graph Embedder algorithm developed by Frick et al. (1998), henceforth referred to as the GEM Frick algorithm, any graph $G = (V, E)$ where V are the vertices or nodes and E are the edges or connections between them, can be drawn using straight lines in which each edge is mapped into a straight line segment in the plane (Ibid.). As a result of this, the nodes can be mapped $\varphi: V \rightarrow R^2$. The GEM algorithm is furthermore developed in such a way that all the nodes and edges are distributed evenly in the area, that the connected vertices are located closely to each other and that the number of edges that cross each other on the plane is minimalized. As a result of this, the nodes that are more central in the graph have a higher centrality and have more direct connections within the network and the closer a node is to the periphery, the lower its centrality.

Using a GEM FRICK algorithm together with Tulip Software a graphical representation of the global liner shipping network structure has been created as shown in figure 4.1. As a result of this algorithm, the nodes that are more central in the graph have a higher centrality within the network and the closer a node is to the periphery, the lower its centrality. Next to this, all the direct connections have been given a color in accordance to the maximum TEU capacity on that particular route: routes that have the highest TEU transport capacity are colored red, and the routes with the lowest capacity blue.

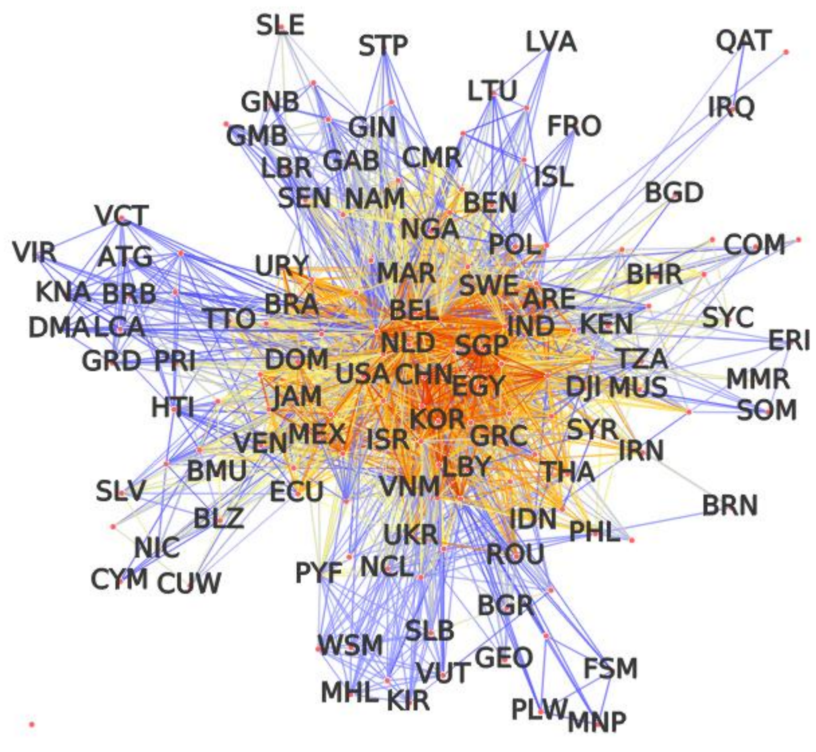


Figure 4.1. Representation of the global liner shipping network 2013 using a GEM FRICK algorithm. (Source: Author)

Figure 4.1 is presented in more detail in appendix C, where several close-ups are presented of each part of the graph.

What is interesting to see in figure 4.1 is that the majority of the edges in the graph that are colored red are located in or near its center implying that those nodes that have a higher network centrality also have increased services between them. Therefore, it can be concluded that countries that are more centrally located in the global liner shipping network are also countries that tend to services operating from and towards them.

Furthermore, this graph shows that there is one main cluster of routes (the center group) and several other smaller separate clusters (the points in the star shaped graph).

The different 'star points' in the global network graph are clockwise starting at the top: 1) the Baltic and Scandinavian states, 2) the Arabian peninsula and countries surrounding the Gulf Aden, 3) Oceania and surrounding island states, 4) Central America and several other Caribbean countries, 5) the northern countries of South America combined with the other Caribbean Islands states and finally 6) the countries on the 'Atlantic' side of Africa. The center group consists out of the expected countries such as Hong Kong, China, US, Netherlands, France, Germany and Singapore.

Another interesting point that is illustrated in this graph is that some countries lie in “between” two of the smaller cluster groups. An example are Jamaica and Haiti, which lie between cluster groups four and five, implying that it acts as a bridge between the two different groups of Caribbean Countries. Also interesting is that Uruguay and Brazil act as a bridge between the Caribbean and the Atlantic countries of Africa.

In the close ups of figure 4.1 as shown in appendix C it can be seen that for the central countries in the graph the TEU capacity per route is between the Asian countries (China, Hong Kong, Korea, Singapore, Malaya and Taiwan). These capacities are slightly less in the connections between these Asian and several European and Mediterranean countries (the Netherlands, Germany, France, Belgium, Malta, Italy, Spain and Egypt). These results are logical when it is considered that the most containerized trade occurs intra-Asia and second largest containerized trade flows are between Asia and North West Europe (UNCTAD, 2012).

Furthermore, it is worth noticing from the close ups that although Great Britain and the USA are also in the center of the graph (despite being on the edge of the center) the majority of their connections is to countries that are in the periphery of the overall network graph and that the majority of their connections have a low TEU transport capacity indicating that although they have many direct connections to other countries, the transport flows over the majority of these routes is limited compared to the cargo flowing from and towards other 'central countries'.

On a final note with regards to the graphical representation of the liner shipping network is the location of Russia in the center of the graph. As a result of Russia's geographical dimensions and location, it has a relative close maritime distance to North (West) Europe, the Mediterranean states and to Asia, resulting in its many direct connections, even if Russia itself is not located at or directly near any of the main containerized trade flows.

4.4 Ways of measuring the position in the network

Various methods of networks have been developed with the goal to compare the relative position of nodes in a network and to highlight groups of nodes within it. Both Ducruet and Notteboom (2010) and Ducruet and Lugo (2011) give an overview of a few of these measures that are often used to define the network structure and centrality for the maritime sector. However, there are many more of these metrics that can be used with this purpose. Six of these parameters that together or individually could act as a relevant representation of the global liner shipping structure are given below and discussed below.

It has been chosen to determine the metrics below for each of the nodes, and not for the existing edges for several reasons. The first is that only a limited number of the country pairs is directly connected and that all these metrics depend on the structure of the network that follows from the direct connections between some of the nodes. Therefore, if the metrics are applied to the connections between the countries, it is not possible to define all country pairs in terms of these metrics. Thus, by calculating the network characteristics per node, the relation between each country pair can actually be determined also resulting in the possibility to compare countries characteristics over time within the network.

4.4.1 Degree Centrality

The degree centrality can be described as the number of direct connections for a node in a network. According to Freeman (1977) the degree of a node can be seen as an index of its potential interactive activity with other nodes in the network: The higher its degree, the higher the number of nodes it can interact directly with without the need of making use of other nodes. The degree centrality is determined by (Ibid.):

$$C_D(P_i) = \sum_{i=1}^{n} \propto (P_i, P_j) \qquad (6)$$

Where

$\propto \left(P_i, P_j\right)$ =1 if and only if P_i and P_j are directly connected and 0 otherwise.

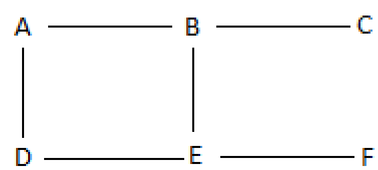


Figure 4.3 Example betweenness centrality (source: author)

In figure XXX, a small network with five nodes is graphically represented to illustrate the idea of the degree centrality. Here, node A is directly connected to two other nodes and has thus a degree centrality of two. Following the same principle, B and E a degree centrality of three, node D a degree centrality of two and nodes C and F each a degree centrality of one.

The degree centrality could potentially be a relevant metric for determining the bilateral liner shipping connectivity as it captures to how many other countries a country is connected. It seems logical that if a country is connected to many others, it has apparently good connections to the rest of the network.

4.4.2 Betweenness Centrality

According to Freeman (Ibid.), another view on network centrality is based upon the idea of the frequency with which a node falls between pairs of other nodes on the shortest path connecting them indicated with the term betweenness centrality. The author mentions that when a node is strategically located on the communication path that links a pair of other nodes, it can be considered to be a 'central' node as it is in such a position that it can network by distorting or even preventing the flow over the edges between the origin and destination nodes. Furthermore, it is mentioned that nodes with a high betweenness centrality have a strong potential to act as the coordinators of the network as they have the option to control all the flows in the network (whether it be communication or transportation of cargo or any other type of flow).

Freeman (1977) also gives how to determine the betweenness centrality. To determine the overall betweenness centrality of a node P_k, the partial betweenness values for all unordered pair of points where $i \neq j \neq k$ and where n is the number of nodes in the graph:

$$C_B(P_k) = \sum_i^n \sum_j^n b_{ij}\ (P_k) \tag{7}$$

Where:

$$b_{ij}\,(P_k) = \frac{g_{ij}\,(P_k)}{g_{ij}} \tag{8}$$

Here, g_{ij} is the number of geodesics linking P_i and P_j and $g_{ij}\,(P_k)$ is the number of geodesics linking P_i and P_j that contain P_k.

The idea of the betweenness centrality can also be illustrated below in the figure using node B. As in this small network there are a total of four nodes, there are a total of six different node pairs possible. Next for each of these node pairs, the total number of possible routes between them has to be identified and finally in which of these routes the path goes through node B. The results are in table 4.3

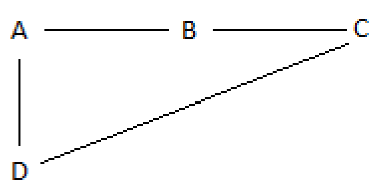


Figure 4.3 example betweenness centrality (source: author)

Table 4.3, example betweenness centrality (source: author)

Node Pair	***Number of shortest paths possible***	***Number of shortest paths that pass through node B***
AB	1	*0*
AC	*2*	*1*
AD	*1*	*0*
BC	*1*	*0*
BD	*2*	*0*
CD	*1*	*0*

Thus, the between centrality of node B in this network is:

$$C_{D_B} = 0 + \frac{1}{2} + 0 + 0 + 0 + 0 = 1/2$$

The betweenness centrality can thus be considered to be a measure of how often a country is used as a location for transshipments implying that if this metric is high, it is located on a strategically strong position within the network.

4.4.3 Clustering Coefficient

Chiricota et al. (2002) state that one of the basic assumptions of research in techniques to help capture the structure of existing systems (and networks) is "that well designed …. systems are organized into cohesive subsystems that are loosely interconnected". This has resulted into several techniques to divide networks into its possible sub components that can considered to be separate clusters in the overall network. One of the most popular methods to find these clusters is by splitting the overall network into several subsets such that the number of edges connecting the nodes between each distinct subset is minimalized (Ibid.).

This concept of a cluster can also be applied to the nodes themselves. The higher the interconnectivity of a nodes neighbors, the more it is part of a (local) subset or cluster. Therefore, a coefficient can be created that indicates to what extend a node is part of a cluster by comparing the interconnectivity of its neighbors to the maximum number of edges that can connect the nodes in the subset N (Chiricota et al. (2002):

$$C_k = \frac{e(N_k)}{\begin{bmatrix} d \\ 2 \end{bmatrix}} \qquad (9)$$

Where N_k are the of neighbors of node k and $e(N_k)$ the number of edges that connects these neighbors and where $\begin{bmatrix} d \\ 2 \end{bmatrix}$ is the binomial coefficient $\frac{d(d-1)}{2}$ that corresponds to the maximum number of edges that can connect the nodes in N_k. The principle of this coefficient is shown in figure 4.4:

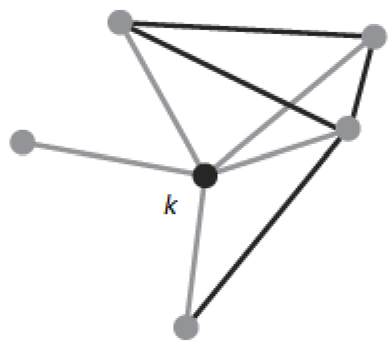


Figure 4.4. Example Cluster Coefficient (source: Author based on Chiricota et al. (2012)

Here, node k has five directly connected neighbors that have four edges connecting them with each other. Thus:

$$C_k = \frac{4}{\binom{5}{2}} = 0.4 \quad (10)$$

If the cluster coefficient of a country i is high, it means that its directly connected neighbors have goods connections amongst themselves, reducing the need of using for them to use country i on their routes when importing and exporting. It is thus an indication of the relative position of a country within the network.

4.4.4 Strength

It is also possible to develop a metric that indicates the relative strength of a node in a network compared to its neighbors. This strength can be also be interpreted as its relative position in a network or sub network: It can show whether it a node is part of a cluster (and whether it is in its center or in its periphery) or whether it acts as an isthmus between clusters (or bridge between two clusters). Chiricota et al. (2002) developed a method to determine this relative strength of an edge in a graph, however, with several mathematical adaptations, their concept of edge strength and their method to calculate it can be used to determine the strength of nodes in a network.

The algorithm by Chiricota et al. (Ibid) determines to what extend the nodes that are directly connected to a node i are connected to other nodes not directly connected to node i. When a the directly connected neighbors of i have more (inter) connections to other nodes not connected to i, it is an indication that these neighbors are part of a cluster where node i is not part of.

The more a node is connected to these clusters, the lower its strength indicating that it acts as an isthmus between clusters whilst a high strength indicates that a node is more at the center of a cluster

4.4.5 Strahler number

Another metric that can help creating a hierarchy in a complex network is by means of the so called Strahler number (Strahler, 1952). The Strahler (sometimes called the Horton-Strahler) number can be considered as an attempt to give the complexity of a network a quantitative qualification: the higher its value, the higher the complexity of a nodes sub trees or networks (REF). It has mainly been used to morphological structures, molecular biology and also in computer graphics and sciences. Furthermore, it has also been applied in combinatorial research in mathematics (Ibid.). Although initially defined for binary trees, it has since its initial use also been applied on general trees, rooted networks and complex networks (see for instance Fedou (1999) and Auber (2003)).

The Strahler number is a numerical measure of the branching complexity in a network. The higher the number, the more edges come together in that particular node. If the number is 1, it means the node is the endpoint with only one direct connection to another node. If the Strahler number is 2, it branches out towards two other nodes with a Strahler number of 1. This principle is the same for all other increasing Strahler numbers. For any node N, the Strahler number can be determined as shown in equation (11) where $N_i, i = l, \ldots, \rho_N$ are the nodes flowing from the node N (Herman et al., 1998):

$$S_N = \begin{cases} 0 & \rho_N = 0 \\ S_{N_l} + \rho_N - 1 & \forall i,j * S_{N_i} = S_{N_j} \\ \max\{S_{N_i}\} + \rho_N - 2 & \text{otherwise} \end{cases} \tag{11}$$

The principle of the Strahler number is illustrated in figure 4.5. Here, nodes C, D, G and H all form the 'end' of the network tree, giving them the Strahler number of 1. Subsequently, F also has a Strahler number of 1 as it does not branch out into multiple edges. However, both E and F branch out in two sub parts each with a Strahler number of 1, giving B and F a Strahler number of 2. Node branches out in two edges each with a Strahler number of 2, giving it a Strahler number of 3.

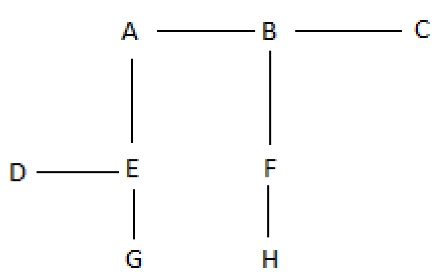


Figure 4.5, example Strahler Number (source: Author)

4.4.6 K-Core

One other way to describe the topology of complex (real world) networks is by using the notion of the K-core. The K-core has been applied to a number of real world networks and has proven to be an important tool for the visualization and interpretation of cooperative processes in complex networks (Dorogovtsev et al., 2006). The K-core is also known as degeneracy: it is the maximal connected sub graph of a network where all nodes have at least a degree of k. Thus, the K-core can be considered to be a representation of the amount of integration of a node or sub network within the overall network.

The K-core is obtained by removing all nodes in a network with a degree lower than k. After this removal, some nodes now also have a degree lower than k after which they also have to be removed. This process has to be continued until no other nodes can be removed in this manner (so until no other nodes have a degree lower than k) (Ibid).For each node, the K-core determined is thus the highest possible degree k in order for the node to be still part of the newly formed sub graph. This is illustrated in figure 4.6

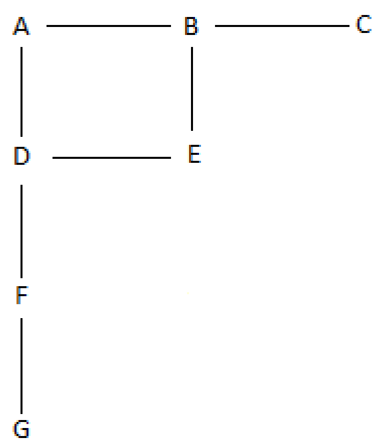


Figure 4.6, Example K-Core (source: Author)

If in figure 4.6, k is one, no nodes have to be removed as they all have a degree of at least one. If k is two, nodes C and G have to be removed, as their degree equals to one. After this removal, the degree of F decreases from 2 as in the figure to 1 as after the removal of G, it is only connected to one other node. Thus, the K-core of nodes C, F and G is 1. For a k of three, all the other remaining four nodes also have to be removed from the subgraph, as each of them now only has a degree of 2. Therefore, the K-cores of the nodes A, B, D and E is two.

The K-core could be a relevant component for this paper as it shows the embedidness of a node within the network. The higher the K-core, the less likely the node will cease being part of the network.

4.5 The network node metrics

The above six measures of network structure have all been calculated for the global liner shipping network of 2013. These results can be found in appendix D and will be further analyzed as being part of the pillar 'network structure' in chapter 6.

It is furthermore worth to mention that if a country has a low clustering degree combined with a high degree centrality it means that it is connected to many other countries that do not have many connections amongst each other. This implies that said country acts as a hub within the network. Two examples of such countries are Italy and Belgium.

Chapter 5 Connectivity

In this chapter, a brief overview is given of of the concept of connectivity. Following this, both LSCI and the LSBCI are discussed

5.1 The concept of connectivity

In the existing body of literature several definitions of transport connectivity are given. The following overview of these definitions is taken from (Tobar et al., 2013):

"Several definitions of transport connectivity are found in the literature. According to Marquez et al. (2006) transport connectivity is an attribute of networks that refers to the quality and costs to move freight between two points in space, meanwhile Wilmsmeier and Hoffmann (2008) describe transport connectivity as the access to regular and frequent transport services and the level of competition in the service supply; additionally, Cullinane and Wang (2009) talk about the connectivity of certain container port as the potential for the movement of containerized cargoes to and from other container ports via the network."

In recent years the concept of maritime connectivity has also gained a lot of popularity. This can be seen in the increasing amount of studies that analyze the relation between maritime connectivity and different maritime aspects such as transport costs, port competitiveness, logistics connectivity and maritime security. (Tobar et al., 2013). Only a very small amount of studies has related connectivity to the liner shipping network using network theory, and they all focused on a limited number of network metrics and on the port level network.

Hoffmann et al., (2013) mention that different measures of country level liner shipping connectivity have been used as explanatory variables in research on trade competitiveness and on transport costs. These recent studies have found strong correlations between liner shipping connectivity and trade costs. The authors also mention that other research has found that liner shipping connectivity has a more significant impact on the formation of international maritime transport costs than the geographical distance.

Furthermore, Hoffmann et al. (2013) also mention that a study performed by ESCAP has shown that a exporting countries LSCI has a higher correlation with the trade costs than the importing countries LSCI. This implies that having access to the global network is more relevant for the exporting country than for the importing country.

5.2 Liner shipping connectivity index

The UNCTAD LSCI is generated yearly for the same 159 countries that are subject to the research for this thesis based on the same database from Lloyds List Intelligence as used to determine the structure of the global liner shipping network in this research.

The liner shipping connectivity index consists out of five components that each is considered to be a representation or possible indictor of the connectivity of a country according to the definition of Wilmsmeier and Hoffmann (2008). These components are as described in Hoffmann et al., (2013):

1. Number of companies providing services from and to the countries port

This forms an indicator of the level of competition in a market: the more companies provide services from and to a country, the higher the contestability of the market, thus reducing the barriers to use that country to connect to the global network.

2. Capacity of the largest ship providing services from and to the countries ports

This component is an indicator of economies of scale and infrastructure as ports require adequate equipment in order to deal with certain vessel sizes. As is discussed in the theoretical framework a good quality of infrastructure of ports is required to attract more services by liner shipping companies and is likely to be positively related to trade volumes. Thus, economies of scale and infrastructure are also likely to reduce the barriers that prevent participating in the global liner shipping network.

3. Number of services that connect the countries ports to other countries

The more services pass through a port, the greater the probability to get a direct connection to a final destination without the need for transshipments. This would imply that this component is a proxy of the number destinations that one can easily reach from the researched country.

4. Total number of ships that are deployed in services from and to the countries ports

Although this component does not necessarily represent the level of frequency of services as this also depends on the length of all individual routes from and to that country and on the sailing speed of the vessels, it is still likely to imply a better connectivity ceteris paribus according to Hoffmann et al. (2013).

5. Total container carrying capacity in TEU of services from and to the countries ports.

Although this component does not mean that the capacity is also utilized, or even can be utilized, the potential for it is still there implying in turn more space available for potential growths in demand for container carrying capacity.

Hoffmann et al, (2013). Show several trends in the LSCI since 2004 when it was first created. They mention that the LSCI of Korea has surpassed the value of Japan and that Russia received a huge boost in 2012. The authors suggest several possible reasons for these two phenomena such as the Russia's joining the WTO and were thus allowing more foreign flagged ships to transport its foreign trade or more investments in Korea's infrastructure.

