



**HAL**  
open science

## The local determinants of interregional shipping flows

César Ducruet, Claudio Ferrari, Hidekazu Itoh, Alessio Tei

► **To cite this version:**

César Ducruet, Claudio Ferrari, Hidekazu Itoh, Alessio Tei. The local determinants of interregional shipping flows. César Ducruet. *Advances in Shipping Data Analysis and Modeling. Tracking and Mapping Maritime Flows in the Age of Big Data*, London, Routledge, pp.231-250, 2017, Routledge Studies in Transport Analysis, 9781138280939. hal-01623620

**HAL Id: hal-01623620**

**<https://hal.science/hal-01623620>**

Submitted on 12 Nov 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The local determinants of interregional shipping flows

César DUCRUET, Claudio FERRARI, Hidekazu ITOH, Alessio TEI

Pre-final version of the chapter published in Ducruet C. (Ed.) (2017) *Advances in Shipping Data Analysis and Modeling. Tracking and Mapping Maritime Flows in the Age of Big Data*, Routledge Studies in Transport Analysis, pp. 231-250.

The analysis of the determinants of shipping flows includes a wide array of approaches, which is mainly the focus of transport and maritime economics as well as operations research. The majority of related studies take the firm as the core unit of analysis, trying to unravel the determinants of port selection and shipping network design from a company perspective (Tiwari et al., 2003; Tongzon, 2009), using either quantitative or qualitative methodologies. While “location” is often cited as one of the numerous determinants, space in itself is considered in very basic ways, such as Euclidian or nautical distance between ports as a complementary proxy for time and cost parameters in various models (Wilmsmeier, 2014). This, for example, is the case of numerous studies using the gravity model to explain the intensity and distribution of trade flows among countries and regions of the world, but often overlooking the maritime mode (Guerrero et al., 2015). Another focus on shipping flows mainly describes the topology of the network made of port nodes and their mutual linkages (Ducruet and Itoh, 2015), with reference to general models of networks (see Kaluza et al., 2010), as a means to detect industry-specific influences such as the emergence of hub-and-spokes systems and global transport players in various contexts (McCalla et al., 2004).

Overall, the local, socio-economic background of shipping flows is much neglected in contemporary literature. The role of the local economy in shaping flows has been investigated in the case of port cities and regions (Ducruet and Itoh, 2016), but without a relational or network perspective (**see also Chapters 13 and 17 on the relationships between shipping, inland accessibility, and regional development, respectively**). Most of the time, scholars preferably look at the economic impacts of ports on local economies (Ng

et al., 2014; Bottasso et al., 2014). This chapter thus adopts a different perspective, by raising the question on how the intensity, distribution, specialization, and topology of shipping flows is determined not only by operational but also territorial factors. Its main objective is not to underestimate the role of supply chain actors and their main requirements (i.e. cost, time, performance, reliability, etc.) but to provide a more balanced view on the influence of demand factors on the pattern of global shipping networks. The relational or network perspective with regard to territorial factors implies that not only port nodes and their interlinkages are considered, but their surrounding socio-economic environment, which is barely mentioned in related studies, except for a few where roads networks (Tavasszy et al., 2011) or land use and other territorial features (Nelson, 2008) are added as complementary indicators. Strangely enough, this approach is new and necessitates a number of methodological choices in terms of database construction and variable selection.

First, a global matrix of inter-port vessel movements was constructed for the year 2008 by aggregating the daily flows of most of the world's merchant fleet, based on original shipping data provided by *Lloyd's List Intelligence* for the months of March, June, September, and December. Such vessels were classified under six main categories (i.e. containers, general cargo, liquid bulks, solid bulks, passengers, and vehicles), while their carrying capacity in deadweight tons (DWT) was attributed to every port of call and inter-port link, the total amount of traffic being a product of ship capacity and call frequency. Links between ports can be either direct calls through a one-stop voyage, or indirect calls with intermediary stops in between, along the whole and uninterrupted vessel voyage. In the meantime, port nodes were aggregated with each other at the level of subnational administrative regions in order to work on a broader unit of analysis (the port region). This allowed collecting supplementary data for the year 2008/2009 on socio-economic activities, population, productivity, etc. from various sources ranging from national accounts and transport ministries to broader online databases such as the OECD Territorial Database or Eurostat. Lastly, network-specific indicators were calculated using the TULIP software on port region centrality but also link centrality, as well as other classic indicators on traffic concentration

and specialization. Most of the time, regional indicators such as those on employment and regional gross product (RGDP) were transformed into location quotients and the share of national total to take into account the national and global context and also to make local economies comparable using relative rather than absolute figures (see Appendix 14.1 for a full list and description of link variables).

The main novelty of the present analysis is thus to provide a joint analysis of regional and shipping indicators at global level for no less than nearly 300 port regions situated in about 40 developed and developing countries. In addition, and as seen in the following sections, it shifts the focus from node level (port region) to link level (inter-port region), creating new variables that translate similarities or gaps in terms of socio-economic and traffic structure by port region pairs. The remainder of this chapter are organized as follows: the second section deals with port regions as nodes in a global maritime network, where we investigate the degree of affinity between regional specialization and traffic specialization, as a first step towards the understanding of the local determinants of shipping flows. The third section adopts a complementary perspective based on the concepts of homophily and assortativity in maritime networks (and more specifically in spatial networks), namely how port regions of similar or contrasted profiles connect through certain types of shipping flows. The fourth section gives more importance to the distance parameter, or friction of trades, by applying a gravity model to interregional (long-distance) shipping flows and contrasting the results with observed flows. In the conclusion we discuss the outcomes of this research and its contribution to existing regional and transport studies, while examining possible transfers in terms of decision-making support.

### ***Traffic structure, regional specialization, and port region centrality***

In this section we examine the interplay between the socio-economic and traffic variables that characterize network nodes, in our case subnational port regions. A Principal Component Analysis (PCA) was applied to two different global datasets, which mainly differ

by number of employment categories; the first dataset comprises 273 regions based on three main economic sectors, while the second includes fewer regions (222) but it details seven economic sectors (agriculture, industry, construction, trade, finance, public, and other services). For each analysis, we retained only the four main load factors (or principal components), which concentrate 54.8% of total variance for the first sample with 25 variables; and 52.2% (as above) for the second sample with 28 variables. Such a result in itself suggests that the sample under study is characterized by heterogeneity; as traffic, network and local economic indicators have complex relationships and accordingly their explanation can be similarly complicated.

