EJTIR

Issue 18(2), 2018
pp. 178-195
ISSN: 1567-7141
http://tlo.tbm.tudelft.nl/ejtir

Logistics clusters, including inter-firm relations through community detection

Joris Beckers[1]
Department of Transport and Regional Economics, University of Antwerp, Belgium.

Isabelle Thomas[2]
Center for Operations Research and Econometrics, Université catholique de Louvain, Belgium.

Thomas Vanoutrive[3]
Research Group for Urban Development, University of Antwerp, Belgium.

Ann Verhetsel[4]
Department of Transport and Regional Economics, University of Antwerp, Belgium.

This paper studies clusters in the logistics sector. Like traditional cluster research, indicators of concentration to detect co-location of employment are calculated. However, this approach is enhanced by including a quantitative analysis of the inter-firm relations between logistics companies through the use of a community detection algorithm on a microeconomic dataset of buyer-supplier relations. Combining both results in a typology of logistics clusters. Next to the big clusters characterized by employment concentration and many internal and external relations, spill-over and polycentric clusters are identified. This approach seems promising to detect in future research clusters in other sectors and places.

Keywords: *clusters, co-location, inter-firm relations, community detection, logistics.*

1. Introduction

Inter-firm relations and proximity are central topics in the economic geography literature (Giuliani, 2010). The geographical clustering of firms in the same or related industries is explained using the concepts of localisation and urbanisation economies. The presence of a dedicated infrastructure, a specialised labour market, easy communication and knowledge exchanges and a network of buyers and suppliers at their doorstep provide agglomeration advantages (Marshall, 1920). Academic scholars and policy makers promote these clusters as the ideal spatial organisation of economic activity due to these agglomeration effects (Martin & Sunley, 2003). Well documented examples of economic clusters are the information and communications technology hub in Silicon Valley (Bresnahan & Gambardella, 2004), the engineering cluster in Baden-Württemberg, Öresund medical cluster (Lundequist & Power, 2002), and the Aerospace Valley in Toulouse (Levy & Talbot, 2015). However, most of the cluster

[1] A: Prinsstraat 13, 2000 Antwerpen, Belgium T: +32 326 541 58 E: joris.beckers@uantwerpen.be
[2] A: Voie du Roman Pays 34 - L1.03.01, 1348 Louvain-la-Neuve, Belgium T: +32 104 721 36 E: isabelle.thomas@uclouvain.be
[3] A: Prinsstraat 13, 2000 Antwerpen, Belgium T: +32 326 540 85 E: thomas.vanoutrive@uantwerpen.be
[4] A: Prinsstraat 13, 2000 Antwerpen, Belgium T: +32 326 542 21 E: ann.verhetsel@uantwerpen.be

literature up till now uses a qualitative methodology. When a quantitative approach is used, the focus is mostly limited to measuring co-location of firms. Although Bennenworth et al. (2003) already note that the presence of one cluster characteristic (e.g. co-location) does not imply the existence of another (e.g. inter-firm relations), most of the cluster studies do not measure linkages when discussing the network's structure but limit themselves to analysing geographical concentrations of employment. While this can be explained by a lack of data, it remains imprudent to derive the existence of relations from the presence of spatial concentration (Benneworth et al., 2003; van den Heuvel, de Langen, van Donselaar, & Fransoo, 2014). The objective of this paper is to study co-location as well as inter-firm linkages in the Belgian logistics industry. We do this by identifying concentrations of employment as well as by analysing the geographical patterns of buyer-supplier linkages. There has been strong criticism on the lack of clarity of the term cluster (Martin & Sunley, 2003). The lack of clear sectoral and geographical boundaries and the confusing use of different terms like agglomeration, concentration, industrial district and cluster for similar concepts underlines the problem. In the present article we will stick to the concepts as presented in figure 1.