Another trend the authors mention is an increasing amount of concentration in the maritime containerized sector: The number of service providers per country has decreased between 2004 and 2013 with 20 percent.

In previous research, several significant relations have been proved between the LSCI and dependent variables such as trade costs. Figueiredo de Oliveira (2010) mentions that following his analysis, he found that the LSCI can be considered to be a key factor in the determination of freight rates its components, the capacity deployed, fleet characteristics and intensity of competition that are according to him all highly positive determinants of freight rates.

Wilmsmeier et al (2006) also found that the individual components of the LSCI are closely related to the freight rates: increasing liner services between a pair of ports leads to a reduction of freight rates. In a study by Wilmsmeier and Hoffmann (2008) the authors found that the number of liner shipping companies that provide direct services between a two countries seem to have a higher impact on the freight rates than the actual geographical distance.

In order to determine the relation between the LSCI and bilateral trade volumes at test on the country level connectivity index is performed here using a gravity equation with a BV-OLS model to test its effects on bilateral trade. The equation has the form:

$$ln\left(X_{ijk}\right) = \beta_0 + \beta_1 \ln(M) + \beta_2[\ln\left(Distance_{ij}\right) - ln(Distance_{PiPj})] + \beta_3\left[\ln(LSCI_i)\ln\left(LSCI_j\right)\right]_{PiPj} + \varepsilon_{ijk} \qquad (12)$$

In order to combine the two country level LSCI into a bilateral parameter that is used in equation XXX, they are averaged by computing the Euclidian norm as given in equation XXX

$$z_{ij} = \sqrt{z_i^2 + z_j^2} \qquad (13)$$

First, the relation between the natural logarithm of the LSCI and the bilateral directed trade is shown in figure 5.1. From this graph a positive relation between the LSCI and trade seems likely. This also follows from the results of the gravity regression as shown in table 5.1.

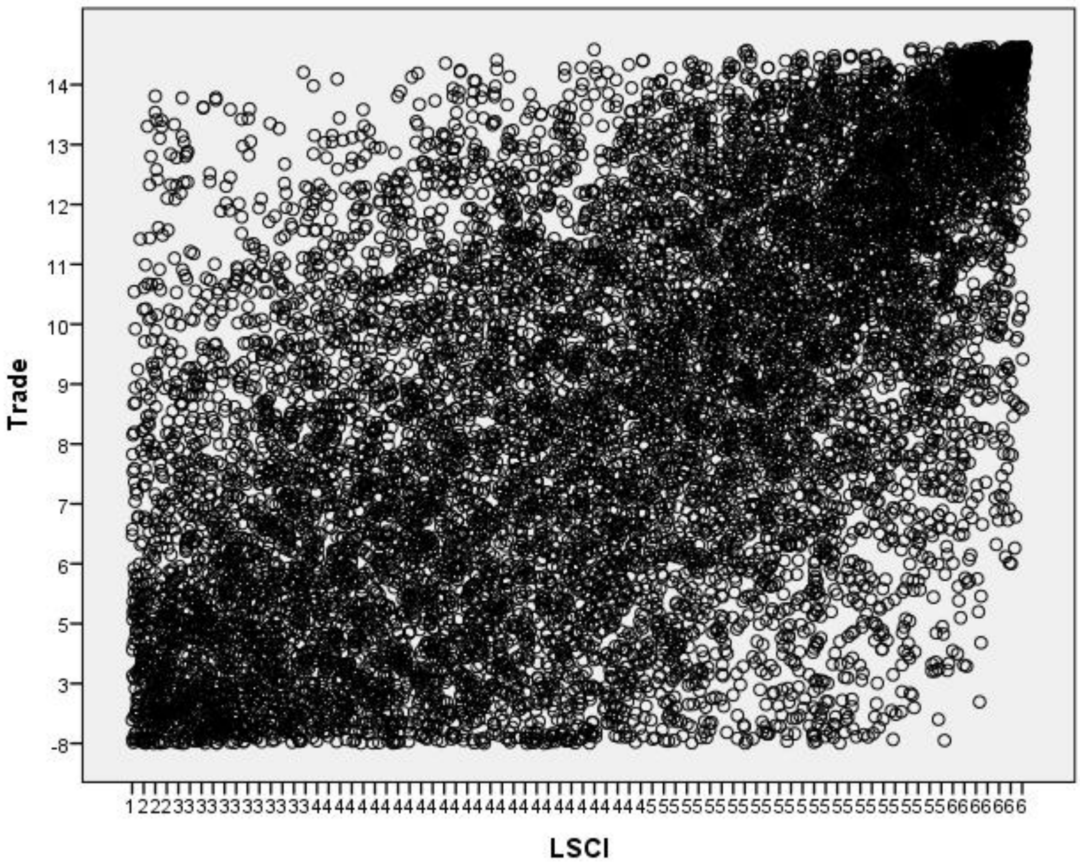


Figure 5.1 Scatterplot LSCI versus Bilateral Trade (Source:Author)

Table 5.1, results Gravity model with LSCI (Source: Author)

Dependent Variable Trade	**Model**
GDP	0.648^{***}
Distance	-0.286^{***}
LSCI	0.156^{***}
Adjusted R-Squared	0.632
N	11633
*, **, *** indicates significance at the 90%, 95%, and 99% level, respectively	

The coefficient of determination is 0.632, indicating that 63.2 percent of all the variation in the bilateral trade is explained by this model. As is expected, the distance is negatively related to trade. The product of the nominal GDP of both importing and exporting countries are positively related to the trade. Furthermore, the LSCI is, also as expected, positively related to bilateral trade. This indicates that the LSCI can be seen as a proxy of a smaller economic distance between two countries.

The above results also indicate that a one unit increase of the LSCI results in a percentage increase of bilateral trade of $100\left(e^{\beta}-1\right)=100(e^{0.156}-1)=16.88\%$.

5.3 Bilateral liner shipping connectivity index

As a result that only around 18 percent of all country pairs are directly connected with each other without the need for transshipment, the country level connectivity index is not sufficient to catch the notion of bilateral global liner shipping connectivity. Next to this, it was shown in chapter 5.2 that the gravity equation for trade is used with the averaged LSCI of both countries using the Euclidian norm, only around 63 percent of the variation in bilateral trade is explained, resulting in the need for a parameter that represents bilateral connectivity more accurately.

As a result of this UNCTAD is trying to create a bilateral liner shipping connectivity index (LSBCI) that can be used to better analyze international maritime containerized trade and transport (Ibid.). As mentioned earlier, the question however is what the components should be that form this bilateral connectivity index.

As previously discussed, the bilateral liner shipping connectivity index needs to determine how easy accessible trade is between any country pair. Thus, its components need to be a proxy of size of bilateral barriers that increase the economics distance between the two countries. Or, if described in terms of a complex network theory, the components need to be forces that attract both nodes towards each other.

In order to do this, first, the possible components will be identified in the next chapter using gravity equations in order to find those variables that can potentially represent (a part of) bilateral liner shipping connectivity. Next, several indices have to be created using these potential components and subsequently tested in order to determine what these indices represent, and thus whether they can act as a proxy of bilateral connectivity. This second part is done in chapter 7.

Chapter 6 Possible indicators of connectivity

In this chapter the four pillars of trade are investigated and tested against the trade values as the dependent variable using the gravity equation of trade. In this way, it can be determined what bilateral relations have positive relation with trade flows and can therefore be assumed to either actually lower the economic distance between two countries or to be a proxy for this reduction in economic distance. The results are subsequently discussed. First, the pillar of the network structure is analyzed, second the pillar of policy, third the pillar of ports and infrastructure and fourth the pillar of institutions and setting. Subsequently, these bilateral variables (or country level variables that have been averaged using the Euclidian norm) are used in the next chapter to generate several potential LSBCI's

In appendix E, an overview of the percentage change in bilateral trade as a result of an increase by one of all of the in this chapter researched independent variables is given.

6.1 Network structure related variables

For this pillar, the independent variables to be researched are those that form a description of the liner shipping network. As mentioned before, some research has previously been done on the structure of the global liner shipping network and its relation with maritime connectivity. However, these studies are either usually focused on port level structure and not on the country level. Furthermore, (Tobar et al., 2013) show that only a limited number of network parameters are used in these studies. Mostly used are the degree and betweenness centrality (with only a few authors also using clustering coefficients).

So, no research has previously been published on the relation between the (more in depth) characteristics of the global liner shipping structure at country level and trade volumes. As a result of this, a large selection of independent variables has been chosen in order to create a good representation of the network structure when investigating its relation with bilateral trade flows: the nodal network metrics that were presented in chapter 4. In the following table, an overview of the independent variables is given.

Table 6.1 Overview variables pillar network structure (source, author)

Independent variable	**Description**
M	The product of the nominal GDP of countries i and j
Distance	The maritime distance between countries i and j
Degree	Degree centrality -> the number of direct connections for a node
Strength	Indication of the relative position of a node compared to its directly connected neighbors
Betweenness	Number of possible position on

	shortstest path routes
Kcore	The maximal connected subgraph of the investigated network where all nodes have at least a degree of k
Cluster	Proportion of observed closed triplets in the sum of all possible closed triplets: a measure of density and tightness
Strahler	Numerical measure of the branching complexity in a directed network

For this regression, the BV-OLS is used again, giving the following equation:

$$
\begin{aligned}
ln\left(X_{ijk}\right) = & \propto_0 + \propto_1 \ln(M) + \propto_2 \left[\ln\left(Distance_{ij}\right) - ln(Distance_{PiPj})\right] \\
& + \propto_3 \left[\ln(Degree_i)\ln(Degree_i)\right]_{PiPj} + \propto_4 \left[\ln(Strenght_i)\ln(Strenght_i)\right]_{PiPj} \\
& + \propto_5 \left[\ln(Betweenness_i)\ln(Betweenness)\right]_{PiPj} \\
& + \propto_6 \left[\ln(Kcore_i)\ln(Kcore_i)\right]_{PiPj} + \propto_7 \left[\ln(Cluster_i)\ln(Cluster_i)\right]_{PiPj} \\
& + \propto_8 \left[\ln(Strahler_i)\ln(Strahler_i)\right]_{PiPj} + \varepsilon_{ijk}
\end{aligned}
$$

In appendix XXX the scatter plots are presented with the natural logarithm of each of the above metrics against the natural logarithm of the bilateral trade. These plots seem to indicate a positive relation between the degree centrality, betweenness centrality and K-core against the bilateral trade. Furthermore, there seems to be a negative relation between the Strength and the Cluster coefficient against the bilateral trade.

As there are a relative large amount of independent variables in this model, they are first checked for their partial correlations as presented below.

Variables	Degree	Strength	Betweenness	K-Core	Cluster	Strahler
Degree	1					
Strength	-482^{***}	1				
Betweenness	0.920^{***}	-0.526^{***}	1			
K-Core	0.693^{***}	-0.163^{***}	0.526^{***}	1		
Cluster	-0.676^{***}	0.672^{***}	-0.650^{***}	-0.455^{***}	1	
Strahler	0.12	0.061^{***}	-0.031^{***}	0.041^{***}	0.021^{**}	1

*, **, *** indicates significance at the 90%, 95%, and 99% level, respectively

As is shown in the results, there are strong significant partial correlations between most of the metrics in this pillar. The only metric whose correlation to the others is limited is the Strahler number. These results can be explained by the fact that they are all based on the same data: the structure of the overall network that follows from the direct connections between several of the countries.

Interesting to see is the strong partial correlation between the degree centrality and the betweenness centrality. This leads to the logical conclusion that countries that more direct connections to other countries are also likely to be countries that lie more on the shortest paths between other countries. Another interesting correlation is the one between the strength and the cluster coefficient.

As a result of the strong partial correlations between the metrics the results of the gravity equation will need to be tested for the multicollinearity, and if necessary, this will lead to several required adjustments of the model. The results are presented here:

Dependent Variable Trade	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**
GDP	0.734^{***} (1.000)	0.757^{***} (1.006)	0.685^{***} (1.510)	0.690^{***} (1.522)	0.698^{***} (1.5980)	0.692^{***} (1.618)	0.698^{***} (1.670)	0.698^{***} (1.674)
Distance		-0.285^{***} (1.006)	-0.284^{***} (1.006)	-0.285^{***} (1.007)	-0.287^{***} (1.010)	-0.285^{***} (1.012)	-0.284^{***} (1.013)	-0.284^{***} (1.014)
Degree			0.123^{***} (1.503)	0.121^{***} (1.505)	0.138^{***} (1.835)	0.099^{***} (2.581)	0.132^{***} (3.966)	0.133^{***} (15.154)
Strahler				-0.045^{***} (1.008)	-0.048^{***} (1.018)	-0.047^{***} (1.018)	-0.047^{***} (1.018)	-0.047^{***} (1.025)
Strength					0.037^{***} (1.592)	0.072^{***} (2.179)	0.085^{***} (2.430)	0.085^{***} (2.465)
Cluster						-0.083^{***} (3.3630)	-0.090^{***} (3.433)	-0.090^{***} (3.434)
Kcore							-0.045^{***} (2.628)	-0.045^{***} (2.975)
Betweenness								-0.002 (10.160)
Adjusted R-Squared	0.539	0.632	0.630	0.632	0.633	0.635	0.636	0.636
N.	11633	11633	11633	11633	11633	11633	11633	11633
*, **, *** indicates significance at the 90%, 95%, and 99% level, respectively The Variance Inflation Factor (VIF) is in parenthesis								

From this table it follows that the multicollinearity is only an issue in model H as is indicated with the high Variance Inflation Factors (the measure of how much the variance of the estimated coefficients are increased over the case of no correlation among the independent variables. If the VIF is around or greater than 5, it is an indication that there is collinearity associated with that variable) for the independent variables Degree and Betweenness.

One way of dealing with multicollinearity is by removing one of the variables that have a VIF greater than or around five individually. Model I is without the degree centrality and model F will be without the betweenness centrality:

Dependent Variable Trade	**I**	**F**
GDP	0.703^{***} (1.655)	0.694^{***} (1.653)
Distance	-0.285^{***} (1.014)	-0.283^{***} (1.013)
Strahler	-0.047^{***}	-0.047^{***}

	(1.018)	(1.018)
Strenght	0.085^{***}	0.086^{***}
	(2.465)	(2.429)
Cluster	-0.100^{***}	-0.090^{***}
	(3.441)	(3.480)
Kcore	-0.011	-0.150^{***}
	(1.982)	(2.794)
Betweenness	0.092^{***}	
	(2.659)	
Degree		0.136^{***}
		(4.190)
Adjusted R-Squared	0.637	0.636
N.	11633	11633
*, **, *** indicates significance at the 90%, 95%, and 99% level, respectively		
The Variance Inflation Factor (VIF) is in parenthesis		

Both models I and F have no issues with multicollinearity (although the VIF of the degree centrality in model F is high, it remains substantially below 5).

These models explain between 63.3 and 63.4 percent of the variation in the bilateral trade. Interesting to see are the significant positive Betas of the degree centrality, and the betweenness centrality. The Betas show that for every increase of the degree centrality by one, the bilateral trade tends to be 14.1% higher and for every increase of the betweenness centrality by one, the bilateral trade is 9.64% higher.

It thus indicates that that countries that have more direct connections to others (the degree centrality), and countries that act more as transshipment hubs (the betweenness centrality) also tend to attract more trade implying that both having more direct connections to other countries means you will attract more trade or that if you trade a lot you tend to get more direct connections to other countries. The same applies for the betweenness centrality: either being a transshipment hub means you will also attract more trade yourself or that if you trade a lot, you tend to increase your transshipments as well.

The significant positive Beta of the strength metric shows that countries that act as a bridge between clusters tend to trade less than countries that are part of a cluster in the overall global network. This could be explained by the fact that there are several production and consumption centers in the world that tend to trade a lot with each other. Countries that lie between those two regions might act as a bridge or as a transshipment country between these two 'clustesr' without actually trading themselves with the countries that are in these clusters. From its Beta it follows that an increase in the strength metric by one tends to lead to an increase in the bilateral trade with 8.9%.

The negative Betas for the cluster coefficient and the Strahler number show that if the directly connected countries of a country i tend to have many amongst each other or with other countries, the trade of country i tends to be lower than when this is not the case. This shows the importance of the relative position of a country in the network compared to its 'surrounding' countries: the better connected the countries are you are connected to, the more likely it is that you trade is lower.

The last metric researched for this subchapter is the K-core, which also has a (small) negative Beta. This result implies that countries that tend to be ‘overconnected’, meaning they are connected to more other countries than is necessary for them in order to remain part of their sub network also tend to be countries that trade less.

The results of this pillar suggest that being a transshipment hub can be a tautology of having a good access to the rest of the network. Furthermore, it also shows that if the attracting forces of a node are stronger than those it is directly connected with, trade also increases, thus proving that the relative position in the global liner shipping network matters when it comes to the trade flows. This corresponds to the idea of the multilateral resistances.

6.2 Policy related variables

This pillar tries to encompass those independent variables that represent effects of policy on trade and could also be used as an indication of or proxy for the bilateral connectivity. In order to achieve this, several variables have been identified in the literature review to have a potential relation with the trade flows. As a result of this, the main requirement for the variables used in this pillar is that they must be likely to be working as either attractive or repellant forces between countries in the global liner shipping network.

Therefore, the independent variables that are selected are the number of documents required for import and export, the days it takes to import or export a container, the costs of importing or exporting a container, the hours a year a company loses in order to pay its taxes, the number of individual procedures required to start new business and the days required in order to start up a new business.

In the following table are the dependant variables given and described:

Table 6.2, overview variables pillar policy (source: author)

Independent variable	Description
M	The product of the nominal GDP of countries i and j
Distance	The maritime distance between countries i and j
Doci	Documents required for export
Docj	Documents required for import
Timei	Days it takes to export
Timej	Days it takes to import
Costi	Costs of exporting container
Costj	Costs of importing container
Taxhours	the hours a year a company loses in order to pay its taxes
Procedures	Number of individual procedures required to start new business

StartDays	Days required when starting up a new business

Using the BV-OLS, it gives the following model to be solved to be solved:

$$
\begin{aligned}
ln\left(X_{ijk}\right) = \beta_0 + \beta_1 \ln(M) + \beta_2[\ln\left(Distance_{ij}\right) - ln(Distance_{PiPj})] \\
+ \beta_3\left[\ln(Doci_i)\ln\left(Doci_j\right)\right]_{PiPj} + \beta_4[\ln(Docj_i)\ln(Docj_i)]_{PiPj} \\
+ \beta_5\left[\ln(Timei_i)\ln\left(Timej_j\right)\right]_{PiPj} + \beta\left[\ln(Timej_i)\ln\left(Timej_j\right)\right]_{PiPj} \\
+ \beta[\ln(Costi_i)\ln(Costj_i)]_{PiPj} + \beta_7[\ln(Taxhours_i)\ln(Taxhours_i)]_{PiPj} \\
+ \beta_8[\ln(Procedures_i)\ln(Procedures_i)]_{PiPj} \\
+ +\beta \quad [\ln(StartDayes)\ln(StartDays_i)]_{PiPj} + \varepsilon_{ijk}
\end{aligned}
$$

In appendix F, the scatter plots of the natural logarithms of each of the independent variables against the natural logarithm of the bilateral trade are presented.

Here again, the partial correlations are determined first and are presented in the following table.

	Doci	Docj	Timei	Timej	Costi	Costj	TaxHours	Procedures	StartDays
Doci	1								
Docj	0.017^{*}	1							
Timei	0.776^{*}	0.009	1						
Timej	0.008	0.883^{***}	-0.002	1					
Costi	0.644^{***}	0.001	0.689^{***}	-0.1	1				
Costj	0.017^{*}	0.832^{***}	0.010	0.797^{***}	-0.046^{**}	1			
TaxHours	-0.049^{**}	-730^{***}	-0.053^{**}	-0.741^{***}	0.102	-0.779^{***}	1		
Procedures	-0.334^{**}	-0.479^{**}	-0.313^{**}	-0.484^{***}	-0.240^{**}	-0.440^{***}	0.428^{***}	1	
StartDays	-0.323^{**}	-0.472^{**}	-0.333^{**}	-0.512^{***}	-0.234^{**}	-0.406^{***}	0.430^{***}	0.769^{***}	1

*, **, *** indicates significance at the 90%, 95%, and 99% level, respectively

As could be expected, there are strong partial correlations beween the different variables that can represent the efficiency of exporting containers (Doci, Timei and Costj) and between those that can represent the efficiency of importing containers (Docj, Timej and Costj). This could imply a possible issue with multicollinearity between the three variables in both efficiency groups (imports and exports). If this is the case, it can be solved by combining the three variables in each group in one single variable.

There is also a strong partial correlation between the number of procedures required and the number of days it costs to start up a period. This is also a logical correlation as it is likely that when the number of procedures required increases, the time it takes to perform all these procedures will also increase.

The negative partial correlations between the variables Procedures and Startdays compared with the six variables that indicate the efficiency of importing and exporting in a country suggests that countries where it is easier to start a new business will have a lower efficiency when processing imports and exports.

Next, the model is tested for its multicolinearity, so the necessary adjustments to the model can be made if so required.