Yet, the analysis based on 273 port regions reveals a number of interesting trends that validate earlier works on port regions (see Ducruet and Itoh, 2016, for a review and empirical test based on port tonnage). As seen in Figure 14.1, the main trend (F1) is an opposition between two different types of port regions. Positive scores underline a concomitant importance of traffic volume (DWT), concentration, network centrality, geographic reach, demographic and economic size and concentration, population density, RGDP per capita, with a specialization in the tertiary sector and in container as well as vehicle flows, the most valued commodities. This is the typical profile of bigger and richer port regions with a dominant position in the maritime network and concentrating the majority of maritime flows. Thus, this profile is a combination of mass or size effect (large urban nodes) and specialization effect (high-valued cargo, economic productivity and prosperity). The opposite profile is characterized by a combined specialization in solid bulks, general cargoes in the primary sector and, to a lesser extent, in liquid bulks and passenger traffic in the secondary sector, with a slightly more pronounced unemployment rate compared with the national average. Such a profile corresponds to regions being more peripheral at national and global scale in terms of size and situation, based on more traditional or labor-intensive activities and less valued cargoes.

The second factor (F2) provides a complementary trend whereby regions specialized in the primary sector can also be very central in the network and have far-reaching spatial connectivity, especially because solid and liquid bulk traffics, unlike liner shipping, are often

operated through direct origin-destination and long-distance flows. In turn, the opposite profile (positive F2 score) is specialized in more valuable traffics (general cargo, containers, vehicles), and tends to have a very crucial importance within its national territory in terms of population, productivity and traffic concentration, with a more tertiary profile, but a lower centrality in the global shipping network and higher unemployment rate. Four types of port regions can thus be defined based on the crossing of the first two components: large urban regions specialized in services and containers/vehicles (top-right quadrant); larger and longer traffic nodes with a strong centrality in the shipping network (bottom-right); primary sector regions specialized in bulk traffics (bottom-left); and general cargo regions with higher unemployment rate and industrial specialization (top-left).

Although it excludes numerous regions situated in developing countries where disaggregated data on employment categories is not available, the analysis of 222 port regions provides additional evidence on such trends. While the main oppositions are very similar to the former analysis, one can observe that the top-right quadrant remains defined by a profile of large urban regions concentrating economic wealth, population and container traffic, but also and more specifically, financial and insurance activities as well as “other services”. These core economic regions are thus also the most advanced in terms of shipping technology and capital-intensive activities. The same opposition occurs with the bottom-left quadrant with port regions specialized in agricultural activities, solid and liquid bulks and passengers, but also public services and manufacturing activities to a lesser extent. Interestingly, centrality indicators remain grouped in the bottom-right quadrant, but together with the specialization in “distributive trade, repairs, transport and accommodation” that concentrates most of the port, logistics and transport-related activities.

[Figure 14.1]

A complementary analysis is proposed based on six quantiles applied to selected key indicators to reveal how regional and shipping indicators, which are standardized by the average of each indicator for the comparison between indicators, statistically fit with each

other (Figure 14.2). When taking demographic population as the dependent variable, we see that most of the shipping indicators increase in relation to the number of inhabitants from the lower to the higher quantile, especially betweenness centrality that exhibits the widest gaps among quantiles. Thus, again here more populated regions tend to be more central in the network (betweenness centrality, degree centrality, clustering coefficient) and to exert longer-distance shipping linkages on average. Very similar results were obtained based on RGDP as dependent variable, with the exception of the average kilometric length of shipping linkages, which, albeit fuzzier than those based on population, still exhibit the highest value for richest regions. Overall, the same goes for population density and RGDP per capita taken as location quotients (no figures because of space constraint). This means that not only the port handling itself but its socio-economic and demographic performance regarding the rest of the host country has a clear and noticeable influence on traffic intensity, interaction range and network centrality. In turn, port regions that are more specialized, based on the Herfindahl-Hirschman Index (HHI), or where traffic structure is more concentrated around fewer cargo types, tend to be more populated and richer.

[Figure 14.2]

### ***Maritime linkages, homophily and assortativity***

This approach consisted in further understanding how maritime linkages are influenced by the socio-economic characteristics of the connected regions. However, since the unit of analysis becomes the inter-port link itself, it is difficult to conceptualize the territorial embedding of this link in two regions simultaneously. The chosen solution has been to reason in terms of homophily and assortativity, two relatively common approaches in network analysis, in particular social networks for the first and complex networks for the second. Homophily in social networks, notably those that are spatially embedded (Wong et al., 2006), is defined by stronger connectivity probability between nodes sharing certain common features than between contrasted nodes. Assortativity is a parent concept, with a

noticeable difference that the same probability is based on the size of nodes. In other words, a network will exhibit assortative mixing when nodes of similar size (i.e. based on degree centrality) are better connected with each other than nodes of dissimilar size, in the latter case the network being more disassortative (Newman, 2002). Empirical studies showed, for instance, that airline networks get more disassortative as the network grows, due to deregulation fostering the emergence of hub-and-spokes configurations (Barrat et al., 2005). In this chapter, both concepts target a better understanding of maritime flows in relation to local socio-economic development.

While the empirical testing of such ideas in a maritime context does not yet exist, it faces a number of conceptual and methodological issues. At conceptual level, it is relatively straightforward to test assortativity since maritime transport, in particular container shipping, adopted similar configurations to airlines in the mid-1990s, namely the establishment of transshipment hub ports at strategic locations. Yet, the whole shipping industry does not follow this network design since it is also based on coastal and short-sea shipping, direct origin-destination flows, such as for bulk shipping, as well as itineraries, such as for cruise shipping, which do not depend on hubs. In terms of homophily, the extent to which similar port regions will be (a) more strongly connected and (b) connected by specific types of flows is not an easy question. In particular, maritime transport had increasingly become one element only of wider value chains (Robinson, 2002), which means that it does not constitute a complete system on its own, if land transport and logistics between the true origins and destinations of flows are not included. Maritime transport is thus a segment connecting regions scattered across the world at different phases of value creation, from raw materials to manufactured goods. This leads us to question the relevance of testing homophily on the sole basis of maritime flows (given that many port regions are transit nodes, not the production or consumption markets, e.g. Ducruet and Itoh, 2016). In addition, our maritime network based on vessel movements is an artefact from which it is difficult to assess whether port regions belonging to the same shipping route truly trade with each other.