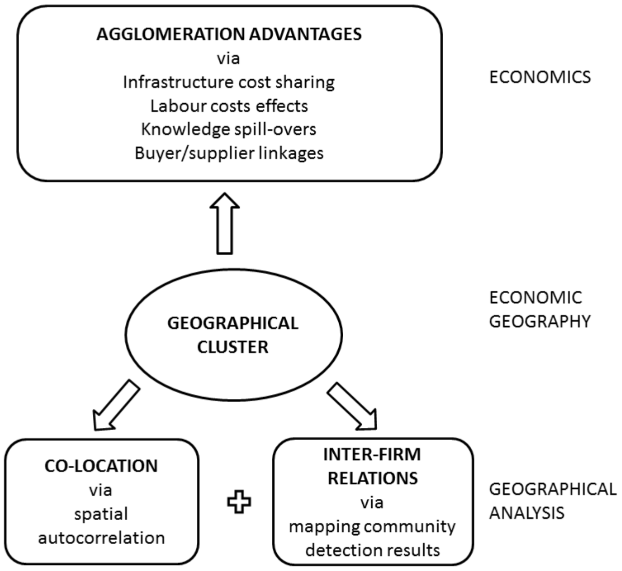


Figure 1. Conceptual scheme

The study of logistics clusters is relevant given the argument the sector drives economic growth in the region (Hesse, 2008; Rivera & Sheffi, 2016; Sheffi, 2012). Despite its key role as facilitator and handler of goods and related information in different value chains, the influence of globalization processes and a shift towards global standardization in the logistics industry has not ruled out the importance of regional characteristics (Akyelken & Keller, 2014; van den Heuvel, Rivera, et al., 2014; Verhetsel et al., 2015). Research has put attention on the (uneven) geographical pattern of logistics activities and several studies discuss the importance of agglomeration advantages by demonstrating the existence of logistics clusters. However, a quantitative approach including the analysis of both co-location and inter-firm relations is missing up till now. In addition, a differentiation through cluster characteristics can lead to a more varied typology than a simple cluster/non-cluster division. Logistics companies act as intermediaries that connect all stages of the supply chain. The heterogeneity of logistics firms, however, makes that neither a consistent nor a standardized notion of logistics exists (Verhetsel et al., 2015). On its homepage, the Council of Supply Chain Management Professionals (CSCMP, 2016) uses the following comprehensive definition: "that part of the supply chain that plans, implements, and controls the efficient, effective forward and reverse flow and storage of goods, services and related information between the point-of-origin and the point-of-consumption in order to meet customers' requirements". For the purpose of the present study, we include all kind of logistics plants, as well those dealing with material flows as those managing the services. As such we consider a quite heterogeneous group of firms ranging from economic activities in peripheral large industrial warehouses to services in inner-city offices.

In the following sections, we first sketch an overview of the cluster literature from the past 25 years and focus on the need for a typology of clusters. In section 3, the methodological steps taken in this paper are listed: the use of traditional indices of agglomeration and of graph clustering methods. Section 4 introduces the available data for Belgium and provides the results in order to study the spatial configuration of the logistic sector. Finally, section 5 provides a typology of logistics concentrations based on the results of the empirical analysis. It constitutes a comparative framework for further research in other regions. The paper ends up with concluding remarks and directions for further research.

2. Literature on clusters, identifying clusters and logistics clusters

The creation of clusters or regional specialized valleys has become an appealing policy strategy after the re-introduction of the cluster concept by Porter in the 1990s. Besides the more theoretical work on clusters (Porter, 1998; Bathelt, Malmberg & Maskell, 2004; Martin & Sunley, 2003), the strong interest and increasing funds resulted in a flow of both quantitative and qualitative research on the agglomeration of economic activities. Most of these studies limit their focus to the co-location of firms and/or employment, measured by the calculation of an aggregated value. An example is the work of Bertinelli et al. that assesses the concentration of manufacturing plants for several industrial sectors in a range of countries in Europe (Bertinelli & Decrop, 2005; Barrios et al, 2009). Similarly, Duranton and Overman (2005) calculate agglomeration by comparing the spatial pattern of several industries in the UK. Further, many more case studies combine several measures of co-location, all to prove the occurrence of clustering in space (Arbia, 2001; Riguelle, Thomas, & Verhetsel, 2007; Rivera, Sheffi, & Welsch, 2014; van den Heuvel, de Langen, van Donselaar, & Fransoo, 2012). Yet, while inter-firm relations are a crucial characteristic of clusters leading to the Marshallian externalities, very few cluster research includes the analysis of the inter-firm relations.