Dependent Variable Trade	A	B	C	D	E	F
M	0.735^{***} (1.000)	0.757^{***} (1.006)	0.750^{***} (1.029)	0.736^{***} (1.078)	0.734^{***} (1.083)	0.735^{***} (1.100)
Distance		-0.283^{***} (1.006)	-0.283^{***} (1.006)	-0.283^{***} (1.006)	-0.283^{***} (1.007)	-0.283^{***} (1.008)
Costi			0.047^{***} (1.023)	0.109^{***} (1.938)	0.118^{***} (2.049)	0.118^{***} (2.050)
Timei				-0.090^{***} (1.900)	-0.062^{***} (2.976)	-0.060^{***} (3.000)
Doci					-0.044^{***} (2.660)	-0.041^{***} (2.722)
Procedures						0.013^{**} (1.150)
Adjusted R-Square	0.540	0.620	0.622	0.626	0.627	0.627
N	11633	11633	11633	11633	11633	11633

*, **, *** indicates significance at the 90%, 95%, and 99% level, respectively
The Variance Inflation Factor (VIF) is in parenthesis

Dependent Variable Trade	G	H	I	J	K
M	0.736^{***} (1.143)	0.736^{***} (1.143)	0.733^{***} (1.239)	0.734^{***} (1.250)	0.734^{***} (1.253)
Distance	-0.283^{***} (1.008)	-0.283^{***} (1.008)	-0.282^{***} (1.013)	-0.282^{***} (1.013)	-0.282^{***} (1.014)
Costi	0.117^{***} (2.069)	0.118^{***} (2.073)	0.119^{***} (2.102)	0.119^{***} (2.107)	0.119^{***} (2.110)
Timei	-0.060^{***} (3.002)	-0.059^{***} (3.006)	-0.060^{***} (3.038)	-0.060^{***} (3.040)	-0.062^{***} (3.095)
Doci	-0.041^{***} (2.726)	-0.041^{***} (2.727)	-0.042^{***} (2.737)	-0.042^{***} (2.739)	-0.042^{***} (2.7409)
Procedures	0.011^{*} (1.413)	0.014^{**} (1.484)	0.011 (1.581)	0.012^{*} (1.595)	0.019^{**} (2.707)
TaxHours	0.004 (1.300)	0.016^{*} (2.684)	0.012 (2.994)	0.013 (2.996)	0.013 (3.006)
Costj		0.018^{*} (2.711)	0.025^{**} (3.596)	0.021^{*} (4.282)	0.022^{*} (4.326)
Timej			-0.014 (3.381)	-0.021 (5.228)	-0.024^{*} (5.4318)
Docj				0.012^{*} (5.793)	0.012^{*} (5.793)

StartDays					-0.011 (2.775)
Adjusted R-Square	0.627	0.627	0.627	0.627	0.627
N	11633	11633	11633	11633	11633
*, **, *** indicates significance at the 90%, 95%, and 99% level, respectively The Variance Inflation Factor (VIF) is in parenthesis					

First of all, the above presented results show the existence of multicollinearity in the model. Especially the variables Cost to import, Documents to import and Time to import show high VIF's. As these three variables tend to measure the same thing (the efficiency of operations when importing), this multicollinearity can be dealt with by joining the three variables together by use of the principle component analysis (PCA).

Furthermore, in the analysis of the partial correlations two other clusters are clear of sets of variables that are strongly correlated (with the correlations being higher than |0.3|). These are the variables that are each an indication of the difficulty of starting up a new business in a country: the variables StartDays and Procedures. The other set of variables that is closely related to each other are the three variables that indicate the efficiency of operations when exporting: Cost to export, Documents to export and Time to export. Therefore, a PCA is also applied to these two sets in order to reduce their dimensions to one.

The first step in the PCA is to assess the suitability of each set for it. The first requirement is that there have to be some partial correlations in the dataset higher than 0.3 (Beaumont, 2012). Other methods to test the suitability of the data set for PCA are by the Bartlett test of Sphericity and by the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (MSA) (Beaumont, 2012). The Bartlett test compares the identity matrix of the correlations with the correlations matrix itself. If the P value of this test is low, it is an indication that it is very unlikely that the observed correlation matrix is obtained from a population with zero correlation (Ibid.). The MSA can be seen as a summary of how small the partial correlations are relative to the original zero order correlations. It should be higher than 0.5 and preferably above 0.8 for the PCA to be suitable for the investigated dataset (Ibid.).

If a dataset is suitable for PCA, it can be performed in order to determine the factors or for each variable in the dataset. By subsequently taking the sum of each of the components in the data set multiplied by their respective factor, the new one dimensional variable has been created. In table 6.3, the three datasets are described after that are the results of the PCA are presented.

Table 6.3 variable sets for PCA (Source: author)

	Variables Set 1	**Variables Set 2**	**Variables Set 3**
Name set	ExportProcedures	ImportProcedures	Startups
Variable 1	Doci	Docj	StartDays

Variable 2	Timei	Timej	Procedures
Variable 3	Costi	Costj	

Table 6.4. Factors of PCA (Source: Author)

	ExportProcedures	ImportProcedures	Startups
MSA	0.763	0.748	0.782
Bartlett's Test of Sphericiy (P)	<0.001	<0.001	<0.001
Factor variable 1	0.952	0.905	0.944
Factor variable 2	0.954	0.835	0.944
Factor variable 3	0.932	0.882	

As the MSA and the Bartlett's Test of Sphericity are both within the acceptable range for all three sets, the PCA is suitable for these sets. As a result, the model as given in equation (???) is retested in the following form:

$$
\begin{aligned}
ln(Trade) = \beta_0 + \beta_1 \ln(M) + \beta_2 [\ln\left(Distance_{ij}\right) - ln(Distance_{PiPj})] \\
+ \beta_3 [\ln\left(ExportProcedures_{ij}\right) - ln(ExportProcedures_{PiPj})] \\
+ \beta_4 [\ln\left(ImportProcedures_{ij}\right) - ln(ImportProcedures_{PiPj})] \\
+ \beta_5 [\ln\left(StartUps_{ij}\right) - ln(StartUps_{PiPj})] \\
+ \beta_6 [\ln(Taxhours_i) \ln(Taxhours_i)]_{PiPj} + \varepsilon_{ijk}
\end{aligned}
$$

This gives the following results for the partial correlations is shown in table XXX

Variables	ExportProcedures	ImportProcedures	StartUps	TaxHours
ExportProcedures	1			
ImportProcedures	-0.027^{***}	1		
StartUps	-0.289^{***}	-0.465^{***}	1	
TaxHours	-0.05	-0.797^{***}	0.448^{***}	1

It shows that although even after the PCA there are a few strong partial correlations, they are less and smaller than there were as shown in table XXX. Interesting to see is that the correlation between the ExportProcedures and the other variables has decreased substantially. For the ImportProcedures, StartUps and Taxhours the correlations are still over 0.3.

When the above variables are solved in the gravity equation using the BV-OLS, it gives the results as shown below.

Dependent Variable Trade	A	B	C	D	E	F
GDP	0.735^{***} (1.000)	0.757^{***} (1.006)	0.753^{***} (1.018)	0.753^{***} (1.0182)	0.753^{***} (1.045)	0.755^{***} (1.076)
Distance		-0.284^{***} (1.006)	-0.284^{***} (1.006)	-0.285^{***} (1.006)	-0.287^{***} (1.006)	-0.285^{***} (1.006)
ExportProcedures			0.033^{***} (1.012)	0.041^{***} (1.102)	0.042^{***} (1.146)	0.042^{***} (1.146)
StartUps				-0.027^{***} (1.091)	-0.0030^{***} (1.443)	-0.028^{***} (1.474)
Importprocedures					0.007 (1.352)	0.023^{*} (2.924)
TaxHours						-0.023^{**} (2.900)
Adjusted R-Squared	0.540	0.620	0.621	0.621	0.621	0.621
N.	11633	11633	11633	11633	11633	11633

From the above table it follows that there are siginificant results for the variables ExportProcedures, StartUps, and TaxHours, but not for the ImportProcedures. Therefore, the regular BV-OLS will be performed with those three variables that gave the siginificant results as shown here.

Dependent Variable Trade	I
GDP	0.755^{***} (1.076)
Distance	-0.283^{***} (1.006)
ExportProcedures	0.040^{***} (1.129)
StartUps	-0.024^{***} (1.398)
TaxHours	0.006 (1.340)
Adjusted R-Squared	0.621
N.	11633
*, **, *** indicates significance at the 90%, 95%, and 99% level, respectively The Variance Inflation Factor (VIF) is in parenthesis	

These results show a significant Beta for the variables GDP, Distance, ExportProcedures and StartUps but not for the TaxHours. The most interesting result here is that the Beta of the ExportProcedures is positive implying that countries where it is easier and cheaper to export trade less. This counterintuitive result can be explained by the composition of this variable. In table XXX negative Beta's are shown for the number of documents required and number of days required to export a container. However, the variable cost to export has a stronger positive Beta. Thus, the fact that ExportProcedūces has a positive relation to the bilateraltrade can be explained by the fact that although it is still very likely that countries that have more days and documents required to export lead are also countries that trade less.

However, a country where the costs of exporting a container is higher, is likely to be a country with a better developed port infrastructure also implying that the country is more developed. And as more developed countries tend to trade more than (cheaper) developing countries, it can explain for the positive Beta of the variable ExportProcedures.

6.3 Ports and Infrastructure

This pillar looks that the quality of the ports and the infrastructure in a country. As has been mentioned before in the theoretical framework, a great variety in port choice determinants have been identified by previous research. However, all these determinants had a few things in common: most of them revolve either around the costs of using the port for its clients, the efficiency of the port and the quality of its services. As the issue of costs of using the port has been addressed in the pillar of policy (with the variables of cost of importing a container and costs of exporting a container), it is not longer researched in this pillar.

The efficiency of a port is measured in this pillar by the port performance indicator by the World Bank. This index is based on a yearly survey where experts are asked what their perception is of the quality of ports in each country. It can be expected that this index will show a positive relation with trade: the higher its value, the more experts like its quality, and thus the higher the incentive to use that port (according the theoretical framework).

The next determinant has to represent the quality of services in the port. This is measured in this pillar by means of the index for logistical performance by the World Bank. This index measures by means of a survey of freight forwarders and shippers the perceived ease of performing logistical operations inside a country. The reasoning behind choosing this variable in this pillar as a proxy for the quality of services is that this quality has been shown in the literature to be one of the main determinants of port choice by shippers and freight forwarders. As they will likely consider quality of services into their perception of how hard it is to do their business in a country, it is thus very likely that this index represents the quality of services. It can also be expected that there exist a positive relation between this index and the trade flows because when people perceive it to be easy to do business in a country, they are more likely to actually do their business in that country.

The final independent variable to be used in this pillar is whether one of the four largest global terminal operators is active in a country: This variable has been chosen as it van be a indication of the amount of privatization in a countries ports. Furthermore, it can also be assumed that these four operators perform at certain standards, allowing this dummy variable to be another indicator of service quality in ports.

In the following table, the independent variables are given:

Table 6.5. overview variables pillar ports and infrastructure (source: author)

Independant variable	Description
M	The product of the nominal GDP of countries i and j
Distance	The maritime distance between countries i and j
PortQuality	An index that shows the perceived quality of ports by experts
LogisticsQuality	An index that shows the perceived ease of performing logistical operations in a country by experts
Operatori	A dummy variable whether one of the four largest terminal operators is active in the exporting country
Operatorj	A dummy variable whether one of the four largest terminal operators is active in the importing country

This gives the following BV-OLS model:

$$\begin{aligned} ln(X_{ijk}) = \beta_0 + \beta_1 \ln(M) + \beta_2 [\ln(Distance_{ij}) - ln(Distance_{PiPj})] \\ + \beta_3 [\ln(PortQuality_i) - \ln(PortQuality_j)]_{PiPj} \\ + \beta_4 [\ln(LogisticsQuality_i) - \ln(LogistisQuality_j)]_{PiPj} \\ + \beta_5 [\ln(Operatori_i) \ln(Operatori_j)]_{PiPj} \\ + \beta_6 [\ln(Operatorj_i) \ln(Operatorj_j)]_{PiPj} + \varepsilon_{ijk} \end{aligned}$$

The partial correlations are again first determined and presented in table XXX and the scatter plots of the relation between the independent variables and the bilateral trade are given in appendix G :

Variables	PortQuality	LogisticsQuality	Operatori	Operatorj
PortQuality	1			
LogisticsQuality	0.313^{***}	1		
Operatori	0.121^{***}	0.135^{***}	1	
Operatorj	0.014	0.081^{***}	-0.398^{***}	1

The analysis of the partial correlations shows only two relatively large correlations: between LogisticsQuality and PortQuality and between Operatori and Operatorj. They imply that countries that have a better perceived quality of their logistics are also more likely to be perceived to be a country with a high quality of ports. Furthermore, the negative correlation between Operatori and Operatorj indicate that if one of the four largest global terminal operators is active in a country, the change that one of them is also active in its trading partner decreases. The results of the stepwise BV-OLS regression are shown in table XXX

Dependent Variable Trade	**A**	**B**	**C**	**D**	**E**	**F**
GDP	0.735^{***} (1.000)	0.757^{***} (1.006)	0.757^{***} (1.007)	0.755^{***} (1.015)	0.739^{***} (1.226)	0.738^{***} (1.227)
Distance		-0.283^{***} (1.006)	-0.283^{***} (1.006)	-0.287^{***} (1.033)	-0.285^{***} (1.0360)	-0.285^{***} (1.036)
Operatorj			-0.028^{***} (1.001)	-0.030^{***} (1.007)	-0.030^{***} (1.007)	-0.026^{***} (1.217)
LogisticsQuality				0.028^{***} (1.044)	0.048^{***} (1.169)	0.039^{***} (1.194)
PortQuality					-0.040^{***} (1.308)	-0.041^{***} (1.318)
Operatorj						-0.011 (1.239)
Adjusted R-Squared	0.540	0.620	0.620	0.621	0.622	0.622
N.	11633	11633	11633	11633	11633	11633

This shows that multicollinearity is not an issue in this model. However, the variable Operatori does not give a significant result. Therefore, the regular BV-OLS has to be regressed using all variables but Operatorj. This result is given in table XXX

Dependent Variable Trade	**I**
GDP	0.739^{***} (1.226)
Distance	-0.285^{***} (1.036)
Operatorj	-0.030^{***} (1.007)
LogisticsQuality	0.041^{***} (1.169)
PortQuality	-0.040^{***} (1.3080)
Adjusted R-Squared	0.622
N.	11633
*, **, *** indicates significance at the 90%, 95%, and 99% level, respectively The Variance Inflation Factor (VIF) is in parenthesis	

In this model all the VIF's remain just above one, showing that multicollinearity is not an issue here. It further indicates that countries that have a higher perceived quality of logistics also tend to be countries that trade more. The siginificant Beta of 0.041 shows that for every increase in the Logistics Quality index by one point for a country, the bilateral trade in manufactured goods tends to be 4.18 percent higher. However, the PortQuality shows a negative Beta, indicating that countries that have a higher perceived port quality tend to trade less. This can be explained by several reasons. The most likely is the response bias by the exports whose answers and opinions form this index. If the experts are more optimistic in assessing the quality of ports in developing countries than in developed countries, it can result in an inflated index for the developing countries explaining the negative Beta as found in this analysis.

The third independent variable with a statistically significant result is the Operatorj. Its negative Beta shows that countries that have one of the main global container terminal operators' active in at least one of their ports also tend to be countries that import less.

6.4 Institutions and international

In the last pillar, institutions and international relations and situations are investigated. It was shown in the theoretical framework that an equal level of quality of institutions between countries increases trade as a result increased similarities of doing business between the two countries. Furthermore, it is likely that common cultural characteristics will also reduce the barriers to trade for the same reasons. Therefore, it has been chosen to use independent variables in this pillar that can represent the cultural relation between two countries and that represent the quality of the institutions. As a result of this, several dummy variables have been included in this pillar that can represent cultural and institutional similarities. These are common language, common currency, a colonial relation but also whether both countries are included in a political or continental union and whether they are joined in a customs union. Furthermore, a dummy has also been added representing whether a war or high civil unrest took place in one of the countries in the country pair in the last five years. It can be expected that this will have negative impact on trade, as it will have likely reduced the quality of the countries institutions.

There are also three more indices included that represent the quality of the institutions. These are the difficulty of enforcing a contract in a country through means of a procedure, an index that shows how much effort the society in a country shows in fighting corruption and the strength of the legal rights: an index that measures the degree to which collateral and bankruptcy laws protect the rights of borrowers and lenders. All these three indices have been acquired from the World Bank.

As was stated in the theoretical framework, it has been shown in the literature that countries that have a common sea border and have no common land border tend to trade more with each other than countries that share no common sea or land border. Therefore it can be stated that a common sea border reduces the effective geographical maritime distance and thus acts as a potential force between a country pair. As a result of this, a dummy as to whether a country pair shares a common sea is included in this analysis.

On a final note, two more dummies have been added. One is whether a country is an island, and 1 is whether a country has a Freeport. In the following table an overview of all the variables is given

Table 6.6 overview variables pillar institutions and setting (source: Author).

Independent variable	Description
M	The product of the nominal GDP of countries i and j
Distance	The maritime distance between countries i and j
Island	Dummy whether country is an island
Freeport	Dummy for the existence of a Freeport in one of the countries
Customs	Dummy for participation in same customs union
PoliticalUnion	Dummy for participation in same political or continental union
Currency	Dummy for a common currency
Border	Dummy for a common land border
Colony	Dummy for colonial relations
Language	Dummy for common language
Enforcing	Index that shows the difficulty of enforcing a contract in a country through means of a procedure
CorruptionFight	Index that that shows how much effort the society in a country shows in fighting corruption
LegalRights	index that measures the degree to which collateral and bankruptcy laws protect the rights of borrowers and lenders
Seaborder	A dummy that indicates whether the coast of both countries in the country pair is at the same sea or ocean.

This gives the following BV-OLS that has to be solved:

$$
\begin{aligned}
& ln(X_{ijk}) \\
& = \beta_0 + \beta_1 \ln(Y_i Y_i) + \beta_2 [\ln(Dist_{ij}) - ln(Dist_{PiPj})] + \propto_3 [\ln(Isl_i) \ln(Isl_j)]_{PiPj} \\
& + \propto_4 [\ln(Frep_i) \ln(Frep_i)]_{PiPj} + \propto_5 [\ln(Customs_{ij}) - ln(Customs_{ij})_{PiPj}] \\
& + \propto_6 [\ln(PoliticalU_{ij}) - ln(PoliticalU_{ij})_{PiPj}] \\
& + \propto_7 [\ln(Currency_{ij}) - ln(Currency_{ij})_{PiPj}] + \propto_8 [\ln(Border_{ij}) - ln(Border_{ij})_{PiPj})] \\
& + \propto_9 [\ln(Colony_{ij}) - ln(Colony_{PiPj})] + \propto_{10} [\ln(Language_{ij}) - ln(Language_{PiPj})] \\
& + \propto_{10} [\ln(war_{ij}) - ln(war_{PiPj})] \\
& + \propto_{11} [\ln(Enforcing) \ln(Enfocing_i)]_{PiPj} + \propto_{12} [\ln(Corruption) \ln(Corruption_i)]_{PiPj} \\
& + \propto_{13} [\ln(legalR_{ij}) - ln(LegalR_{PiPj})] + \varepsilon_{ijk}
\end{aligned}
$$

Also for this pillar, the scatter plots that show the relation between each of the independent variables and the bilateral trade are presented in Appendix H and in appendix I the results of the partial correlations analysis is performed.

As can be seen in Appendix I, the variables in this pillar have very low partial correlations: only the currency and the customs variables have a partial correlation higher than 0.3. This implies that countries that have the same currency are also more likely to be joined in a Customs union. Next, the stepwise regression is performed and results are shown in Appendix I. As was expected following the analysis of the partial correlations, there is no issue of multicollinearity in this model, as all VIF's are below 1.5. Following the analysis, the model for this pillar should include all the variables except for Border and Currency. This gives the following results as shown the following table

Dependent Variable Trade	**Model**
GDP	0.737^{***} (1.349)
Distance	-0.269^{***} (1.350)
LegalRights	-0.085^{***} (1.057)
Language	0.059^{***} (1.057
Freeport	0.056^{***} (1.085)
Enforcing	0.039^{***} (1.099)
Island	0.042^{***} (1.132)
PoliticalUnion	0.057^{***} (1.181)
Colony	0.024^{***} (1.032)
CommonSea	0.024^{***} (1.085)

CorruptionFight	0.020^{***} (1.049)
Customs	-0.020^{***} (1.663)
Adjusted R-Squared	0.640
N.	11633
*, **, *** indicates significance at the 90%, 95%, and 99% level, respectively The Variance Inflation Factor (VIF) is in parenthesis	

This model thus explains 64% of all variations in bilateral trade. The most interesting result in this pillar is that according to this analysis, being in a customs union with another country does not relate to higher bilateral containerized maritime trade with said country.. A possible explanation for this result is that the customs union is related to a increase in land based trade, leading to a reduction of maritime trade. The other interesting result as shown above is the negative Beta for the LegalRights variable. This indicates that countries that that have better laws to protect borrowers and lenders also tend to be countries that trade less. All the other results of this analysis however seem to comply with what is written about them in the scientific literature.

6.4 Further analysis of possible potential components for a LSBCI

From the above analysis it follows that several country level parameters and several bilateral parameters seems to have a clear relation with trade. The question however is whether they can be used together to as components of a bilateral liner shipping connectivity index. Therefore, those who had a significant relation are tested here against undirected bilateral trade in manufactured goods, and these results are further discussed.