Despite such issues, it has been possible to measure in relatively simple ways how homophily and assortativity determine (or not) the topological and spatial distribution of the global maritime network. First, each interregional shipping link is characterized by a number of variables describing its own profile, such as traffic volume, geographic length (orthodromic distance), number of handled commodity types and HHI for their capacities and various centralities (betweenness, clustering, strength). Second, the territorial embedding of each link is measured by socio-economic differentials among region pairs. One group of such variables is a ratio (i.e. gap index, 1 being completely similar size and 0 being very dissimilar size) between the minimum and the maximum value (i.e. population, RGDP), and another looks at the correlation (Pearson) between the employment structure of the connected regions. Such gaps and correlations thus become attributes of the link itself and are compared with traffic/network variables. Two complementary analyses are proposed, namely a multivariate analysis of combined traffic and socio-economic indicators on links, and descriptive statistics based on quantiles.

The Principal Components Analysis (PCA) is a first approach to test the possibility for global maritime flows to exhibit homophily and assortativity (Figure 14.3). In other words, are similar port regions more connected with each other than dissimilar port regions, and through similar or dissimilar types of shipping flows? Given the vast number of links under scrutiny (22,578), our PCA results provided a large number of principal components with an eigenvalue over 1, i.e. fourteen components for all 41 variables included (75.4% of cumulated variance), and eight components for 22 variables excluding certain redundant variables (69.5% of cumulated variance). Our discussion on such results will, however, mainly focus on the first four components, reaching about half of the total cumulated variance in both cases (42.1% for all variables and 47.1% for selected variables), to highlight the main groupings and oppositions at stake. In terms of traffic characteristics, we observed that shipping links carrying bigger traffic volumes (DWTs) are generally more diversified (number of cargo types), with a tendency to specialize in containers, but shorter in terms of kilometeric length.

This is in line with former works showing that most of the world's container traffic concentrates over links of 500 kilometers or less (Ducruet and Notteboom, 2012), while container flows are specific in a sense that they present the greatest overlap with other traffic types in the global network, while most diversified links are also shorter (Ducruet, 2013). This explains why such bigger links are less international on average (i.e. more domestic). When it comes to region pairs in terms of traffic, there is a clear assortativity of the network since bigger links preferentially connect port regions of similar traffic weight, commodity diversity, network position (i.e. degree centrality, clustering coefficient on links) and geographic reach. In turn, links of smaller traffic volumes tend to connect more distant regions, to be more specialized (higher HHI), and to have a more crucial position in the network (strength links, betweenness centrality links), with a pronounced specialization in passengers, general cargo, and solid bulk. In terms of homophily when focusing on the demographic and socio-economic similarity of connected regions, bigger links tend to concentrate in more productive regions (higher RGDP) as well as between regions of similar demographic size (higher gap index) and economic structure to a lesser extent. These results are meaningful in the sense that they empirically demonstrate that the spatial distribution of the global shipping network is far from being randomly organized. On the contrary, spatial structures greatly influence the intensity and profile of shipping flows, notwithstanding noticeable oppositions between bigger and smaller flows.

[Insert Figure 14.3 about here]

A complementary analysis consisted in classifying links according to six quantiles to check the mutual influences between territorial and logistical elements in the network (Figure 14.4, standardized by the average of each indicator for the comparison between indicators). Based on demographic (population) and economic productivity (RGDP) gaps among region pairs, it appeared that connected regions of similar weight share a similar network centrality, as mentioned above, while being connected through less crucial links (i.e. lower betweenness centrality links and strength index), and at the same time more densely connected with their neighbors (clustering coefficient). The gap of hub dependence (vulnerability) is narrower between regions of similar economic weight (RGDP), which is less

apparent from a demographic perspective. This means that the economic productivity of connected regions is perhaps a better indicator than population to highlight similarities in network connectivity. Finally, results were compared based on the similarity (Pearson correlation coefficient) between the respective employment structure of port regions, for 3 and 7 economic sectors. In the case of 3 sectors, connected regions with a similar employment structure (higher quantiles) tend to have a wider gap (lower gap index) in most centrality indicators (i.e. betweenness centrality gap, degree centrality gap, betweenness centrality links) and to be connected through more diversified traffic linkages (i.e. lower HHI). This means that economic differentiation, as a potential complementarity between connected regions, fosters similarity in terms of network performance and specialization in terms of commodity structure. Such results underline that socio-economic homophily is associated with centrality disassortativity, and vice-versa. When it comes to 7 sectors, results are similar only for the betweenness centrality links, since the opposite is shown for all other gaps and link characteristics. One likely reason is the restricted availability of 7-sector data for regions in developing countries: connected regions in more developed (OECD) countries show a slight tendency to share both homophily and assortativity. In other words, such connected regions are more similar and less complementary to each other. In addition, more specialized links (i.e. higher HHI) connect regions with wider socio-economic and demographic gaps (i.e. lower gap index). Thus, more diversified links (i.e. lower HHI) are found between regions of similar size and productivity (higher gap index). These trends confirm previous research showing that the most diversified maritime links actually concentrated between port cities with similar total port traffic size (Ducruet, 2013). In this analysis, diversified links are also seen to concentrate between demographically and economically similar regions.

[Figure 14.4]

## ***Gravity model***

Maritime networks are driven by forces related to connectivity between ports (i.e. the nodes). As noted by Lam and Yap (2011), shipping networks interconnect different kinds of ports, while the role of these ports within the network also affects their competitive position within global supply chains. On the other hand, maritime network development prioritizes main nodes, creating a hierarchy that influences port competitiveness and its role in respect of the main shipping routes, and additionally serving major hinterlands (e.g. Ducruet et al., 2010). This scenario is connected with the possibility of studying maritime flows as driven by the port traffic volume, in which main distribution nodes (i.e. hubs) can have an advantage in comparison with minor nodes (i.e. minor ports, or satellites on the network). Moreover, the presence of frictions (e.g. transport and logistics costs) that differ in relation to both “distance” and “specific route” issues is also a characteristic of the current maritime network. The spatial interaction –within a given network– between nodes that partially compete through their traffic volume and different costs depending on network design naturally calls for the use of a spatial interaction model, such as gravity.

Gravity models are a specific kind of spatial interaction analysis in which the flow between two “nodes” –and the related network– carries “hidden costs” (i.e. frictions) that affect flow distribution with respect to an ideal or optimal scenario. Gravity models are often used in regional economics (e.g. Karemada et al., 2010; Burger et al., 2009) and related disciplines (e.g. Khadarooa and Seetanah, 2008), with several applications in transportation science (e.g. Simini et al., 2012; Anas, 1983). Concerning maritime studies, gravity models were applied to hinterland flows connecting specific ports (e.g. Ferrari et al., 2011; Guerrero, 2014).