Notable exceptions are the work of Bell (2005) who use social network techniques to analyse inter-firm relations at the management level. Balland et al. (2012) explain the importance of geographical proximity of companies in the gaming industry by pointing to the ‘local buzz’ and the increasing technological complexity that demands strong inter-firm relations. Others investigate as well co-location patterns as a-spatial input-output datasets to define agglomeration of related industries (Porter, 2003; Ellison, Glaeser, & Kerr, 2010; Delgado, Porter, & Stern, 2010, 2016). Most often inter-firm relations are supposed to result from co-location without empirical evidence derived from the actual study of inter-firm relations itself. For example, Delgado et al. (2016) state that they derive inter-industry linkages “through the co-location patterns of industries across regions”. Other authors, like Hoffman et al. (2015), are more cautious and explicitly mention the inability to study the buyer-supplier linkages. The main reason for the omission of inter- and intra-industry trade data in cluster studies is due to their limited availability, especially at subnational geographical areas (Hoffmann et al., 2015; Martin & Sunley, 2003).

Although research indicates that economic activities are indeed clustered and that clusters provide at least a part of the advantages expected from agglomeration theory, research on the variety of clusters remains scarce. Markusen (1996) identifies, in addition to the Marshallian agglomeration, three other types of industrial districts, based on the varying roles of firms within the districts, and the orientation of the linkages amongst them: hub and spoke districts, satellite platforms and state-anchored districts. The author does mention several examples for each type, but the research lacks empirical demonstration of the findings. In Sweden Lundequist and Power (2002) classify clusters based on the policy process and vision that has led to their existence.

Although logistics have long been perceived as a consequence of the production process, i.e. as a derived demand (Hesse & Rodrigue, 2004), the actual extent of logistic activities inspire the proposition that logistics clusters, like industrial clusters, enjoy advantages from concentrating activities (Sheffi, 2012). Besides the mentioned agglomeration advantages for companies, logistics clusters are also considered as an important generator of employment and stimulus for growth in other economic sectors (Rivera & Sheffi, 2016; Sheffi, 2012). This partly explains why governments promote their country, region, province or municipality as a logistics cluster, in many cases despite the absence of empirical evidence for their claims (Flämig & Hesse, 2011; Hesse, 2015). But in other cases, the assumed opportunities of logistics clusters have led to the application of cluster research on the co-location of logistics activities. van den Heuvel et al. (2014a; 2014b; 2016) calculated the co-location of logistics companies, but because of a lack of data the inter-firm relations are only analyzed with a qualitative approach.

In contrast to previous studies, a microeconomic dataset of buyer-supplier linkages between the Belgian logistics companies was available for this study. This provides the opportunity to not only measure co-location, but to also empirically analyze the inter-firm relations in the logistics sector. This way, the present paper identifies economic clusters in a more comprehensive way. The use of a community detection algorithm for this purpose is to our knowledge novel.

3. Identifying clusters: integrate both co-locations and inter-firm relations

3.1 Measuring co-location: indicators of spatial association

There exist a large number of different indicators to measure spatial concentration. These indices can roughly be divided into two categories: a-spatial and distance based (van den Heuvel et al., 2014a). First, the a-spatial measures provide a global index of agglomeration for the entire study area under consideration. Examples are the index of Ellison and Glaeser (1997) and the locational Gini coefficient (Krugman, 1991). The latter is here calculated by comparing the employment share in a specific industry with the spatial unit's share of total employment. A value of 0 implies no spatial concentration (total dispersion) while 0.5 points to the opposite. Nevertheless, this category of measures offers no information about concentrations within specific areas. In addition, the estimated concentration score depends on the basic spatial units (MAUP), making comparisons difficult (Duschl et al., 2015; van den Heuvel et al., 2014a).

The second category of indices are distance-based agglomeration metrics, offering the opportunity to identify areas of higher concentration within a study area. The most popular is the Moran's I statistic. This measure calculates the spatial autocorrelation in a study area based on a spatial weights matrix that represents the geographical relationships between the spatial units. These weights can be calculated from a pre-defined number of neighbours, from a contiguity measurement that takes all neighbours into account, or by including only the values of spatial units within a threshold distance. Spatial autocorrelation is positive and high if spatially related units have similar values for a variable, e.g. logistics employment. A Moran's I of 1 indicates perfect spatial autocorrelation, -1 is perfect dispersion and 0 is a random distribution of the value of analysis (Arbia, 2001; Ding & Fotheringham, 1992).

The Local Indicator of Spatial Association (LISA) is further computed to analyse the local contributions to the Moran's I statistic (Anselin, 1995). This index calculates spatial agglomeration values for each spatial unit separately. A Moran's I scatterplot results that consists of four quadrants that correspond to the four types of spatial association: High-High (HH), High-Low (HL), Low-High (LH) and Low-Low (LL) (Anselin, 1995). Values in the HH quadrant thus represent spatial units with high values for the studied variable surrounded by units with similar (high) values, i.e. concentration. LL values correspond to units with low scores for the variable in question, situated in an area characterized by overall low scores. The HH spatial units identify concentrations, i.e. co-location.