An overview of the variables that had a relation with trade following the above analysis is:

Table 6.7, results from gravity equations (source: author)

Variable	Relation with trade
Betweennes	.092
Kcore	048
Strahler	-0.047
Strength	0.086
Cluster	-0.090
Degree	0.136
ExportProcedures	0.040
StartUps	-0.024
Operatorj	-0.030
LogistisQuality	0.041
PortQuality	-0.040
LegalRights	-0.085
Language	0.059

Freeport	-0.056
Enforcing	0.039
Island	0.042
PoliticalUnion	0.057
Colony	0.024
CommonSea	0.024
CorruptionFight	0.020
Customs	-0.020

First, the relation between each of these variables is tested against the natural logarithm of the bilateral trade volumes after again correcting for multilateral resistances. Te result of this analysis is given below:

Table 6.8 relations variables with bilateral trade (source: Author)

Variable	Coefficient of determination	Beta	Significance	1 point increase in variable leads to % change in bilateral trade
Betweennes	0.216	0.465	<0.01	59.20
Kcore	0.167	0.408	<0.01	50.382
Strahler	0	0.18	<0.05	19.72
Strenght	0.143	-.378	<0.01	-31.47
Cluster	0.238	-.488	<0.01	-38.61
Degree	0.259	0.509	<0.01	66.36
ExportProcedures	0.013	0.112	<0.01	11.85
StartUps	0	-0.11	>0.1	-10.41
Operatorj	0	-.017	<0.1	-1.68
LogistisQuality	0.03	0.058	<0.01	5.97
PortQuality	0.102	-.319	<0.01	-27.31
LegalRights	0.036	-.190	<0.01	-17.30
Language	0.05	0.071	<0.01	7.35
Freeport	0.066	0.258	<0.01	29.43
Enforcing	0.026	0.162	<0.01	17.58
Island	0.042	0.206	<0.01	22.87
PoliticalUnion	0.028	0.168	<0.01	18.29
Colony	0.001	0.026	<0.05	2.63
CommonSea	0.09	0.093	<0.01	9.74
CorruptionFight	0.02	0.046	<0.01	4.70
Customs	0.027	0.163	<0.01	17.70

Amongst the researched variables, only one did not give a significant result: the StartUps. All the other results where significant on a 1 percent, 5 percent and in one case on a 10 percent level. However, interesting to see is that this analysis shows none of the variance in trade can be explained by the Strahler number and by the variable Operatorj. The other results are more promising for their use in a LSBCI. Especially the variables related to the network structure itself show strong relations with the bilateral trade volumes.

As all the variables above that are country level data have been joined together using the Euclidian normalization it has to be taken into consideration that 1 point increase in these variables does not mean that actual country level variable increases by 1, but that the square root of the sum of the squares of the variables of both importing and exporting country increases by 1.

Of the variables that represent the network structure, especially the variables of the betweenness centrality, the cluster coefficient and the degree centrality show very strong results. They alone can respectively explain for 21.6, 23.8 and 25.9 percent of all the variation in bilateral containerized trade. The other two remaining variables, the K-Core and the Strength are also able to explain a large portion of the variance in bilateral trade with 16.7 and 14.3 percent. This seems to indicate that the position within the network can act as a accurate proxy for bilateral trade volumes.

For the policy related variables, only the ExportProcedures variable gave statistically significant result that can explain the variation in bilateral trade. Interesting is however that its Beta is positive, indicating that countries that have more and expansive export procedures tend to be the countries that also export more. As mentioned before, this result can be explained with the composition of this variable: the cost of exporting a container is likely to be higher in ports with a more developed infrastructure and services, thus in developed countries. This effect is strong enough to offset the negative relation of the number of documents and number of days required to export a container.

For the variables related to ports and infrastructure, only the PortQuality and the LogistiscQuality show significant results and can explain a portion of the variation in bilateral trade with 3 and 10.2 percent. Noticeable here is the strong negative Beta of the port quality variable. There are several possible explanations for this. It could also be that countries that trade more tend to be more congested at ports reducing the perception of quality of said ports.

With respect to the variables related to the institutions and setting, all variables are able to explain the variance in bilateral trade with values ranging between 0.1 to 9 percent. One of the more interesting results here is the strong positive Beta of the variable Customs, as it is negative in the analysis in chapter 5. This implies that the multicollinearity did play a significant role there even if all the VIF's where relatively low.

6.5 The Components of a LSBCI

The final step in this chapter is to put together all the variables that have been identified above that have a statistically significant relation with bilateral trade and that are a component of liner shipping. Although many of the predictors of trade presented before show statitistically significant relations with the bilateral trade, they cannot all be used as a proxy of bilateral connectivity. The reason for this is that only a limited number of variables can be considered to be largely a component of liner shipping.

All the metrics of the liner shipping structure that show a statistically significant relation with bilateral trade however can be considered to be a potential component of the LSBCI as they are purely a result of the liner shipping structure itself. Next ot this, the variable PortQuality can also be used as a component of the LSBCI as it represent how experts perceive the quality of ports. However, as a result of its negative Beta, It can be assumed that ports with a lower perceived quality tend to have more business and activities. This needs to be taken into consideration when placing this variable as a component it in the index. The final variables that can be used as components of the LSBCI are the LSCI itself as it is an indication of the country level connectivity and whether two countries share a common sea.

An overview of these variables is given in table 6.9:

Table 6.9. Overview possible components for LSBCI (Source: author)

Variable	Reason of choosing
Betweennes	A characteristic that represents the structure of the liner shipping network
Kcore	A characteristic that represents the structure of the liner shipping network
Strenght	A characteristic that represents the structure of the liner shipping network
Cluster	A characteristic that represents the structure of the liner shipping network
Degree	A characteristic that represents the structure of the liner shipping network
PortQuality	Is an indication of the perceived quality of ports in a country
LSCI	Indication of the country level connectivity to the overall liner shipping network

Chapter 7 LSBCI and its components

In this chapter, the possible components of the LSBCI are used to generate a set potential indices. These indices are then tested against bilateral trade using the gravity equation and the results are discussed.

Seven potential components have been identified above. Five of these are metrics of the liner shipping structure, one is an indication of the port quality and the final component is a representation of a countries overall connectivity to the whole network. Therefore, if it is assumed that the LSBCI should consist out of four components as is the case for the LSCI it seems logical to include two of the network metrics, the port quality and the LSCI in the LSBCI. However, as the variable PortCluster does not have a value for all the routes (this index has not been created for all countries in the world) several indices should also be tested where the port quality is not included.

In order to determine the weight of each of the components for the index, a PCA is performed. As a result of this, the following potential LSBCI's are:

7.1 LSBCI model A (source: Author)

Name model	**LSBCI A**
MSA	0.721
Bartlett's Test of Sphericiy (P)	<0.01
Components	**Factors**
Degree	0.972
Betweenness	0.982
LSCI	0.989
PortQuality	0.325

7.2 LSBCI model B (source: Author)

Name model	**LSBCI B**
MSA	0.8
Bartlett's Test of Sphericiy (P)	<0.01
Components	**Factors**
Degree	0.917
Betweenness	0.939
LSCI	0.959
Strenght	0.489

7.3LSBCI model C (source: Author)

Name model	**LSBCI C**
MSA	0.796
Bartlett's Test of Sphericiy (P)	<0.01
Components	Factors
Betweenness	0.957
Cluster	0.842
LSCI	0.957
K-Core	0.838

7.4 LSBCI model D (source: Author)

Name model	**LSBCI D**
MSA	0.838
Bartlett's Test of Sphericiy (P)	<0.01
Components	Factors
Degree	0.968
Betweenness	0.967
Cluster	0.847
LSCI	0.979

Subsequently, the four generated potential LSBCI's have to be tested against the bilateral trade using the gravity model:

$$ln\left(X_{ijk}\right) = \beta_0 + \beta_1 \ln(M) + \beta_2[\ln\left(Distance_{ij}\right) - ln(Distance_{PiPj})] + \beta_3\left[[\ln\left(LSBCI_{ij}\right) - ln(LSBCI_{PiPj})\right] \quad + \varepsilon_{ijk}$$

This gives the following results:

Dependent Variable Trade	**A**	**B**	**C**	**D**
GDP	0.693^{***}	0.693^{***}	0.693^{***}	0.693^{***}
Distance	-0.285^{***}	-0.285^{***}	-0.285^{***}	-0.285^{***}
LSBCI	0.116^{***}	0.118^{***}	0.117^{***}	0.119^{***}
Adjusted R-Squared	0.630	0.631	0.631	0.633
N	11633	11633	11633	11633
*, **, *** indicates significance at the 90%, 95%, and 99% level, respectively				

These results show that there hardly is any difference in the four generated LSBCI's when they are regressed against bilateral trade. This implies that variations amongst the network metrics as components for a LSBCI do not matter much: it can thus be stated one component to represent the network structure is likely to be sufficient. Furthermore, these results are also similar to those in chapter 5, where only the LSCI was used in a gravity equation to explain the variations in bilateral trade. This is also an indication that different types of components need to be added to the LSBCI. On a side note however, the LSBCI does not need to capture and explain trade, but it must capture the amount of goods transported between two countries. Thus, it should be possible for a route to have a high LSBCI but low volumes of trade between the origin and destination but a high volume of cargo being transported between them as a result of transshipments.

Chapter 8 Conclusion and recommendations

In this conclusion, the central research question of this paper will be answered resulting in several recommendations for the creation of a bilateral liner shipping connectivity index. In the end of this chapter, there are also several more recommendations made for future research with respect to this topic of this paper.

8.1 Conclusions

One of the questions that arose in the beginning of this thesis was how to define the connectivity between any two countries in the global liner shipping network. I the theory there are several different definition of this concept, ranging from being an attribute that refers to quality and costs to move freight to the access to regular and frequent transport services combined with the level of competition. The description of the concept of connectivity in this paper was derived using complex network theory combined with the characteristics of liner shipping.

As the main characteristic of the liner shipping network is that that it provides a fixed set of services at a fixed frequency at fixed routes. However, not all countries in the world have a direct connection with each other: only around 17 percent of all countries are directly connected to each other. This results in a complex network of a number of direct and a number of indirect country pair combinations. Although the structure of this network is undergoing a constant change as a result to changing economic parameters and due to shifts in manufacturing and consuming centers in the world, the overall underlying theoretical principles remain the same.

If it is assumed that at any given time the structure of the network is the equilibrium state of all the forces that act on all the nodes, it can be concluded that the relation between each node depends on the relative forces that act between them: the sum of all attracting and all repellent forces between any two countries determines their relationship and how ‘close they are to each other’. Following this, it can thus be stated that two countries have a good and strong connection if the relevant forces that determine the structure of the network have such a direction that the two countries are pulled towards each other. Therefore, the measure of connectivity can thus be described as the economic distance between any two countries.

In order to find a representation of this connectivity or proxy for it, it must thus be identified what the bilateral and country level forces are that determine the structure of the global liner shipping network. From the literature, a set of variables was selected that could potentially be either these forces themselves, or a proxy for them.

These ‘forces’ were then tested against the bilateral trade flows. The results showed that there is a strong relation between the structure of the network itself and the amount traded. Especially the betweenness centrality, cluster coefficient and the degree centrality showed a strong relationship to bilateral trade. Therefore, it can be concluded that the network structure metrics are a good proxy for the economic distance between countries, and could thus potentially be used to indicate the level of connectivity between any country pair. These three mentioned characteristics of the network structure indicate that a country that attracts a lot of transshipment as a result of its position in the network also tends to be a country that trades more. It also shows the importance of the relative position of a country compared to those it is directly connected to.

Furthermore, there is also a strong relation between the country level connectivity to the global network (as measured by the LSCI as created by UNCTAD every year) and the trade flows. This is logical as there is also a strong relation between the LSCI and trade costs and country level trade. From this, it can also be concluded that the LSCI is a likely component of a bilateral connectivity index.

Other possible determinants that can act as a proxy of the forces that work in the network are those related to amount of economic and cultural integration two countries show: if there are many similarities between countries, it is more likely that they have similar business practices resulting in a reduction in economic distance. This was also shown in the performed analyses where a positive relation was proven between maritime trade flows and variables such as common language and common currency.

Therefore, it can be concluded that the main relation between bilateral liner shipping connectivity and its potential determinants is that the determinants must be a proxy for the attractive forces that work between the nodes in the network. Resulting from this, determinants that describe the structure of the network, such as the betweenness centrality and the cluster coefficient should be the main components of a bilateral liner shipping connectivity index together with the already existing LSCI. With these components, a LSBCI represents not only the result of all the attracting and repelling forces acting in the network on each country, but also the country level connectivity to the overall network.

However, several LSBCI's were generated with these components and tested against the bilateral trade flows. The results showed great similarities independent of its components and were similar to using the LSCI alone to explain the varations in bilateral trade. Therefore, it can be concluded that the final LSBCI should consist out of more than a combination of several metrics that describe the network structure and the LSCI.

8.2 Recommendations

Thus, in order to create a LSBCI in the future, more possible components need to be found that can represent bilateral connectivity next to the already existing LSCI and the for this paper determined network characteristics. Further potential components could focus for instance the competition on a route or on the effect of bottlenecks in the network, such as the presence of canals on routes between two countries.

There are also several recommendations to be made for future research. As a result of the high dependency on the quality of the data, this research could be improved if several sets of data on trade volumes and values were available and compared with each other using statistical methods in order to get a more reliable data set to test the independent variables against.

Furthermore, tariffs have not been included in this research despite having a known impact on trade volumes. It can therefore be recommended for future research to look at the relation of tariffs on bilateral connectivity as this might have created a distortion in the results and analysis. On a final note, it is also still required to determine how to weight the different components for a bilateral index as this will also greatly affect the relations of a future index with parameters such as trade costs and trade volumes.

References

Anderson, J. E., & Van Wincoop, E. (2004). *Trade costs* (No. w10480). National Bureau of Economic Research.

Auber, D. (2002). Using Strahler numbers for real time visual exploration of huge graphs. In *International Conference on Computer Vision and Graphics*. Vol.1,No. 1, pp. 3-21.

Bergmann (1998). Comments on Deardoff (1998)"Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical Framework?" in J.A. Frankel (ed.) The Regionalization of the World Economy, Chicago: University of Chicago Press

Bensassi, S. Martinez-Zarzoso, I. (2010). Piracy and international trade: an empirical study. *Proceedings of the 20101 IAME Conference. Lisbon.*

Bensassi, S. Martinez-Zarzoso, I. Burguet, C. (2011).*The uneven effect of maritime transport costs on international trade. Proceedings of 2011 IAME conference*

Baier, S. L., & Bergstrand, J. H. (2007). Do free trade agreements actually increase members' international trade?. *Journal of international Economics*,*71*(1), 72-95.

Baier, S. L., & Bergstrand, J. H. (2009).Estimating the effects of free trade agreements on international trade flows using matching econometrics. *Journal of International Economics*, *77*(1), 63-76.

Baier, S. L., Bergstrand, J. H., &Feng, M. (2011). Economic integration agreements and the margins of international trade. *University of Notre Dame, mimeo*.

Beaumont, R. (2012). Principal Component Analysis & Factor Analysis. *Virtual Classroom.*

Bergstrand, J. H., Egger, P., & Larch, M. (2013). Gravity Redux Estimation of gravity-equation coefficients, elasticities of substitution, and general equilibrium comparative statics under asymmetric bilateral trade costs.*Journal of International Economics*, *89*(1), 110-121.

Burger, M., Van Oort, F., &Linders, G. J. (2009). On the specification of the gravity model of trade: zeros, excess zeros and zero-inflated estimation. *Spatial Economic Analysis*, *4*(2), 167-190

Caldeirinha, V. Augusto Felicio, J. Dionisio, A. (2013). Effect of the container terminal characteristics on performance.*Proceedings of the 2013 International Association of Maritime Economics (IAME) Conference*, Marseille, Marseille.

Cariou, P. (2010). Is slow steaming a sustainable mean for reducing liner shipping CO2 emissions? In: *Euromed Management Mare Forum.*

Castillo-Manzano, J. Gonzalez-Laxe, F. (2011).*Analyzing the factors that determine the evolution of traffic captrued in shared and highly disputed hinterlands.The Spanish Experience.* Proceedings of the 2011 IAME conference.

Caschilli, S. &Medda, F. R. (2012).A review of the Maritime Container Shipping Industry as a complex adaptive system. In: *Interdisciplinary Description of Complex Systems.* Vol. 10, No. 1, pp. 1-15.

De Groot, H. L., Linders, G. J., Rietveld, P., & Subramanian, U. (2004).The institutional determinants of bilateral trade patterns. *Kyklos*, *57*(1), 103-123.

Dorogovtsev, S. N., Goltsev, A. V., & Mendes, J. F. F. (2006). K-core organization of complex networks. *Physical review letters*, Vol. *96, No* 4, pp. 140-160.

Ducruet, C., &Notteboom, T. (2012). The worldwide maritime network of container shipping: spatial structure and regional dynamics. *Global Networks*,*12*(3), 395-423.

Ducruet, C. Zaidi, F. Inria, B. (2012). Maritime Constellations: A complex network approach to shipping and ports. In: Maritime policy and Management. Vol. 39, No. 2,pp 151-168.

Ducruet, C. (2013). Network Diversity and Maritime Flows.*Proceedings of the 2013 International Association of Maritime Economics (IAME) Conference*, Marseille, Marseille.

Ducruet, C., & Lugo, I. (2013). Structure and dynamics of transportation networks: Models, methods and applications. *The SAGE Handbook of Transport Studies*, 347-364

Dutt, P. &Traca, D. (2010). Corruption and Bilateral Trade Flows: Extortion or Evasion? In: *The Review of Economics and Statistics.* Vol. 92, No. 4, pp. 843-860.

Egger, P. (2000).A note on the proper econometric specification of the gravity equation. In: Economic Letters. Vol. 66, No. 1, pp. 25-31,

Eichengreen, B. & Irwin, D. (1998).The Role of History in Bilateral Trade Flows.in J.A. Frankel (ed.) The Regionalization of the World Economy, Chicago: University of Chicago Press

Fedou, J. (1999). *Nombre de Strahler sur les arbres généraux*. Ecole jeunes Chercheur en algorithmique et calcul formel, Bordeaux.

Figueiredo de Oliveira, G. (2010). European andMediterraneanliner trade routes: determinants of freight rates. Proceedings of the 20101 IAME Conference.

Figueiredo De Oliveira, G. (2013). Determinants of European Freight Rates: the role of Market Power and Trade Imbalance. *Proceedings of the 2013 International Association of Maritime Economics (IAME) Conference*, Marseille, Marseille.

Fageda, X. Gonzalez-Aregall, M. (2013). Regulation of Port Charges in Spain: Global versus Local Competition.*Proceedings of the 2013 International Association of Maritime Economics (IAME) Conference*, Marseille, Marseille.

Farole, T. (Eds). (2013). *The internal geography of Trade: Lagging regions and global Markets.* The World Bank.

Feenstra, R. (2004). *Advanced International Trade: Theory and Evidence.* Princeton University Press, Princeton.

Frick, A., Ludwig, A., & Mehldau, H. (1995, January). A fast adaptive layout algorithm for undirected graphs (extended abstract and system demonstration). In *Graph Drawing* (pp. 388-403). Springer Berlin Heidelberg.

Garcia-Alonso, L. Sanchez-Soriano, J. (2009).Port Selectionfrom a hinterland perspective. In: *MaritimeEconomics&Logistics.* Vol. 11, No. 8, pp 260-269

Glick, R. & Rose, A. (2002). Does a currency union affect trade? The time series evidence. In: *European Economic Review.* Vol. 46, pp. 1125-1151.

Guan, C. Yahalom, S. (2013). Potential Changes in Vessel Deployment Patterns after Panama Canal Expansion. *Proceedings of the 2013 International Association of Maritime Economics (IAME) Conference*, Marseille, Marseille.

Guerrero, D. Rodrigue, J. (2013). The Waves of Containerization: Shifts in Global Maritime Transportation. Proceeding of the International Association of Maritime Economists AnnualConference, Marseille.Marseille.

HARALAMBIDES, H. E., 2004, Determinants of price and price stability in liner shipping, Workshop on The Industrial Organization of Shipping and Ports National University ofSingapore.

Herman, I., Delest, M., & Melancon, G. (1998). Tree visualisation and navigation clues for information visualisation. In *Computer Graphics Forum.* Vol. 17, No. 2, pp. 153-165

Hoffmann, J. Benson. Wilmsmeier, G. (2013). Bilateral Liner Shipping Connectivity.*Proceedings of the 2013 International Association of Maritime Economics (IAME) Conference*, Marseille, Marseille

Hu, Y., & Zhu, D. (2008, June).Empirical analysis of the worldwide maritime transportation network.Physica A, 388(1061-2071).

Hutchinson, W. (2005). Linguistic Distance As a Determinant of Bilateral Trade. In: Southern Economic Journal. Vol. 72, No. 1, pp. 1-15.

Karamperides, S. Mangan, J. Jackson, E. (2013*).* Development of an index for maritime container transport costs, connectivity and risks. *Proceedings of the 2013 International Association of Maritime Economics (IAME) Conference*, Marseille, Marseille.

Marquez Ramos, L. (2006). Trade Policy barriers versus institutional trade barriers: An application using "Good Old" OLS.

Notteboom, T. &Rodrigue, JP. (2008). Containerisation, Box Logistics and Global supply chains: The integration of ports and liner shipping networks. In: *Maritime Economics and Logistics.* No 10, pp. 152-174.

Jacobs, W. De Langen, P. Notteboom, T. (2011) *Institutional plasticity and path dependence in seaports: interactions between institutions, port governance reform and port authority routines. Proceedings of 2011 IAME conference*

Liu, X. Ye, H. Yuan, X. (2011). Tactical planning models for managing container flow and ship deployment. In: *Maritime Policy and Management.* Vol. 38, No. 5, pp. 487-508.

Luo, M. Fan, L. Liu, L. (2009). An Econometric analysis for container shipping market.In: *Maritime Policy & Management.* Vol. 36, No. 6, pp. 507-523.

Maatsch, S. (2010).A multinomial logit model of container fleet deployment.*Proceedings of the 20101 IAME Conference. Lisbon.*

Martinez Zarzoso, I. Marquez-Ramos, L.Wilmsmeier, G. (2011).*Repeat Pater Trade Facilitation, transport costs and maritime imports in Latin America. Proceedings of 2011 IAME conference*

Nazemzadeh, M. Meersman, H. Vanelslander, T. (2013).Ship operator's port selection behaviour in north European Ports..*Proceedings of the 2013 International Association of Maritime Economics (IAME) Conference*, Marseille, Marseille.

Oh, C. H., &Fratianni, M. U. (2010). Do Additional Bilateral Investment Treaties Boost Foreign Direct Investments?. *Available at SSRN 1607989*.