As suggested by Haynes and Fotheringham (1984), a general gravity model appears as:

$$T_{ij}=A_i O_i D_j B_j (d_{ij})$$

(1)

in which T equals the “interactions (i.e., maritime traffic volume)” between regions i and j; O and D are the “production” and “destination” flows; while A and B are “balancing

constraints” for the origin or the destination, respectively. If A and B are used simultaneously, the gravity model is called “doubly” constrained, while if only one of them is used the model can be “production” (in the case of A) or “destination” (in the case of B) constrained.

In general terms, where:

$$A_i = \left[ \sum_j B_j D_j f(d_{ij}) \right]^{-1} \quad (2)$$

$$B_j = \left[ \sum_i A_i O_i f(d_{ij}) \right]^{-1} \quad (3)$$

A and B should be solved interactively in order to identify the constraints related to the in- and out- flows of origin and destination.  $(d_{ij})$  represents the effect of the physical distance between regions in the spatial interaction. Often  $f(d_{ij})=d_{ij}^\beta$  where  $\beta$  shows the sensitivity to the distance of the studied variables and the effect of potential frictions. In our case, Euclidian distance has been used between port regions attracting different kind of cargoes. This choice is linked to the main differences in the organization of the shipping services that is actually affecting the shipping distance (e.g. transshipment, not direct call services).

In order to evaluate spatial interaction between the maritime regions, SIMODEL software has been used. This software allows the estimation of the gravity model, comparing the actual  $T_{ij}$  and the predicted  $T'_{ij}$ . As said, the application of either A or B (or both) changes the model; in particular when both are applied traffic between each node is kept constant, whereas if only one is applied, the traffic generated by each origin remains constant both in the observed ( $T_{ij}$ ) and the predicted ( $T'_{ij}$ ) matrix, while the traffic attracted by each destination is redistributed among them.

In order to solve the model, a hypothesis on the  $f(d_{ij})$  is necessary; in accordance with main literature (e.g. Ferrari et al., 2011), we used a power function equal to  $(d_{ij}^\beta)$ . In such a model, the parameter  $\beta$  represents the “probability” for the estimating values to be equal to the real value, therefore, representing the relation between the distance and the “generalised

costs" (i.e. frictions that affect the model) that impact in the development of the network. Since distance should have a decay effect, the parameter  $\beta$  is expected to be negative. Moreover,  $\beta < -1$  means costs have an effect more than proportional in respect to distance, while if  $-1 < \beta < 0$  costs are less proportional than distance. When  $\beta$  is equal to zero there are no frictions and the distribution has no connections with the distance. Another interesting outcome is the standard deviation ( $e$ ) that represents a goodness of fit in the model. Moreover, the standard deviation also provides information on the redistribution among the regions of total traffic and containers only (TEU).

The analysis targets the NUTS2 regions belonging to our database, representing the maritime network. The main results provide a comparison of NUTS2 interactions along three main shipping routes (e.g. Transpacific, Transatlantic, and Far East-Europe) for the discussion on long-distance trade excluding short-sea and coastal shipping. In fact, the gravity model only takes into account the regions belonging to the above listed route, as representative examples of the differences within the current maritime network. Table 14.1 shows the results of the analyses for the production model and the doubly-constrained model.

[Table 14.1]

The analysis underlines how "frictions" act differently among the main trade routes while distance impacts on the distribution of commercial flows. Moreover, the standard deviation ( $e$ ) –that is a proxy for the degree of redistribution- is higher in the transatlantic matrix in comparison with the other routes, demonstrating a difference between predicted and observed values. In particular, higher values of both  $\beta$  and  $e$  for the Europe-America routes can be explained by a different traffic organization (e.g. fewer main nodes and different use of transshipment) in respect to the other two main routes, or including Asian trade.

Moreover, the high value of the standard deviation highlights a great difference among predicted and observed values. This issue underlines how some regions are able to attract more than what they might be, generating agglomeration effects at port regions. In general Asian and American regions seem to have a better distribution pattern than European ones, Figure 14.5, for instance, compares the results related to the trade routes of American and

European regions. In the figure, the axes represent the difference between the expected and real values of the studied routes per region (i.e. the European graph represents the differences characterising the European-Asia routes and European-Americas route). The fact that the majority of the regions are more or less concentrated around the origin means that expected and actual values do not differ too much (even if differences can be underlined). Thus, outliers in both graphs have specific explanations depending on the quarter to which they belong. For instance, ES70 (i.e. Canary Islands) and ES61 (i.e. Andalusia), representing two outliers located on the first quarter of “Europe”, demonstrate an unexpected flow from/to these regions in both the EU-Asia and transatlantic routes. Similarly, there is an imbalance situation in the ES52 (i.e. the Community of Valencia) outlier belonging to the fourth quarter of the same graph, but only for the EU-Asia route, while it has a negative difference in the transatlantic route. These outliers are often related to big hub ports (e.g. Algeciras, Valencia) that provoke a distortion to the “estimated” connections, deviating flows that would otherwise be “naturally” and more directly reaching other regions. Similar cases occur in the “Americas” graph, with regions hosting port hubs (e.g. California, South Carolina, and Florida) that became regional pivotal nodes. Interestingly, among the outliers there are also some few regions that register either heavy underperforming (e.g. Balearic Islands) or over-performing values (e.g. Alaska), not for structural reasons but for demographical and market specialisation (e.g. bulk and cruise) characteristics. Nevertheless, Figure 14.5 demonstrates how the shipping network reflects not only distance effects but also organisational effects (e.g. transshipment) that promote both scale and scope economies to overcome the distance factor.

[Insert Figure 14.5 about here]

## ***Conclusion***

This chapter underlines how contemporary maritime flows remain very much linked to the economic structure of the regions that they help to interconnect, contributing to existing

literature on the impact of the local economy of port regions and maritime flows, an area so far insufficiently researched. Preliminary findings from current research underline how it is possible through the study of maritime networks to differentiate regions according to their role within the network and with their economic characteristics. The PCA –and related analyses– showed how regional centrality in respect to specific traffic flows (e.g. liquid bulk, container, general cargo) depends on the regional environment (e.g. population, sectoral specialization, economic productivity), demonstrating how the international transport network is still shaped by local strategies and policies. Moreover, links carrying bigger traffic volumes are generally more diversified (in terms of lower HHI), with a tendency to specialize in containerized cargo, but shorter in terms of kilometric length. On the other hand, the gravity model showed that the main commercial routes are differently affected by distance- and transport-related costs. These findings showed how the different organization on main routes is still affecting main trade volumes among different world regions. The differences between real and expected values estimated by the gravity model also underline an unbalanced situation among the flows that is not fully explained by distance and is principally due to specialization and agglomeration effects (especially in the container business) in the port regions and in their neighboring regions too.