3.2 Measuring inter-firm relations: community detection within networks

An industry's network consists of a myriad of financial (buyer-supplier), material and information flows, strongly distributed amongst different scales. Such a network is the result of a history of microeconomic behaviors that led to emergent macro structures and adaptive behavior of the stakeholders involved, e.g. public and private actors. As a result, the combination of local, regional, national and international flows complicate the delineation of the network's borders. Given these characteristics, it is fair to approach such network as a complex system (Martin & Sunley, 2007). An appropriate way to investigate such a network of non-deterministic nature is by using community detection algorithms.

Networks can be described as a series of nodes that are interconnected by links, the strength of a link between two nodes is called the weight. In random networks, nodes and links are homogeneously distributed. Observed networks are almost never random but are characterized by groups of nodes with higher link densities. These groups of nodes are here referred to as 'communities' and are important entities in the complex network (Fortunato, 2010; Newman, 2003). Their influence on the behavior and structure of the encapsulating network has already attracted a fair amount of academic attention. As such, community detection methods are often used to understand and visualize the organization of such complex networks (Lancichinetti & Fortunato, 2009; Newman & Girvan, 2004; Porter et al., 2009). One method that provides good results for large networks is the approach proposed by Blondel et al. in (2008). Their algorithm – the 'Louvain method' – groups the nodes by optimizing the modularity gain. Due to its strong performance, the method is already widely used in academic research, though mainly in engineering and natural sciences. Examples include Lancichineti et al. (2011) who simplified a social network dataset with nearly 20 million nodes and 300 million edges, Blondel et al. (2010) who applied the method on telephone data, Tranos et al. (2015) on migration networks and Croitoru et al. (2014) on social network groups.

When using the Louvain method, one has to note that the results depend on the order of the input data. This means that several iterations have to be run before robust results can be presented. For all our analyses, 100 iterations were applied. To visualize the summary of the hundred runs, each node is classified in its major community. Areas corresponding to nodes with a membership value of only 50-75% are indicated by dashes on the maps.

The Louvain method only groups nodes that are strongly linked. In order to obtain a deeper insight into the role of individual nodes in the network, the within-module degree z and the participation coefficient P, defined by Guimèra and Amaral (2005), are computed. The within-module degree z estimates how well connected a node is to the other nodes in its community. A high z value indicates that the node is well embedded in its cluster, i.e. it contains many internal links. The participation coefficient P, in turn, indicates how well-distributed the node's links are across the other communities. Values close to 1 mean that the node's links are uniformly distributed among the other clusters, a value close to 0 indicates that the node's links are almost all within the cluster. The combination of both parameters determine the role of the node in the network, seven types of regions (Table 1) are defined by Guimèra and Amaral (2005). Ultra-peripheral and peripheral nodes lie at the periphery of the network, with few connections to other nodes. Non-hub connector nodes are already better embedded in the overall network, however, their relative low z-degree indicates a low importance in the local community. Next non-hub kinless nodes are better connected with nodes outside their community then with nodes within their own community, this can point to a wrong classification of the nodes. Provincial hubs are characterized by a very few external connections combined with a high within-module degree z (i.e very well connected internally). Connector hubs are important nodes in a network, tying nodes in a group together and connecting the group to other communities Finally kinless hubs are almost equally connected to the whole network and thus do not enforce local communities.

Table 1. Parameters of node connectivity in a network (Guimerà & Amaral, 2005)

Node role	*Within-module degree z*	*Participation coefficient P*
Ultra-peripheral nodes	<2.5	<0.05
Peripheral nodes	<2.5	0.05 < P < 0.62
Non-hub connector nodes	<2.5	0.62 < P < 0.8
Non-hub kinless nodes	<2.5	>0.8
Provincial hubs	>2.5	<0.3
Connector hubs	>2.5	0.3 < P < 0.75
Kinless hubs	>2.5	>0.75

In this paper the Louvain method is used to detect communities within the buyer-supplier network of logistics firms in Belgium. Subsequently the calculation for all nodes of the within-module degree and the participation coefficient allows the mapping of the node's role in order to grasp the geography of the Belgian logistics buyer-supplier network. Finally the results of the measurement of co-location and of inter-firm relations are combined in a typology of logistics clusters in Belgium