Panayides, P. Polyviou, M. (2010).The Effect of logistics attribues and services of ports on the economic and supply chain performance of shipping firms.*Proceedings of the 20101 IAME Conference. Lisbon.*

Rodriguez, F. Rodrik, D. (2001). Trade Policy and Economic Growth: A skeptic's Guide to the Cross-National Evidence. In: *NBER Macroeconomics Annual 2000 Volume 15.* Eds. Bernanke, B and Rogoff, K. MIT Press. pp. 261-338.

Santos Silva, J. Tenreyo, S. (2006). The log of Gravity. In: *Review of Economics and Statistics.* No. 88, Vol. 4, pp. 641-658.

Seoane, M. Gonzalez Laxe, F. Pais Montes, C. (2013). Emergent containership traffic in the south atlantic port system: A complex networks analysis (2007-2011). *Proceedings of the 2013 International Association of Maritime Economics (IAME) Conference*, Marseille, Marseille.

Sou, W. Ong, G. Chin, A. (2013).An Exploratory Study to Forecast Container Demand Using Economic Trade Models: A Case Study of the Asean Region. *Proceedings of the*

2013 International Association of Maritime Economics (IAME) Conference, Marseille, Marseille.

SYS, C.(2008). Is container liner shipping an oligopoly? In: *Proceedings of the International Forum on Shipping, Ports and Airports.* Vol. 4, No. 1, pp. 25-28.

The Economist. (2013). *The Humble Hero.*
http://www.economist.com/news/finance-and-economics/21578041-containers-have-been-more-important-globalisation-freer-trade-humble
Accessed on 19 August 2013

Tobar, B. Hernandez, R. Rodriguez, D. (2013). Port competitiveness and connectivity: the Canary Islands mainports case. *Proceedings of the 2013 International Association of Maritime Economics (IAME) Conference*, Marseille, Marseille.

UNCTAD (2012).Review of maritime Transport 2012).UNCTAD, Geneva.

UNESCAP (2008): *Gravity Models: Theoretical Foundations and related estimation issues.* Presentation made at ARTNetCapacity Building Workshop for Trade Research Phnom Penh, Cambodia 2-6 June 2008.

Ungo, R. Ducreux, A. (2013). A Competitive Analysis of Panama Canal routes: a shipper's perspective. *Proceedings of the 2013 International Association of Maritime Economics (IAME) Conference*, Marseille, Marseille.

Wang, Y., &Cullinane, K. (2008).Measuring container port accessibility: An application of the Principal Eigenvector Method (PEM). *Maritime Economics & Logistics*, *10*(1), 75-89.

Wei, D.Peijian, TIAN Guishi WU. (2008). "Analysis of Network Effect in Port and Shipping System Characterized by Scale-Free Network." *Chinese Journal of Management* Vol. 3. No 1, pp. 1-19

Wilmsmeier, G., Hoffmann, J., & Sanchez, R. J. (2006). The impact of port characteristics on international maritime transport costs. *Research in Transportation Economics*, *16*(1), 117-140.

Wilmsmeier, G., & Hoffmann, J. (2008). Liner shipping connectivity and port infrastructure as determinants of freight rates in the Caribbean. *Maritime Economics & Logistics*, *Vol. 10, No.1,* pp. 130-151

World Bank (2013).*Doing Business*
Www.doingbusiness.org
Accessed on 20 July 2013

World Bank (2013). Key performance indicators
www.worldbank.org/indicators
Accessed on 20 July 2013

WTO .(2013). Lamy: Aid for Trade aims to “connect the least connected”WTO speech

http://www.wto.org/english/news_e/sppl_e/sppl291_e.htm
Accessed on 14 August 2013

Yap, W. &Notteboom, T. (2011). Dynamics of liner shippi;ng service scheduling and their impact on container port competition. In: *Maritime Policiy and management.* Vol. 28, No. 5, pp. 471-485

Zondag, B. Bucci, P. Gützkow, P. Jong, G. de. (2010). Port competition modeling including maritime, port, and hinterland characteristics. In: *Maritime Policy & management. Vol. 37, No. 3, pp. 179-194*

Appendices

A List of countries

COUNTRY	ISO3	COUNTRY2	ISO33	COUNTRY4	ISO35
Albania	ALB	Ghana	GHA	Pakistan	PAK
Algeria	DZA	Greece	GRC	Palau	PLW
Angola	AGO	Greenland	GRL	Panama	PAN
Antigua & Barbuda	ATG	Grenada	GRD	Papua New Guinea	PNG
Argentina	ARG	Guam	GUM	Peru	PER
Aruba	ABW	Guatemala	GTM	Philippines	PHL
Australia	AUS	Guinea	GIN	Poland	POL
Bahamas	BHS	Guinea-Bissau	GNB	Portugal	PRT
Bahrain	BHR	Guyana	GUY	Puerto Rico	PRI
Bangladesh	BGD	Haiti	HTI	Qatar	QAT
Barbados	BRB	Honduras	HND	Republic of Ireland	IRL
Belgium	BEL	Hong Kong	HKG	Romania	ROU
Belize	BLZ	Iceland	ISL	Russia	RUS
Benin	BEN	India	IND	Samoa	WSM
Bermuda	BMU	Indonesia	IDN	Samoa, American	ASM
Brazil	BRA	Iran	IRN	Sao Tome & Principe	STP
Brunei	BRN	Iraq	IRQ	Saudi Arabia	SAU
Bulgaria	BGR	Israel	ISR	Senegal	SEN
Cambodia	KHM	Italy	ITA	Seychelles	SYC
Cameroon	CMR	Ivory Coast	CIV	Sierra Leone	SLE
Canada	CAN	Jamaica	JAM	Singapore	SGP
Cape Verde	CPV	Japan	JPN	Slovenia	SVN
Cayman Islands	CYM	Jordan	JOR	Solomon Islands	SLB
Chile	CHL	Kenya	KEN	Somali Republic	SOM
China	CHN	Kiribati	KIR	South Africa	ZAF
Colombia	COL	Kuwait	KWT	South Korea	KOR
Comoros	COM	Latvia	LVA	Spain	ESP
Congo	COD	Lebanon	LBN	Sri Lanka	LKA
Congo, The	COG	Liberia	LBR	St. Kitts-Nevis	KNA
Costa Rica	CRI	Libya	LBY	St. Lucia	LCA
Croatia	HRV	Lithuania	LTU	St. Vincent & Grenadines	VCT
Cuba	CUB	Madagascar	MDG	Sudan	SDN
Curacao	CUW	Malaysia	MYS	Suriname	SUR
Cyprus	CYP	Maldives	MDV	Sweden	SWE
Denmark	DNK	Malta	MLT	Syria	SYR
Djibouti	DJI	Marshall Islands	MHL	Taiwan	TWN
Dominica	DMA	Mauritania	MRT	Tanzania	TZA
Dominican Republic	DOM	Mauritius	MUS	Thailand	THA
Ecuador	ECU	Mexico	MEX	Togo	TGO
Egypt	EGY	Micronesia	FSM	Tonga	TON
El Salvador	SLV	Montenegro	MNE	Trinidad & Tobago	TTO
Equatorial Guinea	GNQ	Morocco	MAR	Tunisia	TUN
Eritrea	ERI	Mozambique	MOZ	Turkey	TUR
Estonia	EST	Myanmar	MMR	U.K.	GBR
Faroe Islands	FRO	Namibia	NAM	U.S.A.	USA
Fiji	FJI	Netherlands	NLD	Ukraine	UKR
Finland	FIN	New Caledonia	NCL	United Arab Emirates	ARE
France	FRA	New Zealand	NZL	Uruguay	URY
French Polynesia	PYF	Nicaragua	NIC	Vanuatu	VUT
Gabon	GAB	Nigeria	NGA	Venezuela	VEN
Gambia	GMB	Northern Marianas	MNP	Vietnam	VNM
Georgia	GEO	Norway	NOR	Virgin Islands, American	VIR
Germany	DEU	Oman	OMN	Yemen	YEM

B Minimum number of transshipments required to reach all other countries.

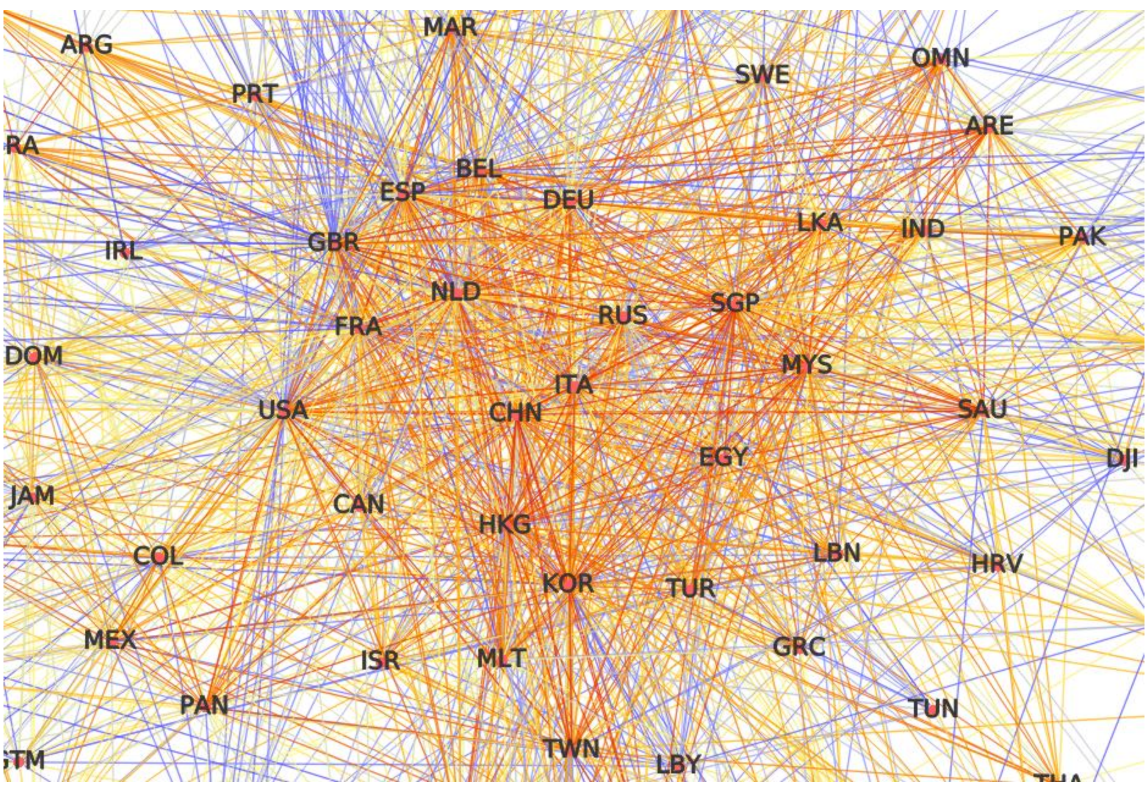
Nations	Tzero	Tone	Ttwo	Tthree	Nations 2	Tzero 3	Tone 4	Ttwo 5	Tthree 6
ALB	6	82	70	0	LBN	32	122	4	0
DZA	20	114	24	0	LBR	19	113	26	0
AGO	29	122	7	0	LBY	24	110	24	0
ATG	14	109	35	0	LTU	10	101	47	0
ARG	31	120	7	0	MDG	12	76	70	0
ABW	5	94	58	1	MYS	75	83	0	0
AUS	37	120	1	0	MDV	7	77	74	0
BHS	24	124	10	0	MLT	45	113	0	0
BHR	11	109	38	0	MHL	13	78	67	0
BGD	4	80	74	0	MRT	17	98	43	0
BRB	17	106	35	0	MUS	19	117	22	0
BEL	87	71	0	0	MEX	49	96	13	0
BLZ	15	105	38	0	FSM	6	79	73	0
BEN	26	130	2	0	MNE	12	112	34	0
BMU	7	118	33	0	MAR	55	103	0	0
BRA	33	118	7	0	MOZ	18	115	25	0
BRN	3	78	77	0	MMR	2	79	77	0
BGR	10	81	67	0	NAM	23	130	5	0
KHM	8	95	55	0	NLD	87	71	0	0
CMR	28	121	9	0	NCL	28	118	12	0
CAN	45	112	1	0	NZL	28	118	12	0

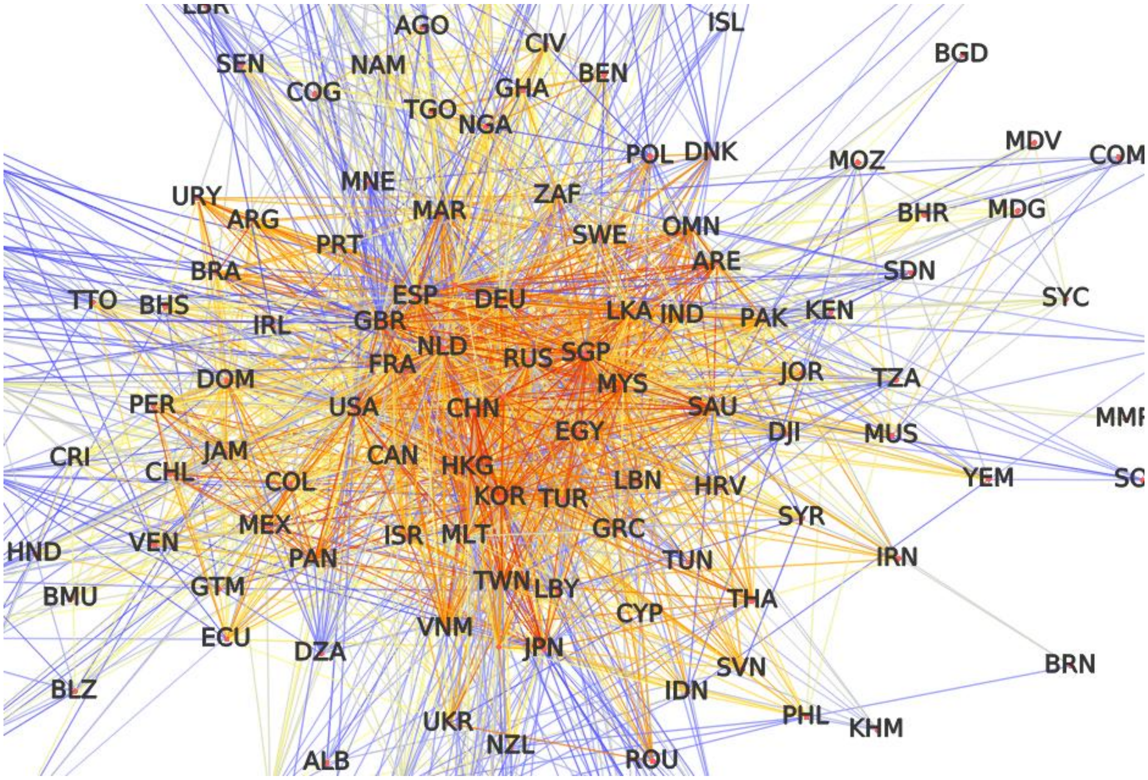
CPV	9	102	47	0	NIC	10	90	57	1
CYM	4	95	58	1	NGA	35	121	2	0
CHL	26	119	13	0	MNP	5	78	75	0
CHN	87	71	0	0	NOR	12	114	32	0
COL	47	98	13	0	OMN	37	120	1	0
COM	7	62	88	1	PAK	39	118	1	0
COD	10	100	48	0	PLW	6	79	73	0
COG	29	115	14	0	PAN	50	95	13	0
CRI	23	108	27	0	PNG	13	91	54	0
HRV	27	127	4	0	PER	27	118	13	0
CUB	16	106	36	0	PHL	18	93	47	0
CUW	6	94	57	1	POL	29	118	11	0
CYP	21	103	34	0	PRT	48	98	12	0
DNK	29	116	13	0	PRI	18	110	30	0
DJI	33	123	2	0	QAT	2	54	101	1
DMA	8	101	49	0	IRL	22	107	29	0
DOM	43	102	13	0	ROU	16	107	35	0
ECU	27	118	13	0	RUS	49	104	5	0
EGY	58	100	0	0	WSM	14	106	38	0
SLV	11	89	57	1	ASM	14	106	38	0
GNQ	14	125	19	0	STP	8	100	50	0
ERI	6	58	93	1	SAU	46	111	1	0
EST	10	105	43	0	SEN	27	106	25	0
FRO	6	100	52	0	SYC	11	76	71	0
FJI	27	119	12	0	SLE	6	83	69	0

FIN	14	105	39	0	SGP	77	81	0	0
FRA	93	64	1	0	SVN	20	121	17	0
PYF	23	115	20	0	SLB	18	86	54	0
GAB	25	116	17	0	SOM	7	60	90	1
GMB	15	99	44	0	ZAF	48	108	2	0
GEO	6	62	90	0	KOR	83	75	0	0
DEU	83	75	0	0	ESP	87	70	1	0
GHA	29	127	2	0	LKA	50	107	1	0
GRC	40	118	0	0	KNA	12	87	58	1
GRL	1	28	116	13	LCA	15	108	35	0
GRD	13	110	35	0	VCT	13	110	35	0
GUM	8	115	35	0	SDN	16	111	31	0
GTM	34	111	13	0	SUR	17	104	37	0
GIN	24	96	38	0	SWE	40	118	0	0
GNB	9	102	47	0	SYR	17	106	35	0
GUY	16	105	37	0	TWN	61	96	1	0
HTI	20	105	33	0	TZA	24	112	22	0
HND	21	105	32	0	THA	27	126	5	0
HKG	82	76	0	0	TGO	29	127	2	0
ISL	9	111	38	0	TON	14	77	67	0
IND	53	104	1	0	TTO	31	113	14	0
IDN	23	127	8	0	TUN	14	108	36	0
IRN	15	109	34	0	TUR	48	110	0	0
IRQ	3	56	98	1	GBR	93	65	0	0
ISR	38	116	4	0	USA	98	59	1	0

ITA	82	76	0	0	UKR	21	116	21	0
CIV	31	122	5	0	ARE	56	101	1	0
JAM	50	95	13	0	URY	25	126	7	0
JPN	56	98	4	0	VUT	16	75	67	0
JOR	27	128	3	0	VEN	30	115	13	0
KEN	28	116	14	0	VNM	29	125	4	0
KIR	13	78	67	0	VIR	6	93	58	1
KWT	3	59	95	1	YEM	15	93	50	0
LVA	5	85	68	0					

C Liner shipping network structure



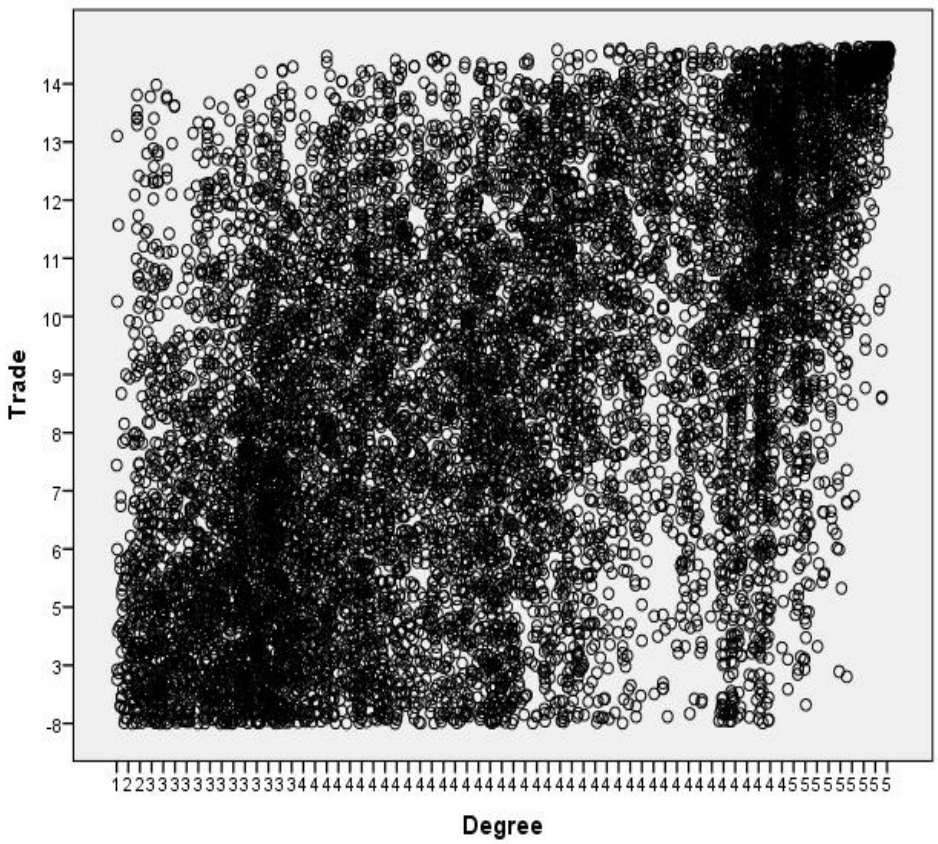

AGO
ISL
CIV
BGD
SEN
NAM
BEN
COG
GHA
TGO
NGA
MDV
POL
DNK
MOZ
COM
MNE
URY
ZAF
BHR
MDG
MAR
ARG
OMN
SWE
PRT
ARE
BRA
SDN
TTO
BHS
ESP
DEU
SYC
LKA
IND
PAK
KEN
IRL
GBR
NLD
SGP
RUS
FRA
JOR
TZA
DOM
MYS
PER
USA
CHN
SAU
EGY
DJI
MUS
JAM
CAN
CRI
CHL
HKG
LBN
YEM
COL
HRV
KOR
TUR
SYR
MEX
HND
VEN
ISR
MLT
GRC
IRN
PAN
TUN
TWN
LBY
GTM
THA
BMU
CYP
ECU
VNM
DZA
JPN
SVN
BRN
IDN
BLZ
PHL
KHM
UKR
NZL
ALB
ROU