Moreover, even if the assumption related to distance (i.e. Euclidian distance) seems quite reasonable considering the differences between various maritime transport sectors, further research could improve the overall findings of the gravity model. In fact, nautical distance is only partially representative for all those kinds of traffic that do not use direct calls (e.g. containers) or, generally speaking, do not load/unload completely at every port of call (e.g. general cargo, passengers). As a matter of fact, the complexity of the freight distribution – and the related network– leads to a low representation of the connectivity using simple nautical distance within the region and for this reason Euclidian distance has been chosen. Nevertheless, further investigations could try to overcome this issue, differentiating the distance used for different kinds of services (e.g. container, bulk).

Thus, even if further research needs to be done, the findings of this chapter shed new light on the understanding of maritime network development and characteristics. Eventually, as



potential extensions of the current research, one could further differentiate the effects among different traffic categories in order to identify specific patterns that characterise the different shipping markets and their network development.

### ***Acknowledgements***

The authors would like to thank Ali El Hosni for his help with the database construction. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. [313847] "World Seastems".

### ***References***

Anas A. (1983) Discrete choice theory, information theory and the multinomial logit and gravity models. *Transportation Research Part B*, 17(1): 13-23.

Barrat A., Barthélemy M., Vespignani A. (2005) The effects of spatial constraints on the evolution of weighted complex networks. *Journal of Statistical Mechanics*, 5: P05003.

Bottasso A., Conti M., Ferrari C., Tei A. (2014) Ports and regional development: A spatial analysis on a panel of European regions. *Transportation Research Part A*, 65: 44-55.

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: 167-190.

Ducruet C. (2013) Network diversity and maritime flows. *Journal of Transport Geography*, 30: 77-88.

Ducruet C., Itoh H. (2015) The mutual specialization of port regions connected by multiple commodity flows in a maritime network. In: Ducruet C. (Ed.), *Maritime Networks: Spatial Structures and Time Dynamics*, London and New York: Routledge Studies in Transport Analysis, pp. 285-302.

Ducruet C., Itoh H. (2016) Regions and material flows: Investigating the regional branching and industry relatedness of port traffic in a global perspective. *Journal of Economic Geography*, 16(4): 805-830.

Ducruet C., Lee S.W., Ng A.K.Y. (2010) Centrality and vulnerability in liner shipping networks: revisiting the Northeast Asian port hierarchy. *Maritime Policy and Management*, 37(1): 17-36.

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

Ferrari C., Parola F., Gattorna E. (2011) Measuring the quality of port hinterland accessibility: The Ligurian case. *Transport Policy*, 18, pp. 382-391.

Guerrero D. (2014) Deep-sea hinterlands: Some empirical evidence of the spatial impact of containerization. *Journal of Transport Geography*, 35: 84-94.

Guerrero D., Grasland C., Ducruet C. (2015) Explaining international trade flows with shipping-based distances. In: Ducruet C. (Ed.), *Maritime Networks: Spatial Structures and Time Dynamics*, London and New York: Routledge Studies in Transport Analysis, pp. 303-321.

Haynes K.E., Fotheringham A.S. (1984) *Gravity and Spatial Interaction Models*. Beverly Hills: Sage Publications.

Kaluza P., Kölzsch A., Gastner M.T., Blasius B. (2010) The complex network of global cargo ship movements. *Journal of the Royal Society Interface*, 7(48): 1093-1103.

Karemera D., Oguledo V.I., Davis B. (2010) A gravity model analysis of international migration to North America. *Applied Economics*, 32: 1745-1755.

Khadarooa J., Seetanah B. (2008) The role of transport infrastructure in international tourism development: A gravity model approach. *Tourism Management*, 29: 831-840.

Lam J.S.L., Yap W.Y. (2011) Dynamics of liner shipping network and port connectivity in supply chain systems: analysis on East Asia. *Journal of Transport Geography*, 19: 1272-1281.

McCalla R.J., Slack B., Comtois C. (2004) The geographical hierarchy of container shipping networks in the Caribbean basin and Mediterranean Sea. Paper presented at the 10th World Conference on Transport Research (WCTR), Istanbul, July 4-8.

Nelson A. (2008) *Travel time to major cities: A global map of accessibility*. Global Environment Monitoring Unit, Joint Research Centre of the European Commission, Ispra, Italy.

Newman M.E.J. (2002) Assortative mixing in networks. *Physical Review Letters*, 89: 208701.

Ng A.K.Y., Ducruet C., Jacobs W., Monios J., Notteboom T.E., Rodrigue J.P., Slack B., Tam K.C., Wilmsmeier G. (2014) Port geography at the crossroads with human geography: between flows and spaces. *Journal of Transport Geography*, 41: 84-96.

Robinson R. (2002) Ports as elements in value-driven chain systems: the new paradigm. *Maritime Policy and Management*, 29(3): 241-255.

Simini F., González M.C., Maritan A., Barabási A.L. (2012) A universal model for mobility and migration patterns. *Nature*, 484: 96-100.

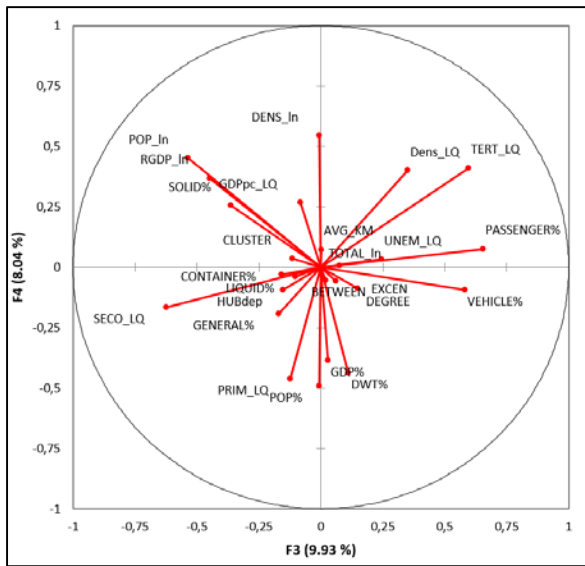
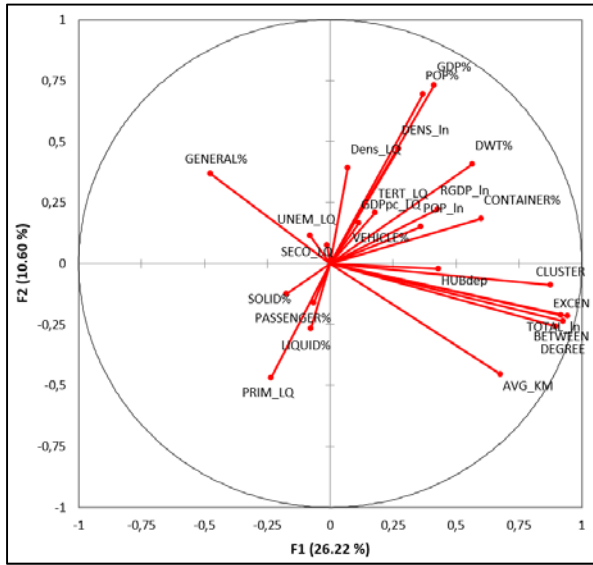
Tavasszy L.A., Minderhoud M., Perrin J.F., Notteboom T.E. (2011) A strategic network choice model for global container flows: Specification estimation and application. *Journal of Transport Geography*, 19(6): 1163-1172.