4. Logistics clusters in Belgium

4.1 Data: employment statistics and buyer-supplier linkages of Belgian logistics companies

The case study is carried out with data for Belgium. Because of the country's location nearby the main European markets and good multimodal accessibility, the logistics economic sector is important in Belgium. In 2016, Belgium ranked sixth in the Worldbank's Logistic Performance Indicator, only behind neighboring countries Germany, Luxembourg and The Netherlands, Sweden and Singapore. Furthermore, the northern part of the country is one of the 44 global city logistics regions that handled half of the worldwide air freight and two thirds of total sea freight in 2006 (O'Connor, 2010). Internally, the direct share of the logistics sector to Belgian GDP is almost 8% (Lagneaux, 2008; Mathys & Van Kerckhoven, 2012; The World Bank, 2017).

This paper uses the classification defined by the National Bank of Belgium to identify logistics firms. These firms are not only transportation companies but concern also storage and warehousing and other supporting transport and logistics activities (Lagneaux, 2008). Two different datasets are used to measure geographical clustering of Belgian logistics. First, in Section 4.2, we calculate the co-location in the logistics sector using the employment in full time equivalents (FTE) per municipality in 2010 provided by the Balanscentrale (Central Balance Sheet Office) of the National Bank of Belgium (NBB) (https://www.nbb.be/en/central-balance-sheet-office). Second, in Section 4.3, a dataset is provided by the NBB of more than 800.000 buyer-supplier linkages in the logistics sector in 2011 to analyze the inter-firm relations. It concerns micro-economic data derived from the invoices between companies where either the buyer and/or the supplier is a logistics company. For reasons of anonymity the linkages are aggregated at the zip code level, of which there are 1192 in Belgium, which allows to conduct a detailed geographical analysis. Some zip codes have no interaction, while the highest frequency is observed within Antwerp with over 10.000 buyer-supplier linkages. Besides the linkages that occur between two logistics companies, e.g. an invoice from a transport company to a freight forwarder, the wholesale and retail sector is the largest provider and receiver of services and goods from the logistics sector. Most of the service providers are related to the sales or maintenance of vehicles. Most frequent customers are farmers, construction companies and machinery vendors.

The use of these heterogeneous buyer-suppliers linkages to measure agglomeration advantages is not straightforward. First, the inclusion of non-logistics buyers and suppliers may blur the image of a logistics cluster. Yet, an economic cluster is not only characterized by a high concentration of activities within the same sector, but also encompasses the presence of related industries. As such, the presence of vehicle services, which are the most prominent non-logistics suppliers in the dataset, near a concentration of logistics activities is a prime example of an agglomeration advantage. Second, despite the significant amount of non-logistics companies in the inter-firm linkages, the majority of the flows do consist of logistics-logistics linkages, ensuring their principal role within the analysis. Because of this reason, a high geographical density of the heterogeneous buyer-supplier linkages is assumed to represent agglomeration advantages and can result in clustering.

The dataset contains only buyer-supplier linkages with their origin and destination within the Belgian territory. This might seem a limitation, but as our main interest is to identify the logistics clusters within Belgium the data are sufficient. We are conscious that considering the international linkages could have generated in some places different patterns. This would be the case in places with many international exchanges, such as the big cities and major (air)ports. Another drawback of the data is the allocation of all linkages of a firm to the zip code of the headquarter, this leads to an overestimation in especially urban areas where generally more head offices are located (this is also the case for the employment data). But as nearly 170.000 unique firms in total, of which more than 83.000 unique logistics companies are involved, we can observe that the logistics sector in Belgium still has a lot of small and medium sized companies which have their headquarters on the location of the actual activities. Therefore the data seem appropriate to calculate co-location and inter-firm relations.

4.2 Global indices of logistics employment

The locational Gini coefficient is reported in Table 2 and compared to two studies in the same field.