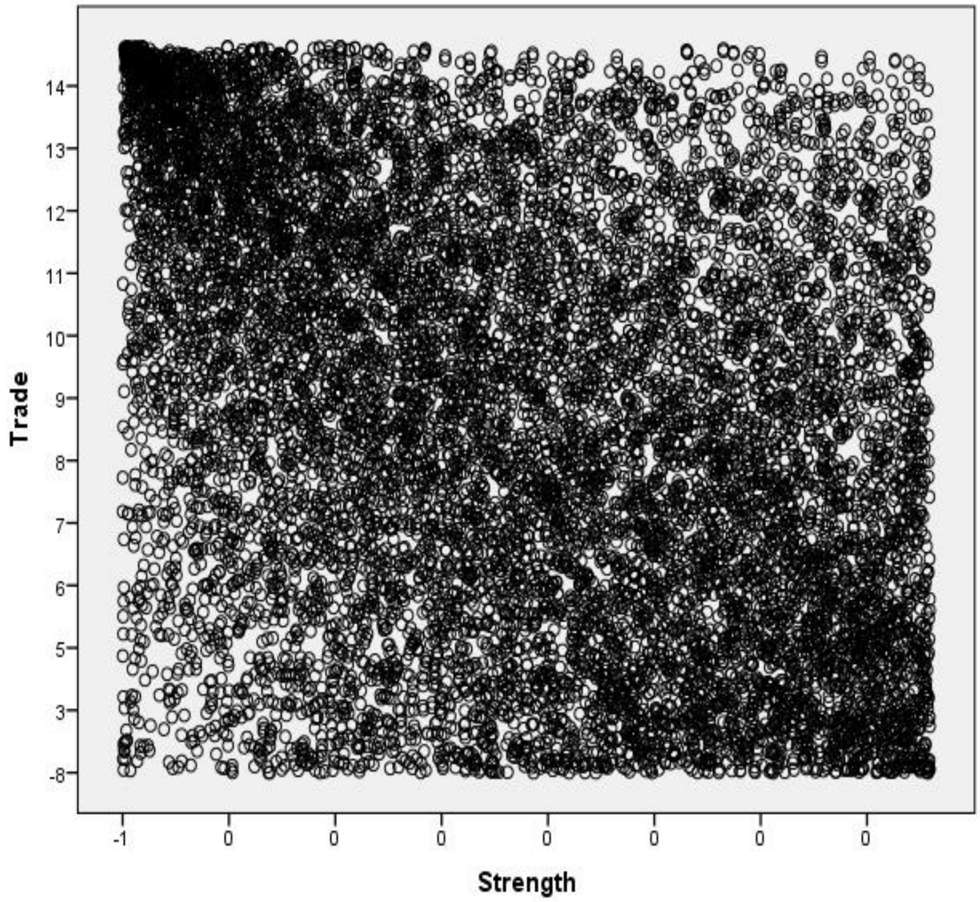
D Network structure determinants

Country	Degree	Betweenness centrality	Cluster	Strenght	K-Core	Strahler
ABW	5	0	1	0.54144	5	71
AGO	29	22.3093	0.724138	0.511892	22	84
ALB	6	0.589476	0.73333	0.409016	6	66
ARE	56	520.275	0.465584	0.535223	23	1
ARG	31	10.1795	0.772043	0.640799	23	84
ASM	14	4.05251	0.945055	0.554456	13	19
ATG	14	0.18458	0.813187	0.479198	11	68
AUS	37	46.3777	0.657658	0.377412	23	77
BEL	87	491.815	0.361935	0.58174	23	83
BEN	26	5.29197	0.870769	0.578358	22	76
BGD	4	0	1	0.578358	4	42
BGR	10	1.28947	0.888889	0.579878	9	68
BHR	11	9.59791	0.854545	0.568999	10	76
BHS	24	9.47156	0.757246	0.538365	22	83
BLZ	15	10.1222	0.704762	0.517379	12	68
BMU	7	0	1	0.541505	7	76
BRA	33	17.5724	0.727273	0.515315	23	77
BRB	17	10.6534	0.691176	0.51338	11	68
BRN	3	0	1	0.470539	3	42
CAN	45	70.0037	0.59798	0.457042	23	76
CHL	26	3.78803	0.873846	0.559627	23	76
CHN	87	450.752	0.369153	0.559627	23	76
CIV	31	12.9545	0.767742	0.554427	22	42
CMR	28	13.5035	0.767196	0.560987	22	67
COD	10	0.697447	0.84444	0.552502	10	67
COG	29	15.4482	0.731527	0.548148	22	66
COL	47	99.074	0.556892	0.452716	23	71
COM	7	0.129167	0.904762	0.56454	7	42
CPV	9	0.0714286	0.972222	0.564385	9	66
CRI	23	15.2389	0.715415	0.52354	16	68
CUB	16	5.20678	0.75	0.53534	13	67
CUW	6	0.217857	0.93333	0.579405	5	33
CYM	4	0	1	0.496049	4	34
CYP	21	9.03423	0.780952	0.561719	18	68
DEU	83	458.656	0.382016	0.395594	23	60
DJI	33	72.7654	0.681818	0.507617	23	67
DMA	8	0	1	0.583563	8	66
DNK	29	191.363	0.608374	0.457709	20	66
DOM	43	59.9065	0.576965	0.458716	23	67
DZA	20	1.48908	0.915789	0.536609	19	83
ECU	27	15.4493	0.746469	0.522207	22	66
EGY	58	153.093	0.746439	0.451334	23	66
ERI	6	0	1	0.575623	6	51
ESP	87	524.808	0.361401	0.384019	23	16
EST	10	3.73312	0.82222	0.540387	9	61
FIN	14	6.84776	0.758242	0.531086	10	60
FJI	27	36.0329	0.660969	0.499009	18	66
FRA	93	709.703	0.335671	0.354633	23	66
FRO	6	0	1	0.573247	6	60
FSM	6	0.27833	0.93333	0.591841	5	20
GAB	25	16.9057	0.71333	0.548771	19	50
GBR	93	776.076	0.321879	0.347782	23	5
GEO	6	0	1	0.581396	6	34
GHA	29	10.6858	0.82265	0.561515	22	57
GIN	24	16.1877	0.673913	0.537034	18	49
GMB	15	6.78797	0.704762	0.517066	11	49
GNB	9	0.0714286	0.972222	0.564385	9	34
GNQ	14	0.350623	0.835165	0.5715	13	50
GRC	40	49.6964	0.646154	0.503896	23	57
GRD	13	1.89673	0.871795	0.570235	11	43
GRL	1	0	0.000001	0.00001	1	1
GTM	34	37.4043	0.638146	0.49614	22	57
GUM	8	8.97384	0.75	0.520025	5	43
GUY	16	5.68985	0.75	0.51995	11	43
HKG	82	461.753	0.384824	0.332837	23	57
HND	21	17.4299	0.66667	0.5163358	14	34
HRV	27	6.58176	0.843305	0.560494	23	68
HTI	20	19.328	0.594737	0.459399	14	43
IDN	23	12.1586	0.73913	0.525887	18	49
IND	53	188.558	0.517417	0.452064	23	51
IRL	22	10.0036	0.718615	0.510095	18	16
IRN	15	0.27489	0.952381	0.620728	15	49
IRQ	3	0.633074	0.66667	0.397119	2	23
ISL	9	1.74111	0.861111	0.553789	8	32
ISR	38	61.0893	0.651494	0.491836	23	49
ITA	82	411.666	0.393857	0.404138	23	49
JAM	50	158.568	0.46449	0.402338	23	33
JOR	27	13.5715	0.769231	0.558619	21	43
JPN	56	292.919	0.428571	0.388325	23	43

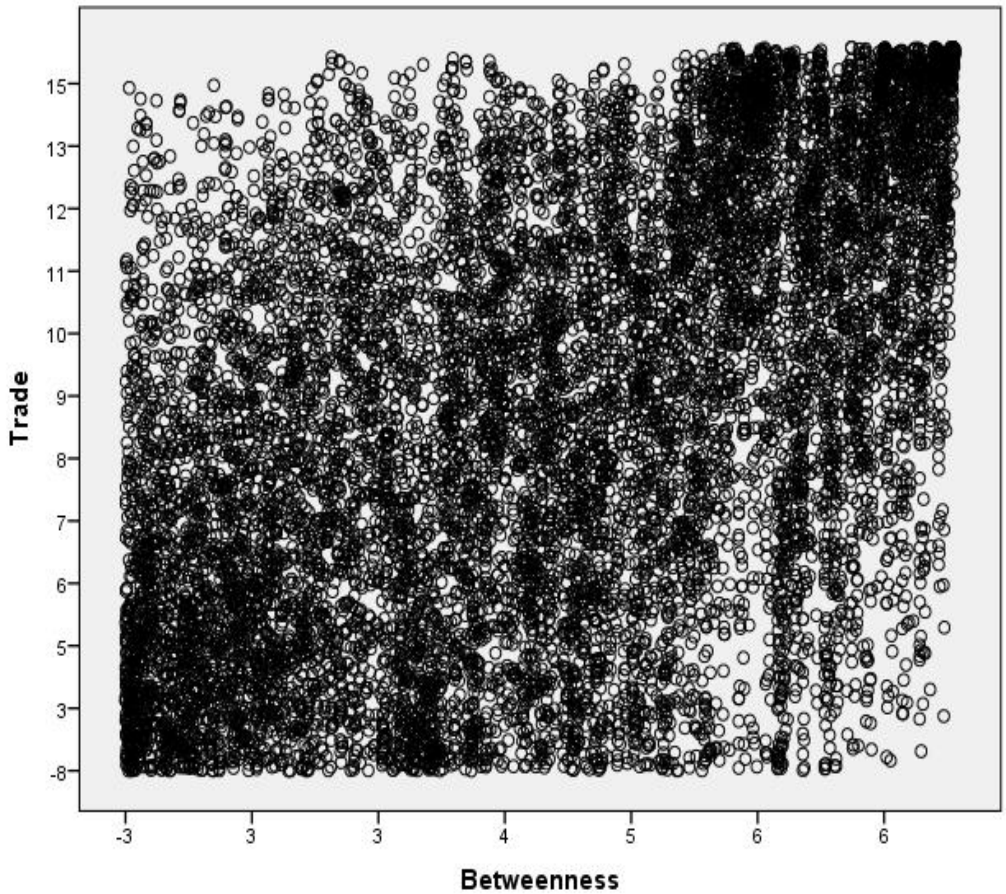
Column1	Degree	Betweenness centrality	Cluster	Strenght	K-Core	Strahler
KEN	28	24.8055	0.73545	0.528664	21	42
KHM	8	0	1	0.603562	8	76
KIR	13	0	1	0.670519	13	27
KNA	12	2.43059	0.772727	0.532305	11	9
KOR	83	687.026	0.363503	0.362268	23	18
KWT	3	0	1	0.454545	3	21
LBN	32	14.1459	0.752016	0.543834	23	42
LBR	19	17.0392	0.654971	0.501051	14	33
LBY	24	4.46388	0.818841	0.567938	21	34
LCA	15	5.22151	0.771429	0.540576	11	9
LKA	50	116.941	0.560816	0.474404	23	12
LTU	10	2.87942	0.84444	0.546208	9	32
LVA	5	0	1	0.537483	5	32
MAR	55	153.735	0.514478	0.446598	23	32
MDG	12	2.3172	0.8181812	0.576366	10	42
MDV	7	0	1	0.586751	7	21
MEX	49	119.02	0.528061	0.451487	23	33
MHL	13	0	1	0.670519	13	26
MLT	45	121.406	0.581717	0.477129	23	34
MMR	2	0	1	0.469883	2	20
MNE	12	5.57566	0.818182	0.565107	11	32
MNP	5	0	1	0.600145	5	19
MOZ	18	11.4283	0.797386	0.567236	15	23
MRT	17	8.56553	0.720588	0.542858	13	32
MUS	19	17.0713	0.754386	0.541996	15	24
MYS	75	491.651	0.389189	0.401188	23	42
NAM	23	9.32728	0.794466	0.559301	21	32
NCL	28	39.9937	0.642857	0.498501	18	26
NGA	35	33.2979	0.690756	0.523781	22	22
NIC	10	0.245833	0.93333	0.593209	10	21
NLD	87	549.613	0.350441	0.374504	23	32
NOR	12	3.2981	0.78789	0.54806	10	16
NZL	28	28.2251	0.690476	0.499849	19	23
OMN	37	95.0893	0.696697	0.504097	23	23
PAK	39	149.503	0.591093	0.482605	23	21
PAN	50	111.8188	0.518367	0.444611	23	21
PER	27	12.3521	0.823362	0.544004	23	21
PHI	18	14.8152	0.718954	0.515682	16	21
PLW	6	0.270833	0.93333	0.591814	5	18
PNG	13	1.35911	0.910256	0.571291	13	20
POL	29	21.9681	0.672414	0.519101	22	21
PRI	18	17.5524	0.647059	0.468473	14	16
PRT	48	107.712	0.478723	0.445235	23	20
PYF	23	27.4889	0.731225	0.507344	16	57
QAT	2	0	1	0.342508	2	1
ROU	16	11.711	0.816667	0.569878	14	21
RUS	49	203	0.528061	0.42572	23	220
SAU	46	140.151	0.595169	0.482833	23	21
SDN	16	0.974587	0.93333	0.598516	15	2
SEN	27	12.6415	0.706553	0.545793	21	18
SGP	77	537.822	0.37013	0.393077	23	20
SLB	18	6.44086	0.79085	0.579933	13	18
SLE	6	0	1	0.533729	6	16
SLV	11	1.13083	0.8	0.571076	10	57
SOM	7	0.206407	0.952381	0.593873	6	2
STP	8	0.212234	0.892857	0.579072	8	5
SUR	17	6.95406	0.720588	0.506617	11	7
SVN	20	13.0169	0.784211	0.558705	18	18
SWE	40	64.8824	0.620513	0.49556	23	8
SYC	11	1.4044	0.836364	0.579752	10	20
SYR	17	1.56977	0.882353	0.606129	17	6
TGO	29	6.93345	0.852217	0.576169	22	6
THA	27	15.181	0.737892	0.529709	20	5
TON	14	0.381818	0.956044	0.660794	13	1
TTO	31	71.698	0.584946	0.431553	20	6
TUN	14	0.17549	0.967033	0.630814	14	6
TUR	48	121.933	0.574468	0.473833	23	6
TWN	61	320.432	0.435519	0.392578	23	10
TZA	24	33.2857	0.699275	0.519385	18	1
UKR	21	20.3766	0.719048	0.537704	18	1
URY	25	4.30664	0.856667	0.562166	22	1
USA	98	1310.54	0.29939	0.316766	23	5
VCT	13	6.26447	0.782051	0.562225	11	8
VEN	30	41.4132	0.65977	0.484818	21	1
VIR	6	0.283333	0.86667	0.493593	6	1
VNM	29	16.7034	0.770936	0.537626	22	1
VUT	16	1.79491	0.866667	0.6248427	13	1
WSM	14	4.05251	0.945055	0.640799	13	20
YEM	15	9.73703	0.752381	0.564364	13	1
ZAF	48	108.56	0.557624	0.464367	23	18

E Network structure relations to trade

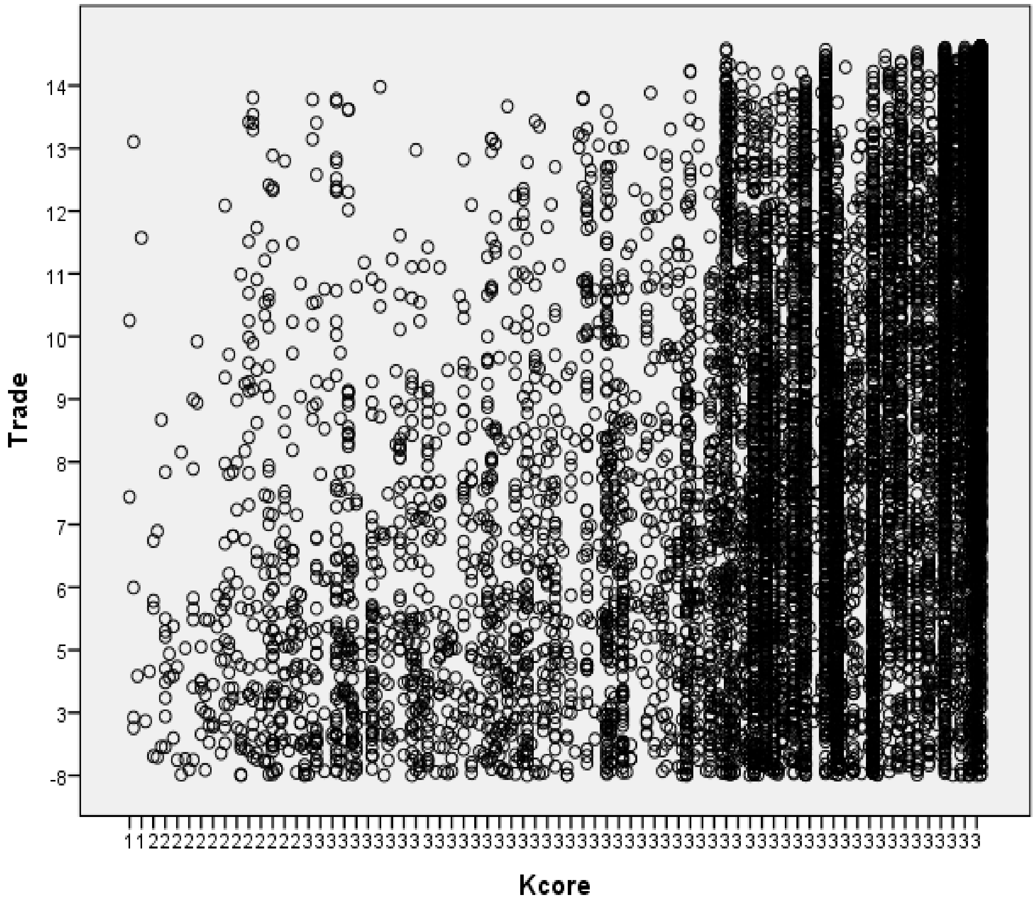


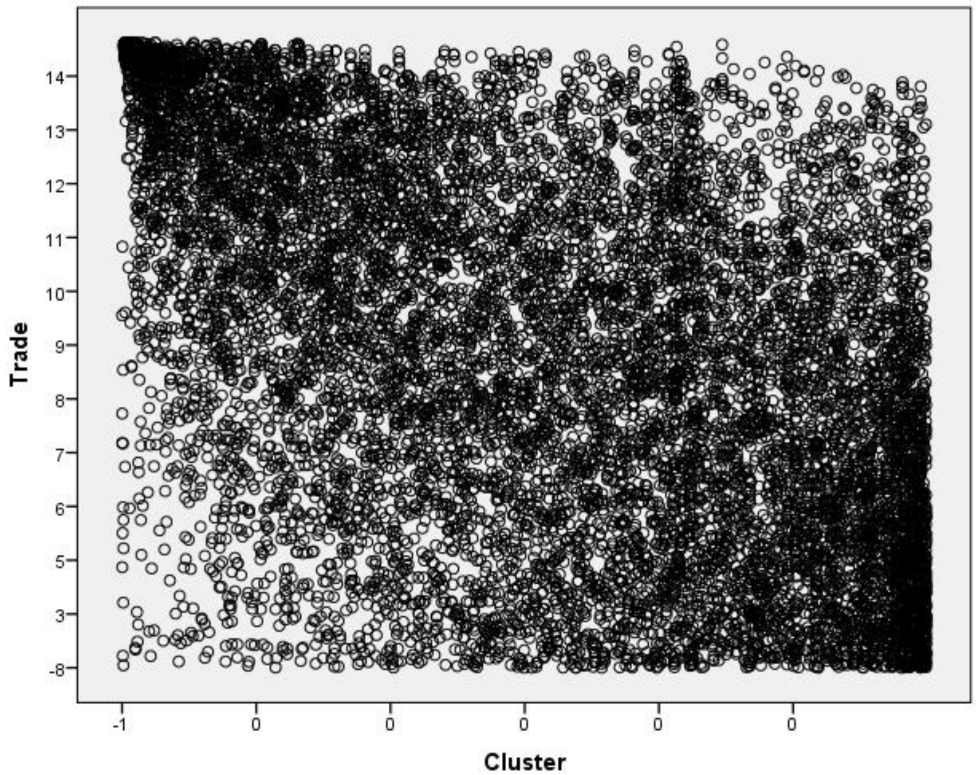


Trade
14
13
12
11
10
9
8
7
6
5
3
-8
-1
0
0
0
0
0
0
0
Strength



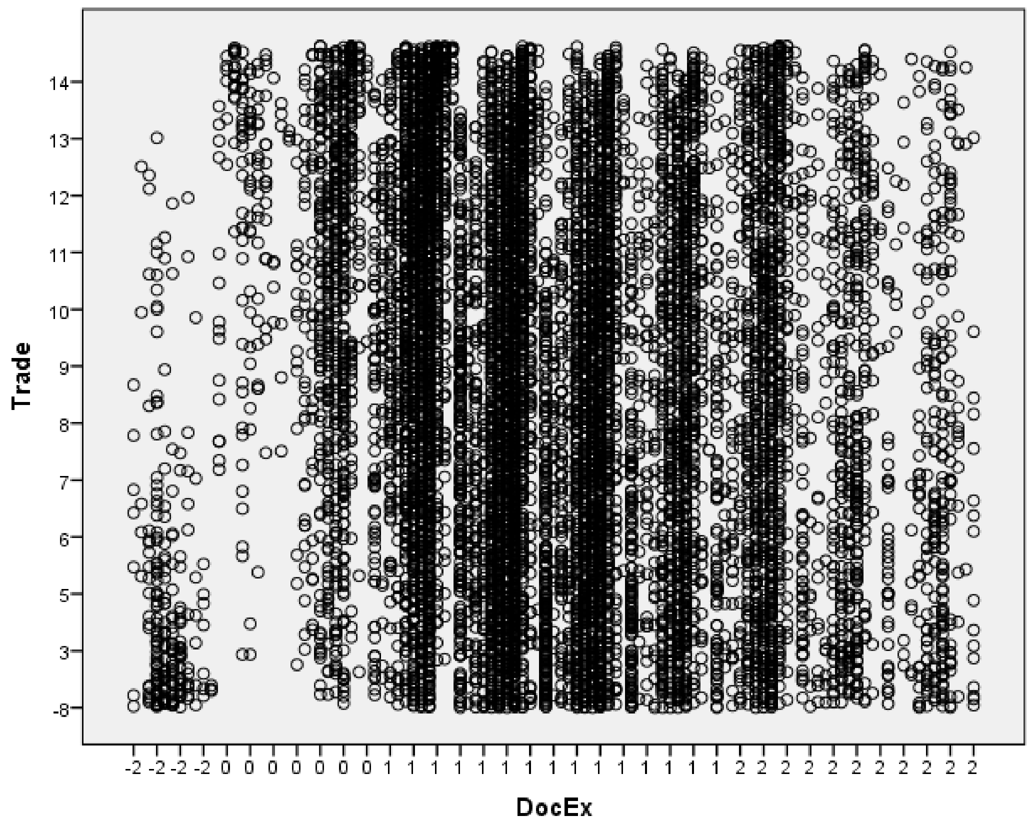
Trade
15
13
12
11
10
9
8
7
6
5
3
-8
-3
3
3
4
5
6
6
Betweenness