Tiwari P., Itoh H., Doi M. (2003) Shippers' port and carrier selection behaviour in China: a discrete choice analysis. *Maritime Economics and Logistics*, 5(1): 23-39.

Tongzon J.L. (2009) Port choice and freight forwarders. *Transportation Research Part E*, 45: 186-195.

Wilmsmeier G. (2014) *International Maritime Transport Costs. Market Structures and Network Configurations*. Aldershot: Ashgate.

Wong L.H., Pattison P., Robins G. (2006) A spatial model for social networks. *Physica A*, 360: 99-120.



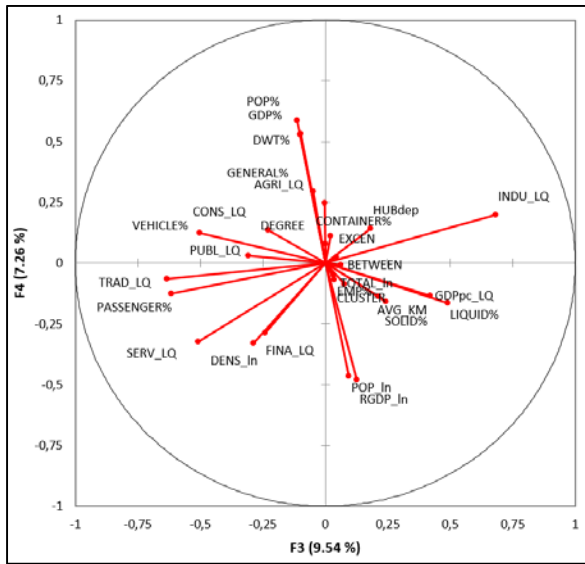
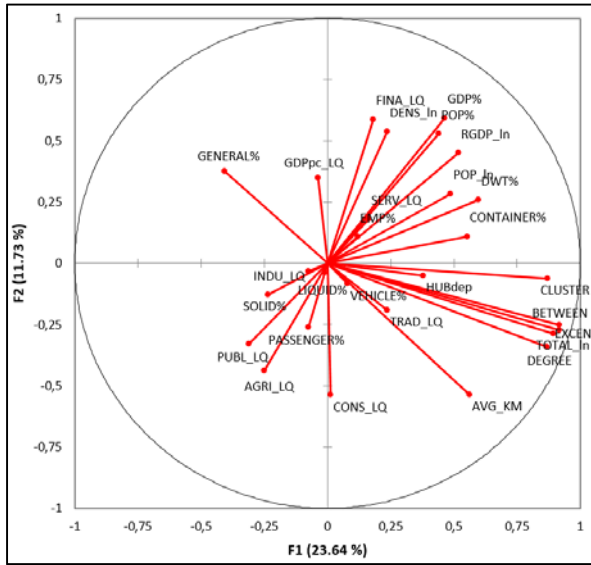


Figure 14.1: Principal Component Analysis of 273 port regions (top) and 222 port regions (bottom)

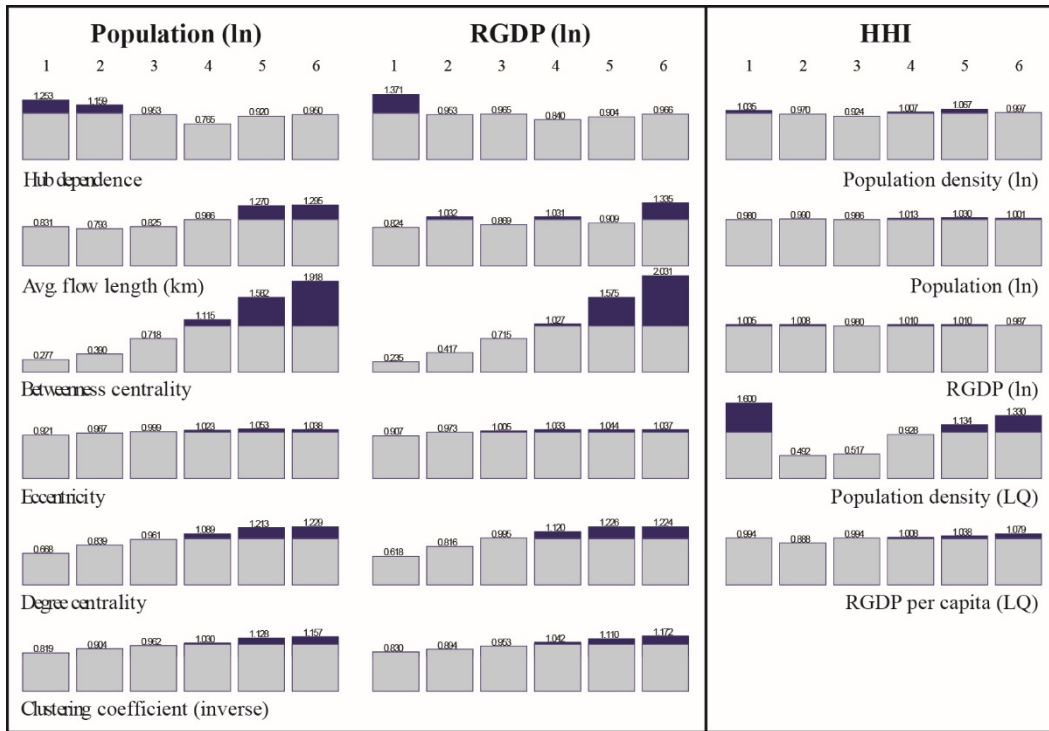


Figure 14.2: Mutual influences between selected regional and shipping variables a node level