Table 2: Locational Gini coefficients

Region	*Locational Gini coefficient*	*Spatial unit*	*Average spatial unit size (km²)*	*Size of study area (km²)*
Belgium	0.2991	Municipality	52	30,528
Province North Brabant (NL)	0.2984	Zip code	15	4,919
Ile-de-France and surroundings (F)	0.3797	Municipality	10	77,976

Source: own calculation and van de Heuvel et al. (2014a) and Guillain and Le Gallo (2010)

With a value of almost 0.3, the Gini coefficient gives a first indication of the concentration of logistics employment in Belgium. This result is comparable with the results found in the province of North Brabant and – to a lesser extent – in Paris and surroundings. However, these comparisons have to be interpreted with caution since there exist spatial differences among the three datasets. First, while Belgium is six times larger than North Brabant, its size is only 40% the size of Ile-de-France and its surroundings. In addition, the average spatial unit size is much larger in Belgium than in both other studies. In the case of local concentrations, a larger size of the units will probably lead to an underestimation of the locational Gini coefficient. It is thus possible to assume that the first index of agglomeration indicates at least a similar level of concentration in the logistics sector in Belgium.

The different Moran's I calculations are given in Table 3, these were calculated on the employment density per municipality to partially correct for the effects of the delineation of municipalities. With an average value of 0.4, the Moran I supports the idea of agglomerations of municipalities with higher densities of logistics employment or co-location of logistics. In addition, the values show that the wider the spatial context, the lower the spatial agglomeration, this is an extra indication of local higher densities of employment.

Table 3: Moran's I for different spatial contexts

Spatial context	*Moran's I*	*p-value*
contiguity	0.43	<0,001
nn1	0.42	<0,001
nn3	0.42	<0,001
<15km	0.38	<0,001
<30km	0.32	<0,001

nn1 = municipality plus neighbour with longest common border
nn2 = municipality with the 3 neighbours with longest common border

The LISA is then calculated to disaggregate the Moran's I to the spatial units. The municipalities that are statistically relevant ($p<0.05$) are mapped in figure 2. The map shows the logistics employment density in Belgium. One observes mainly colocation in the northern part of Belgium on the axis Antwerpen-Brussels, in the port areas of Gent and Antwerpen, around Kortrijk and Roeselare and near the Albert canal. In the southern part only in Seraing near Liège co-location is identified, most of the southern region is characterised by low employment per km² in the sector.

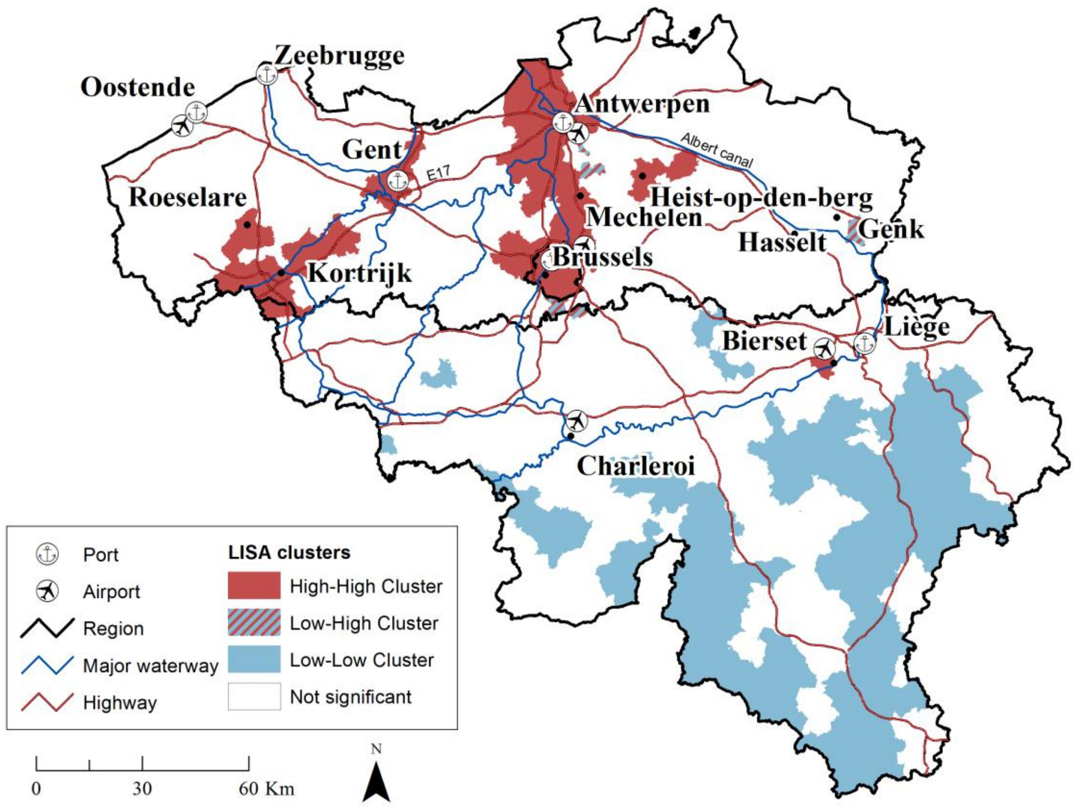


Figure 2: LISA map of employment density in logistics (contiguity spatial relation) (2010)