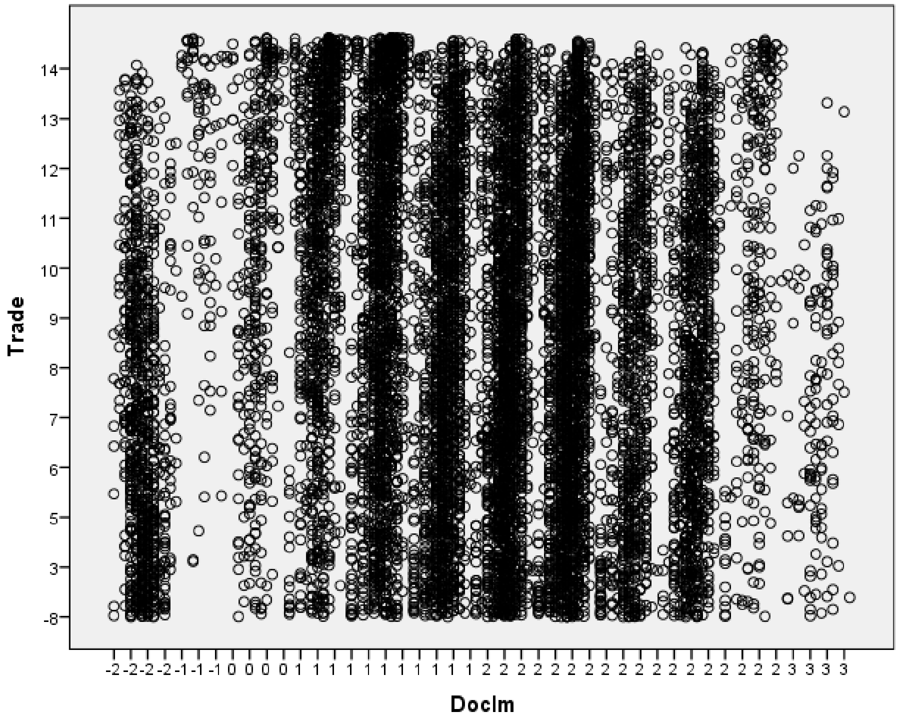

Trade
14
13
12
11
10
9
8
7
6
5
3
-8
Kcore

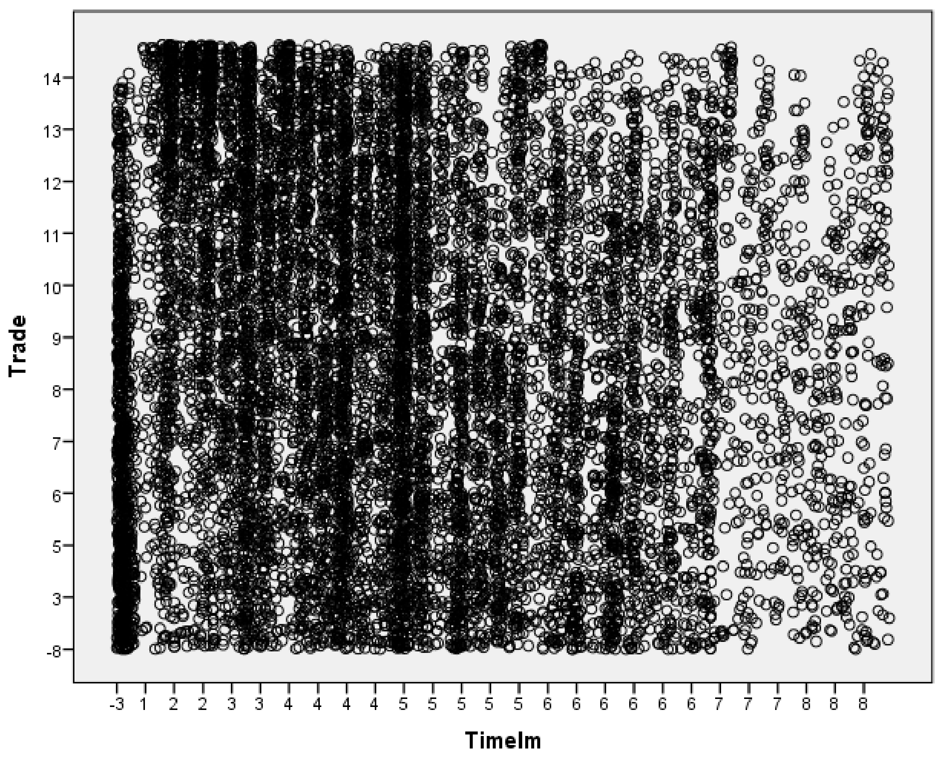


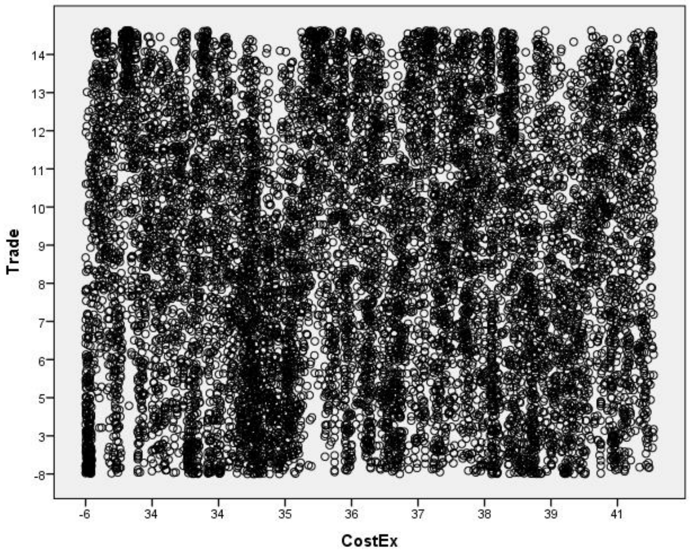
Trade
14
13
12
11
10
9
8
7
6
5
3
-8
-1
0
0
0
0
0
Cluster

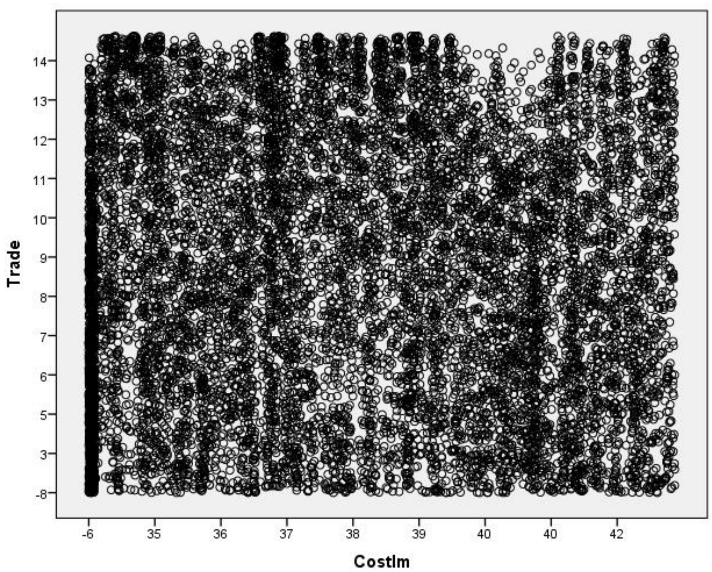
F Policy relations to trade

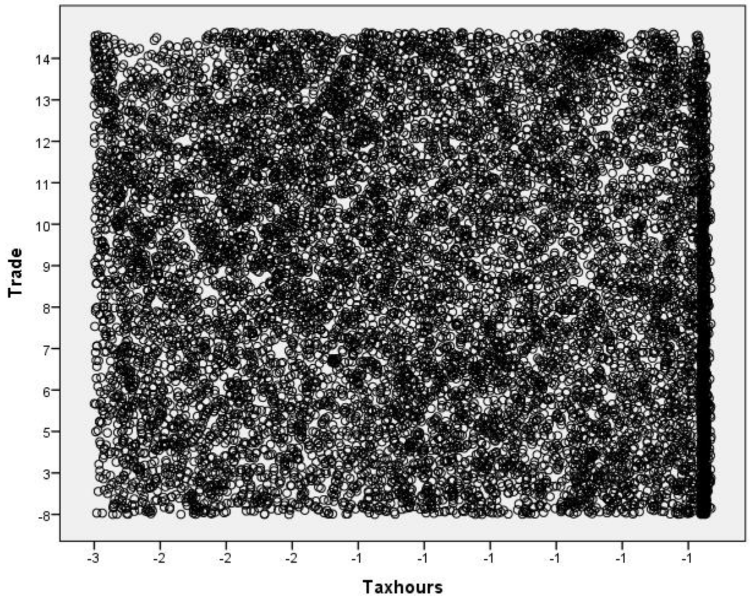


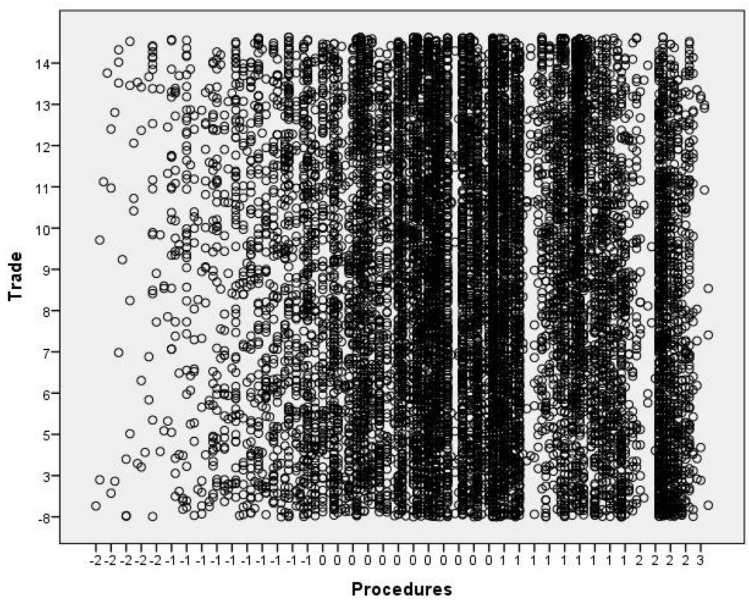

Trade
DocIm

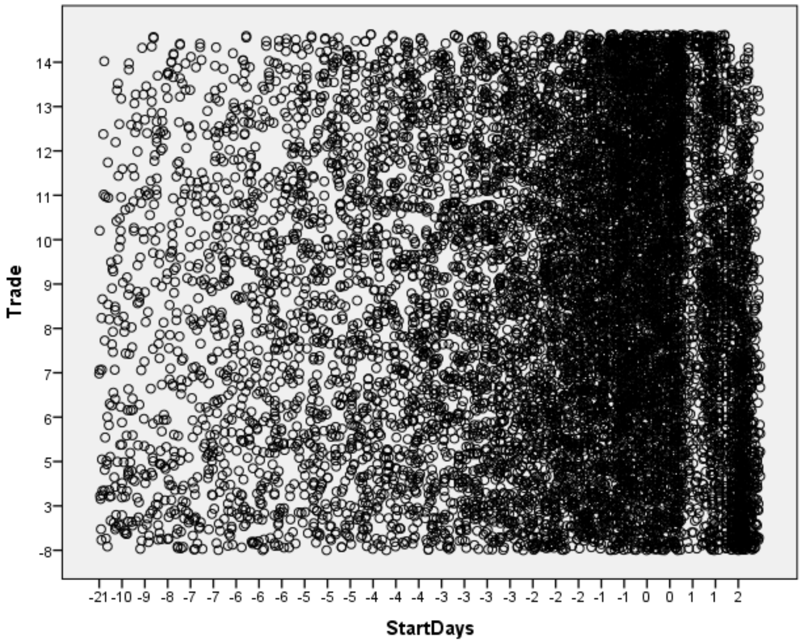

Trade
TimeIm

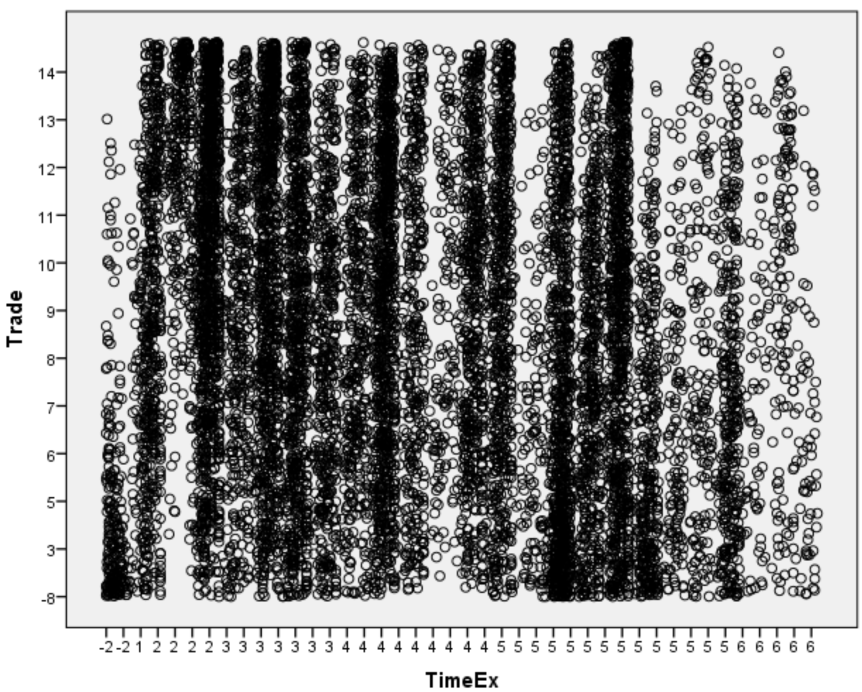

Trade
14
13
12
11
10
9
8
7
6
5
3
-8
-6
34
34
35
36
37
38
39
41
CostEx


Trade
14
13
12
11
10
9
8
7
6
5
3
-8
-6
35
36
37
38
39
40
40
42
CostIm

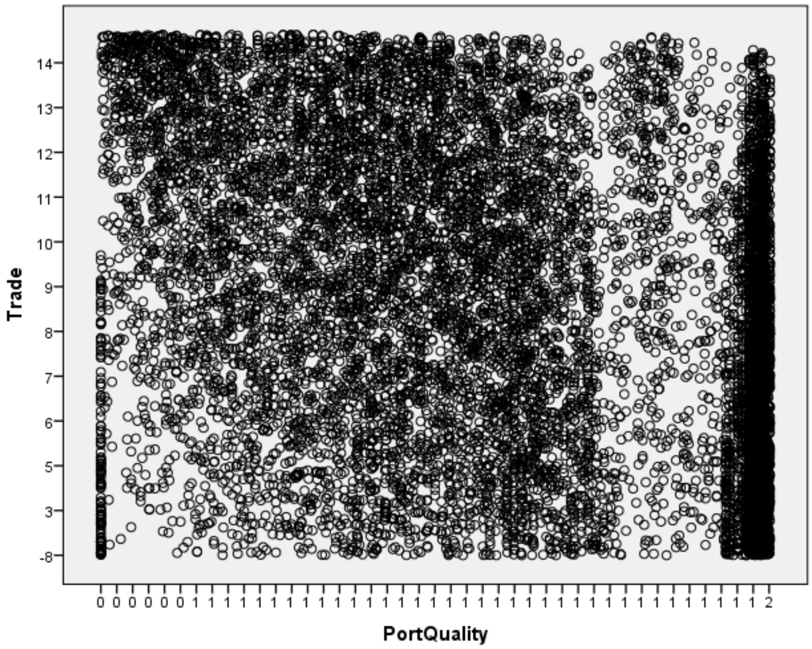

Trade
Taxhours

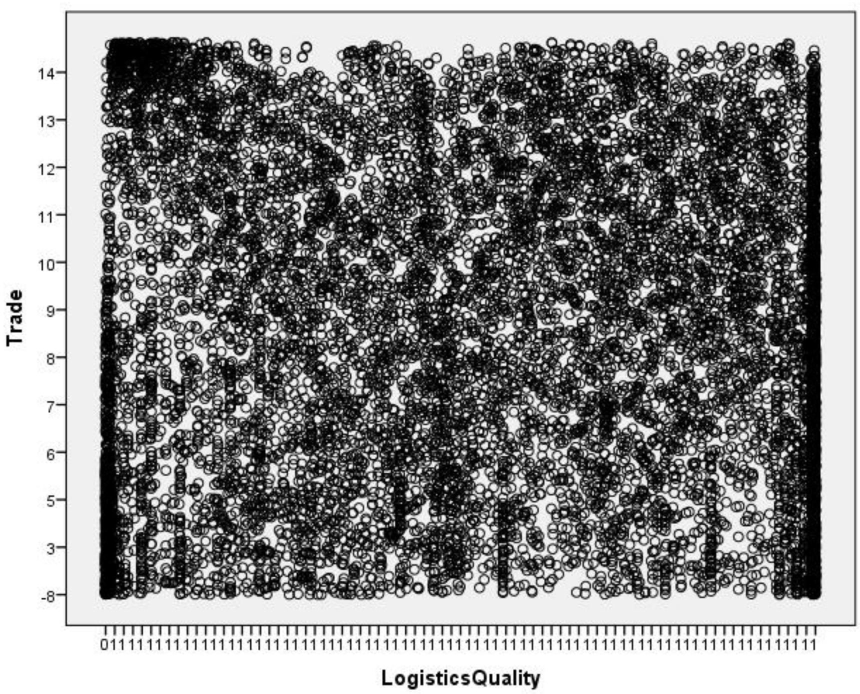

Trade
Procedures

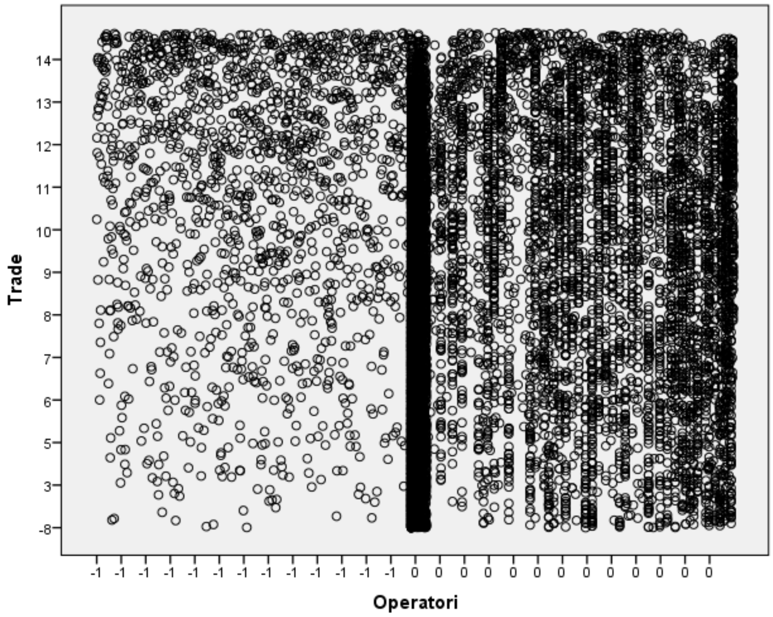

Trade
StartDays


Trade
TimeEx

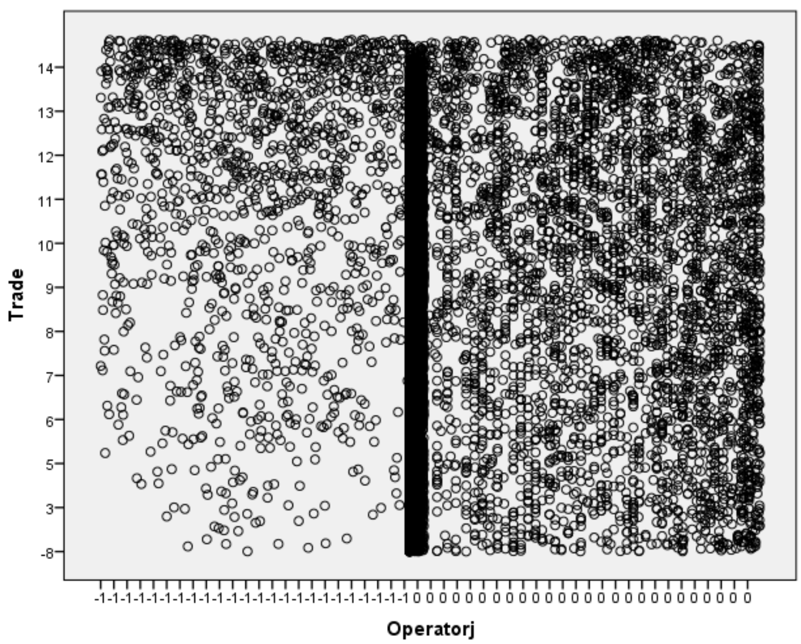
G Port and infrastructure structure relations to trade