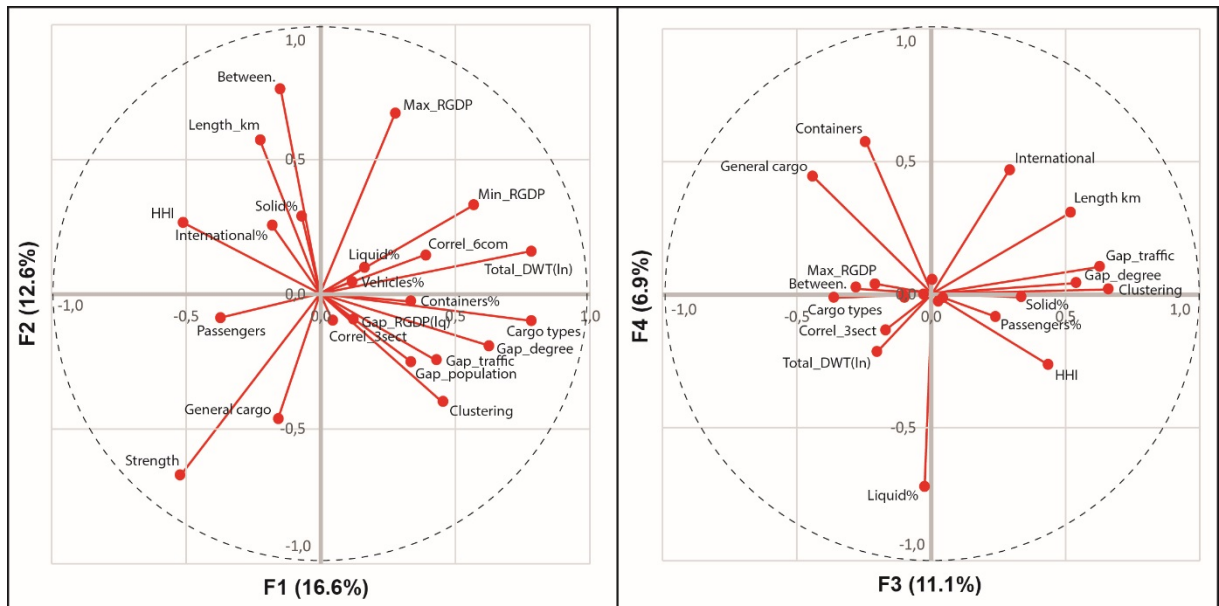


Figure 14.3: Principal Component Analysis of global maritime linkages

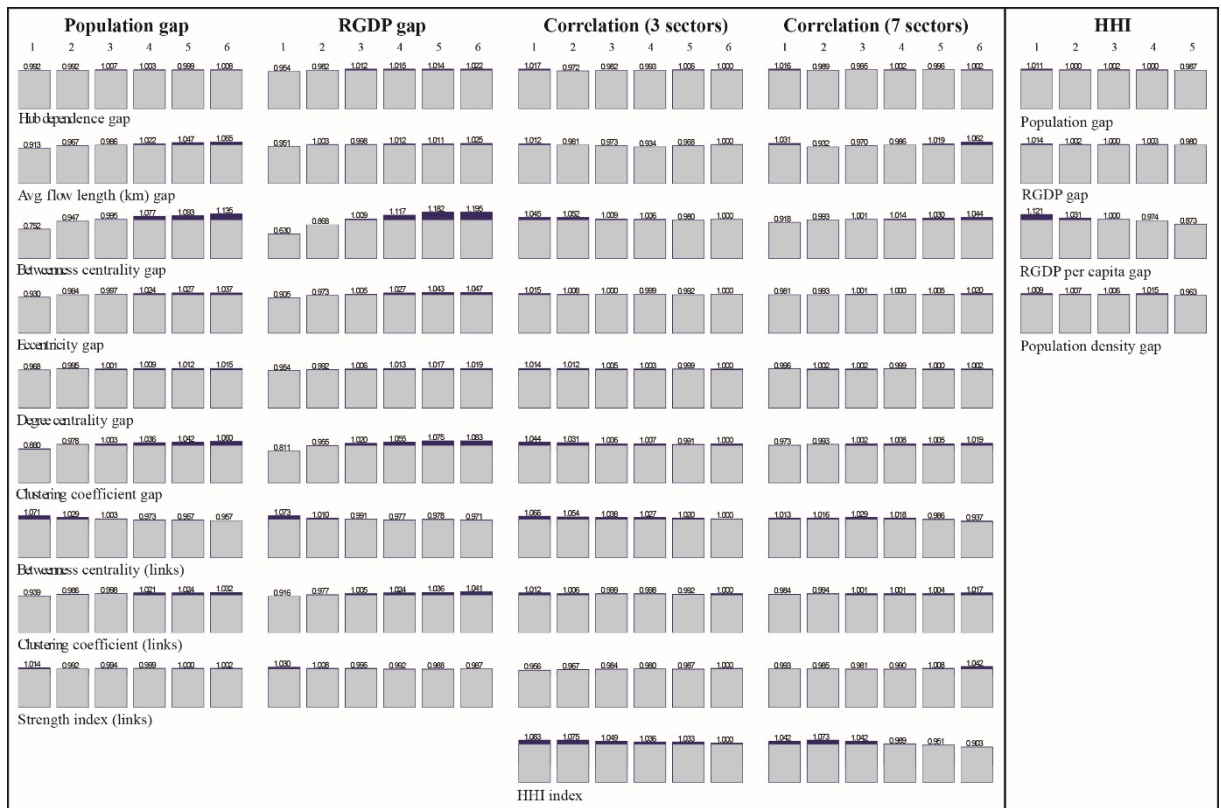


Figure 14.4: Mutual influences between selected regional and shipping variables a link level