4.3 Community detection in logistics buyer-supplier linkages

In order to find out if the local concentrations of logistics employment also imply more intense inter-firm relations, the buyer-supplier network is now analyzed. A community detection analysis is used to analyze the frequency of logistics buyer-supplier flows between zip codes, which is the aggregation of the number of linkages between logistics firms of which at least one belongs to the logistics sector. Figure 3 visualizes the amount of buyer-supplier linkages per zip code, the map shows a higher density of flows in the northern region between Brussels, Hasselt/Genk, Antwerpen and Kortrijk, including links to Liège.

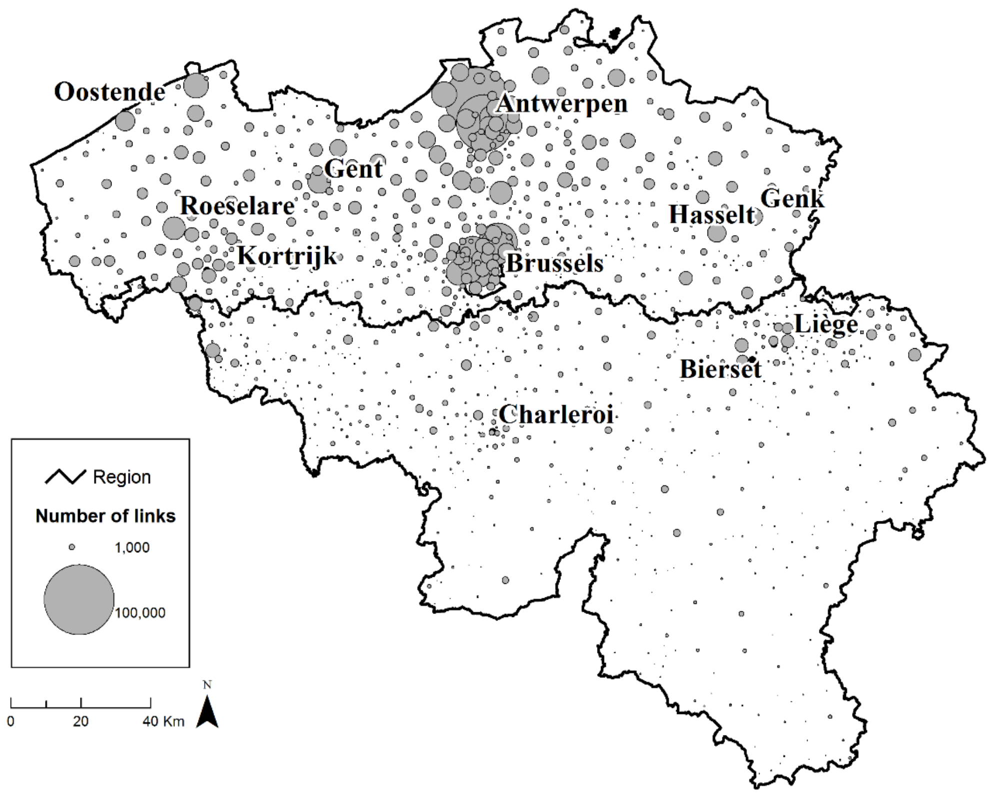


Figure 3: Total buyer-supplier linkages per zip code

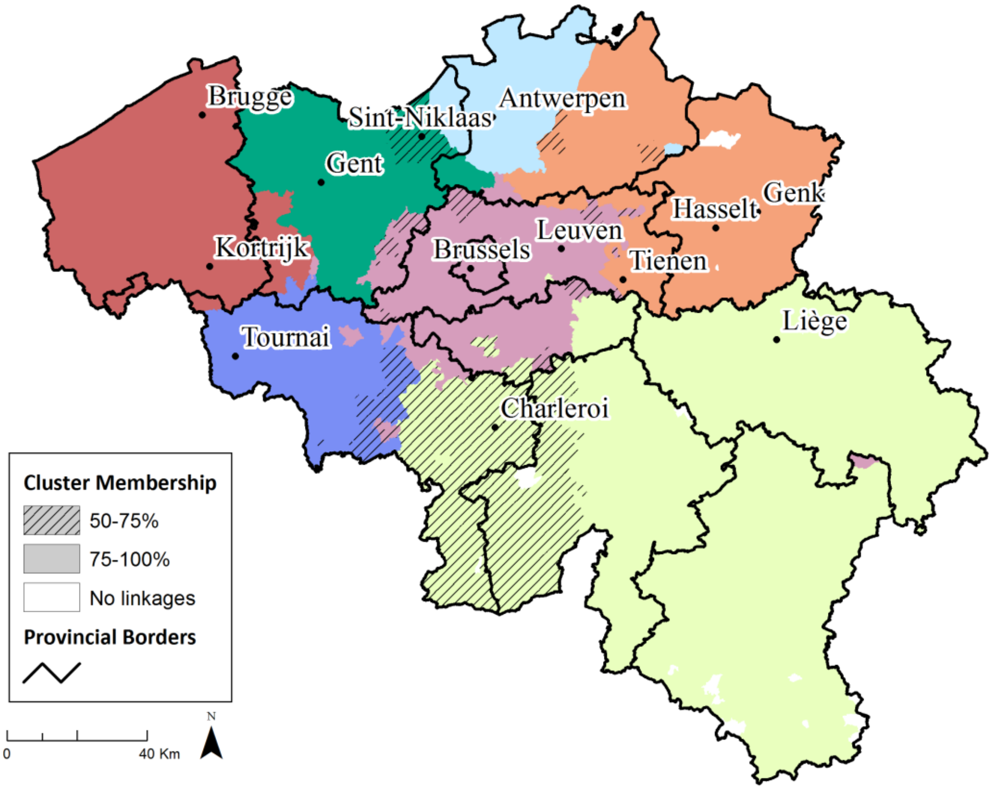


Figure 4: Communities in the buyer-supplier network

Figure 4 maps the communities resulting from the runs of the Louvain-method on the logistics buyer-supplier network, optimizing the modularity of the overall network. In total seven major communities are found in Belgium, with an unexpected spatial contiguity (geography/proximity still matters!). Communities are delineated around the capital city Brussels in the central part, around the major cities Antwerpen and Gent that both are port-cities, in the east where the important Albert canal links Antwerpen with Liège, in the province in the west including Kortrijk and the port of Zeebrugge, around Tournai, and finally a stretched out community in the southern part. The broad area around Charleroi is dashed, indicating there are alternative connections, mainly to the Brussels community. The community detection algorithm provides a delineation of the different communities, which include logistics clusters with their hinterland. The main cities and (air)ports have a major impact on the spatial structure of the communities. The community detection algorithm proves to be a useful tool for analyzing big data on complex networks since the high amount of linkages made it very hard to detect geographical patterns when analyzing the raw data in a traditional explorative way.