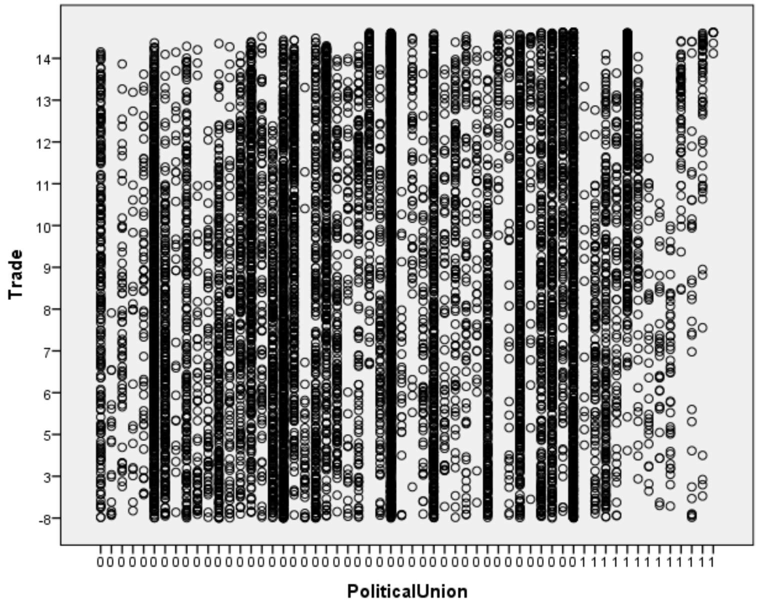


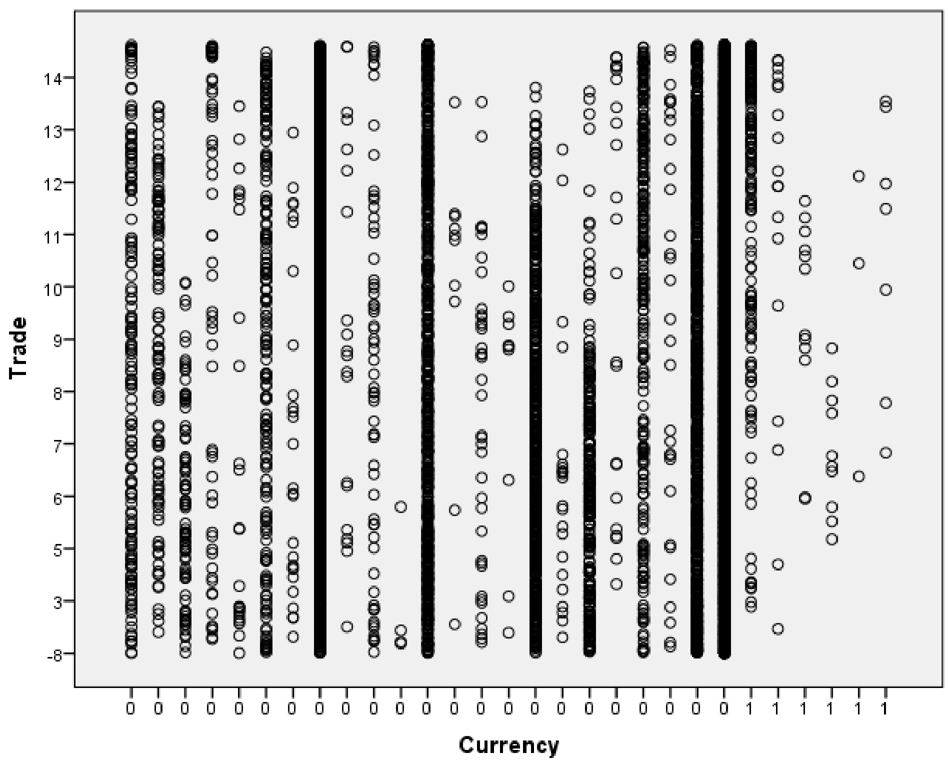
Trade
Operatori

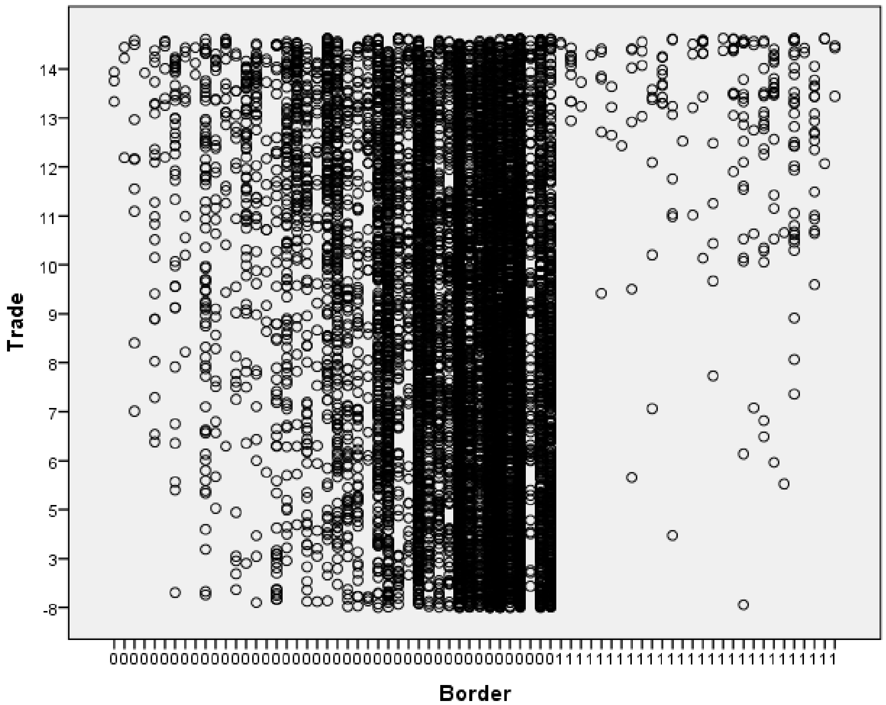


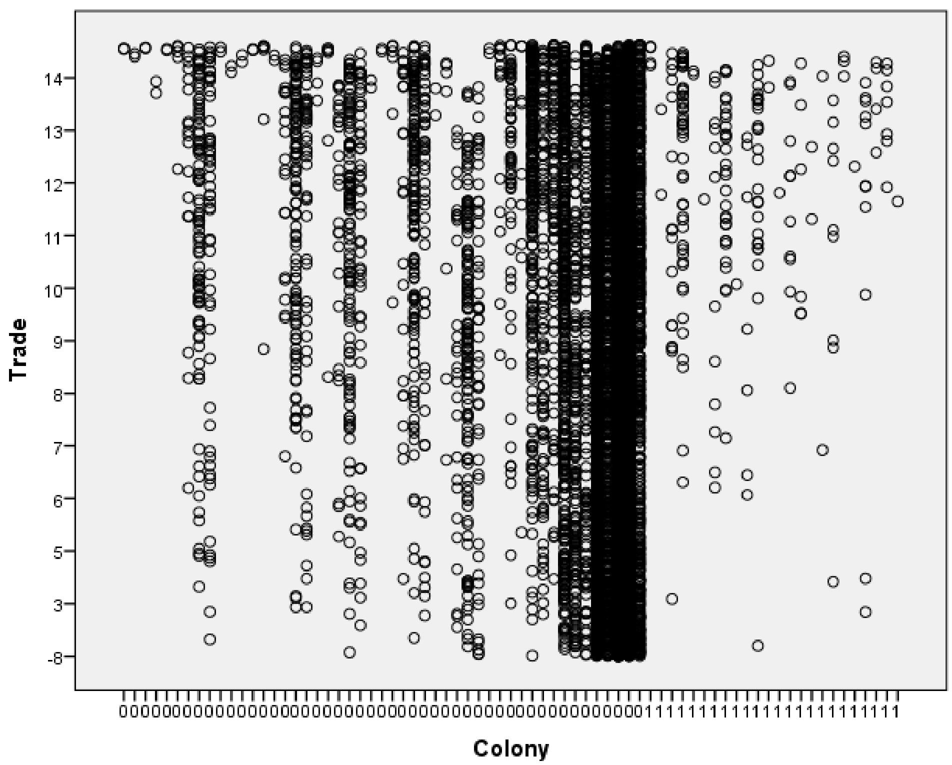
Trade
Operatorj

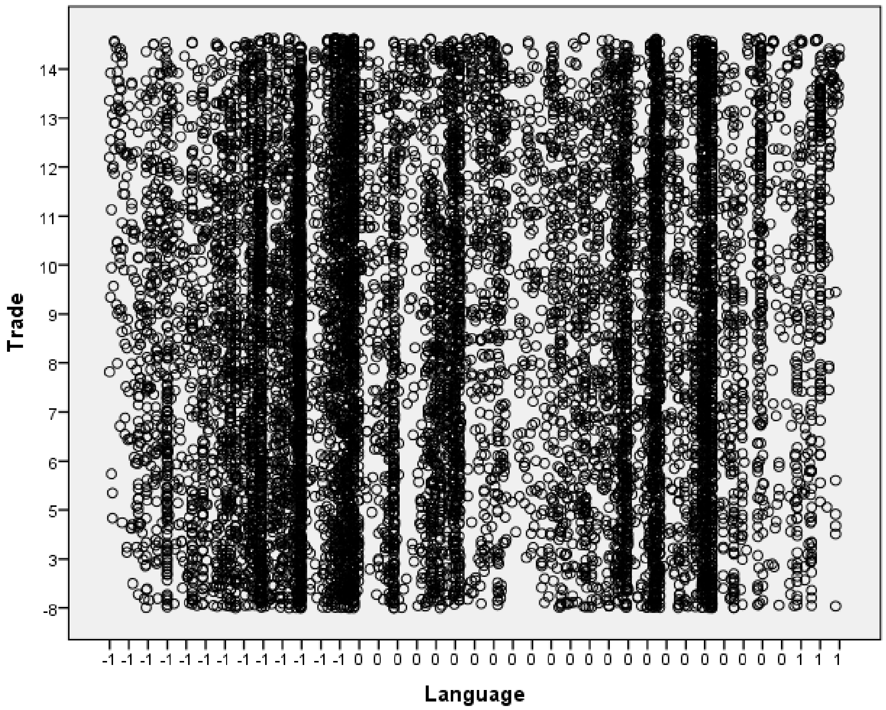
H Institutions and Settings relations to trade



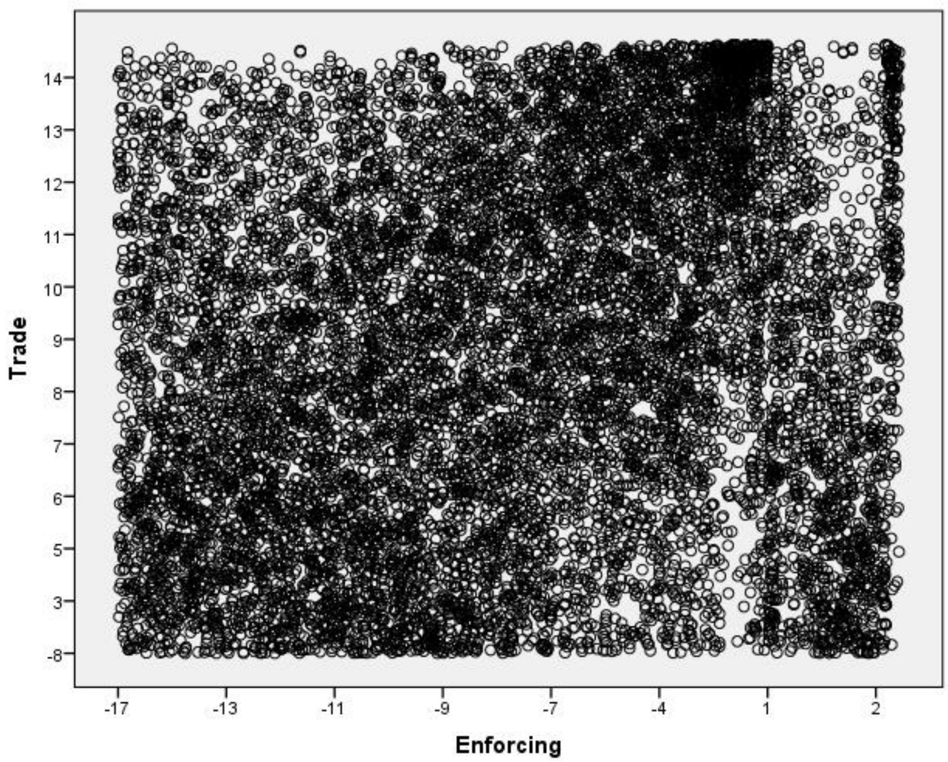



Trade
Border

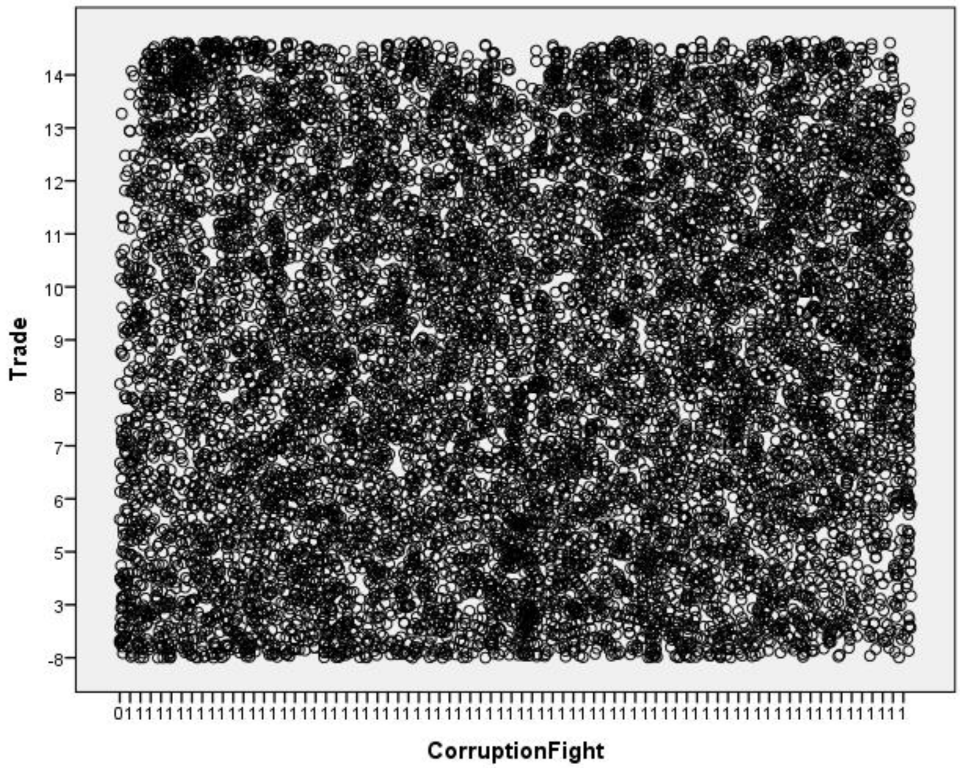

Trade
Colony

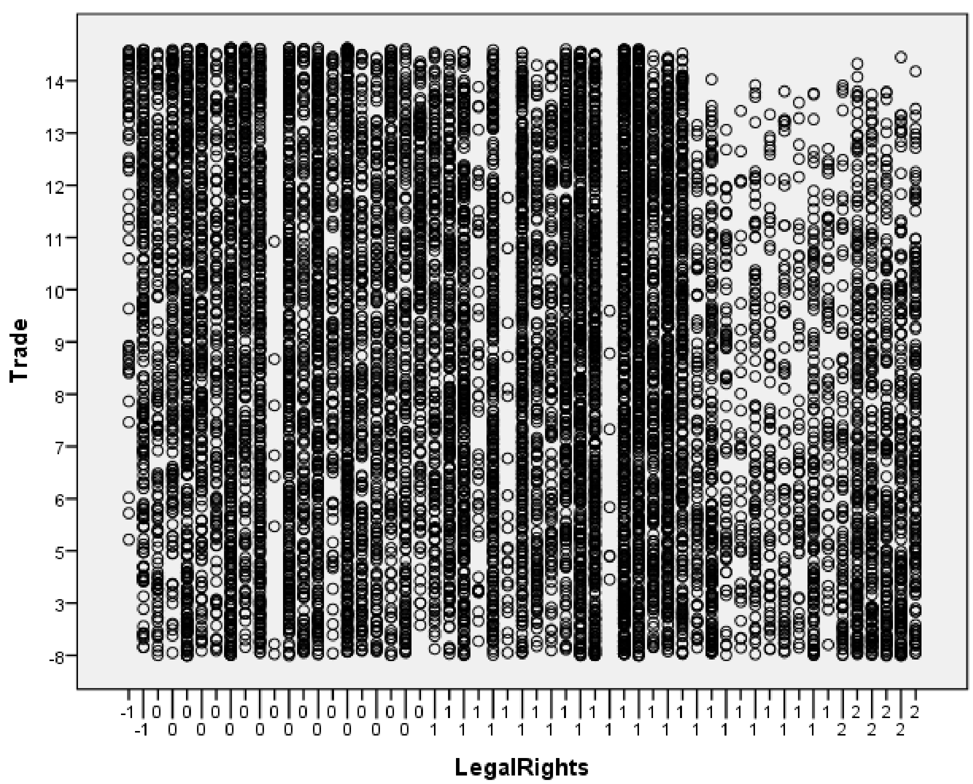


Trade
14
13
12
11
10
9
8
7
6
5
3
-8
-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1
Language



Trade
14
13
12
11
10
9
8
7
6
5
3
-8
-17
-13
-11
-9
-7
-4
1
2
Enforcing


Trade
CorruptionFight


Trade
LegalRights

I partial Correlations and model results Institutions and settings

	Island	Freeport	Customs	PoliticalUnion	Currency	Border	Colony	Language	CommonSea	Enforcing	CorruptionFighting	LegalRights
Island	1											
Freeport	0.011	1										
Customs	−0.0	−0.0	1									
PoliticalUnion	−0.0	−0.0	0.60	1								
Currency	−0.0	−0.0	0.39	0.399**	1							
Border	0.00	−0.0	0.154	0.154**	0.232	1						
Colony	0.014	0.007	−0.0	−0.076	−0.0	0.011	1					
Language	0.060	−0.0	0.063	0.063**	0.155	0.156	0.13	1				
CommonSea	0.036	0.004	0.062	0.062**	0.079	−0.0	0.05	0.045	1			
Enforcing	0.184	0.025	0.007	−0.007	−0.0	−0.0	−0.0	−0.0	−0.008	1		
CorruptionFighting	0.39*	0.017	−0.1	−0.134	−0.1	−0.0	0.01	−0.0	−0.01	0.143	1	
LegalRights	0.40*	0.026	−0.0	−0.030	−0.0	0.013	−0.0	−0.0	0.014	0.144	−0.002	1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
Trade	0.73 (1.000)	0.757 (1.006)	0.744 (1.031)	0.748 (1.036)	0.731 (1.130)	0.721 (1.173)	0.736 (1.329)	0.735 (1.329)	0.735 (1.333)	0.737 (1.399)	0.736 (1.340)	0.737 (1.349)	0.738 (1.352)	0.738 (1.352)
M		−0.2 (1.006)	−0.2 (1.007)	−0.2 (1.048)	−0.2 (1.048)	−0.2 (1.049)	−0.2 (1.060)	−0.2 (1.049)	−0.2 (1.259)	−0.2 (1.381)	−0.2 (1.381)	−0.2 (1.350)	−0.2 (1.500)	−0.2 (1.503)
Legal Rights			−0.0 (1.026)	−0.0 (1.030)	−0.0 (1.049)	−0.0 (1.050)	−0.0 (1.051)	−0.0 (1.060)	−0.0 (1.055)	−0.0 (1.050)	−0.0 (1.056)	−0.0 (1.057)	−0.0 (1.057)	−0.0 (1.057)
Language				0.067 (1.0490	0.067 (1.049	0.057 (1.095	0.067 (1.053)	0.063 (1.051)	0.060 (1.084)	0.060 (1.055)	0.060 (1.084)	0.059 (1.057	0.059 (1.057)	0.060 (1.095)

				)										
Freeport					0.058 (1.095	0.057 (1.059	0.055 (1.098)	0.056 (1.053)	0.056 (1.099)	0.056 (1.099)	0.056 (1.099)	0.056 (1.085)	0.056 (1.095)	0.056 (1.099)
Enforcing						0.051 (1.059)	0.042 (1.110)	0.042 (1.098)	0.042 (1.110)	0.042 (1.110)	0.039 (1.131)	0.039 (1.099)	0.039 (1.099)	0.039 (1.099)
Island							0.041 (1.176)	0.041 (1.110)	0.041 (1.176)	0.042 (118076)	0.042 (1.181)	0.042 (1.132)	0.042 (1.181)	0.042 (1.181)
PoliticalUnion								0.042 (1.176)	0.044 (1.267)	0.043 (1.268)	0.046 (1.286)	0.057 (1.181)	0.058 (1.869)	0.058 (1.869)
Colony									0.024 (1.026)	0.025 (1.030)	0.025 (1.030)	0.025 (1.032)	0.025 (1.032)	0.024 (1.032)
Commonsea										0.024 (1.085)	0.024 (1.085)	0.024 (1.085)	0.024 (1.095)	0.024 (1.095)
CorruptionFight											0.021 (1.049)	0.020 (1.049)	0.020 (1.049)	0.020 (1.049)
Customs												−0.0 (1.663)	−0.0 (1.663)	−0.0 (1.663)
Border													−0.0 (1.174)	−0.0 (1.174)
Currency														0.002 (1.174)
Adjusted R-Squared N	0.54	0.620	0.626	.631	0.634	0.636	0.637	0.639	0.639	0.640	0.640	0.640	0.640	0.640