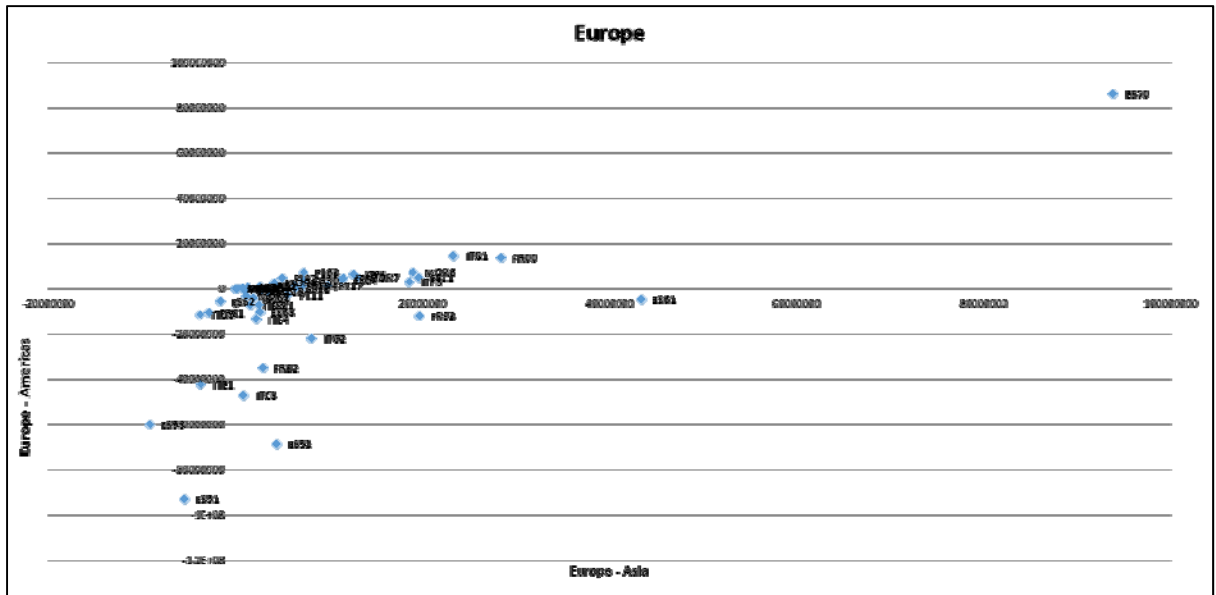
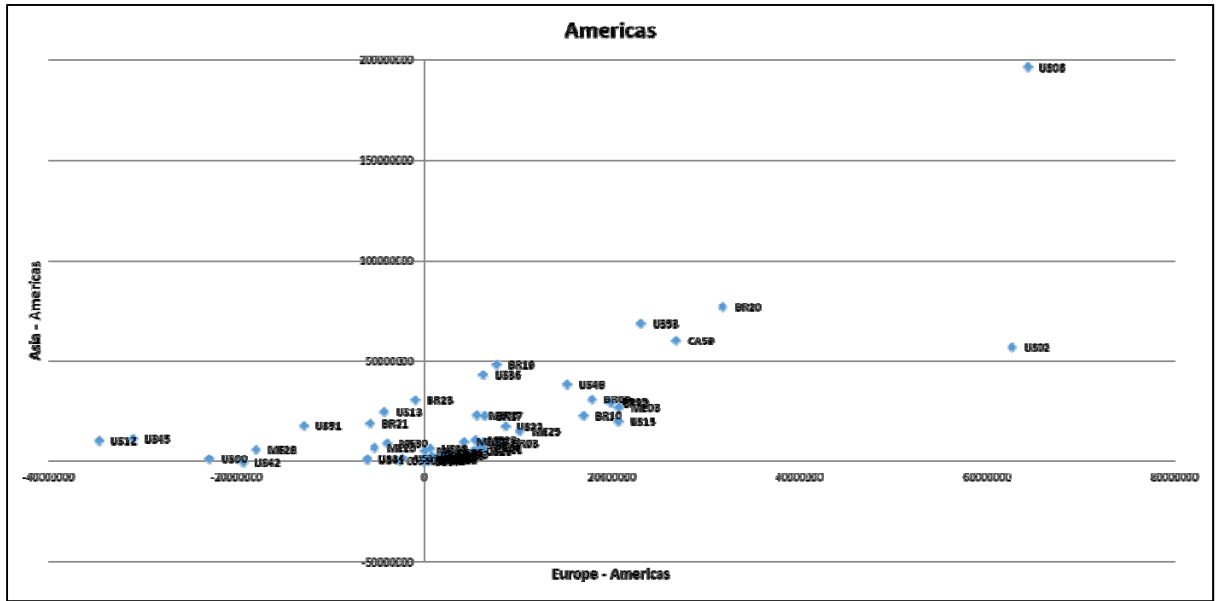


Figure 14.5: Position of port regions by observed/estimated traffic volume along selected major shipping routes



Production constraints model			
<b>TOTAL</b>	EU-Far East	Transpacific	Transatlantic
<i>B</i>	-0.44	-0.36	-0.50
<i>(e)</i>	65%	58%	81%
<b>Container (TEU)</b>	EU-ASIA	ASIA-AM	EU-AM
<i>β</i>	-0.34	-0.30	-0.37
<i>(e)</i>	52%	52%	63%
Doubly constraints model			
<b>TOTAL</b>	EU-Far East	Transpacific	Transatlantic
<i>β</i>	-0.48	-0.39	-0.53
<b>Container (TEU)</b>	EU-ASIA	ASIA-AM	EU-AM
<i>β</i>	-0.40	-0.33	-0.40

**Table 17.1: Gravity model results**

**Appendix 17.1: List of indicators retained in this study**

VARIABLE TYPE	VARIABLE NAME	DEFINITION
Node characteristics	POP_In	Natural log of total port region population, gross domestic product, population density, and vessel traffic
	RGDP_In	
	DENS_In	
	LN_DWT	
	POP%	Share of port region's population, RGDP or vessel traffic in host country total
	RGDP%	
	DWT%	
	RGDP_pc_LQ	Location quotient of RGDP per capita, population density, unemployment, primary, secondary, and tertiary sectors
	DENS_LQ	
	UNEM_LQ	
	PRIM_LQ	
	SECO_LQ	
	TERT_LQ	
	Clustering	Clustering coefficient or "hub power"
	Eccentricity	Topological proximity to other nodes
	Hub dependence	Share of the largest flow link in total traffic
	Degree	Number of adjacent neighbor nodes
Betweenness	Occurrences on shortest paths	
Avg_km	Average kilometric distance of shipping linkages	
Link characteristics	INTERNATIONAL	Dummy for domestic (0) or international (1) links
	CONTAINER	% of this commodity in total link traffic (%dwt)
	GENERAL	
	LIQUID	
	PASSENGER	
	SOLID	
	VEHICLE	
	LN_DWT	Natural log of total link traffic
	LN_DIST	Natural log of link's kilometric distance
	NB_COM	No. commodity types carried
	CC	Clustering coefficient
	STRENGTH	Strength index
Differentials among region pairs	COR_3SECT%	correlation coefficient between respective employment shares (%), absolute scores (ABS) or location quotients (LQ) based on 3 or 7 economic sectors
	COR_3SECT_LQ	
	COR_7SECT_ABS	
	COR_7SECT_LQ	
	COR_6COM%	correlation coefficient between respective traffic distribution (%) based on 6 commodity types
	GAP_POP	minimum score / maximum score among the 2 regions

GAP_RGDP
GAP_DWT
GAP_HUBDEP
GAP_AVGKM
GAP_BETW
GAP_CC
GAP_EXCENT
GAP_DEGREE