What is lacking is a good understanding of the roles of the different nodes in the various communities. In the final step we therefore use the coefficients defined by Guimerà and Amaral (2005) to analyze the local and overall connectedness of the different nodes in the network. In figure 5 each node is classified according to the z-P parameter space using Table 1. Attachment 1 shows details on the parameters for the nodes classified as connector hub or kinless hub. Especially the connector hubs can be defined as geographical clusters as they tie the nodes together in the community and they make the connection to other communities. The kinless hubs are equally tied to the whole network and as such their role is less geographically localized. Central in the 'Brussels community' four zip codes are kinless hubs including the area around the national airport, the presence of many headquarters explains the many ties with the whole network; two areas in Brussels are connector hubs, these are known as main industrial regions, apparently a logistics cluster is linked to it. Other kinless hubs are found in Gent and Tournai. The Antwerp port area clearly is a logistics cluster tying together the surrounding community and connecting to the rest of the network with an impressive amount of linkages. The ports of Zeebrugge and Oostende next to Roeselare are the clusters of the community in the west. In the northeast, Hasselt and Genk comprise the logistics cluster, next to two more smaller clusters. Also in and around Liege a series of connector hubs can be found, they are linked to different industrial areas and transport terminals spread over the region. Further most of the nodes have z-vales under 2.5 combined with an average participation coefficient around 0.6, resulting in mostly peripheral and non-hub connector nodes. These nodes belong to the hinterland within the community, but are only loosely connected to the hubs.

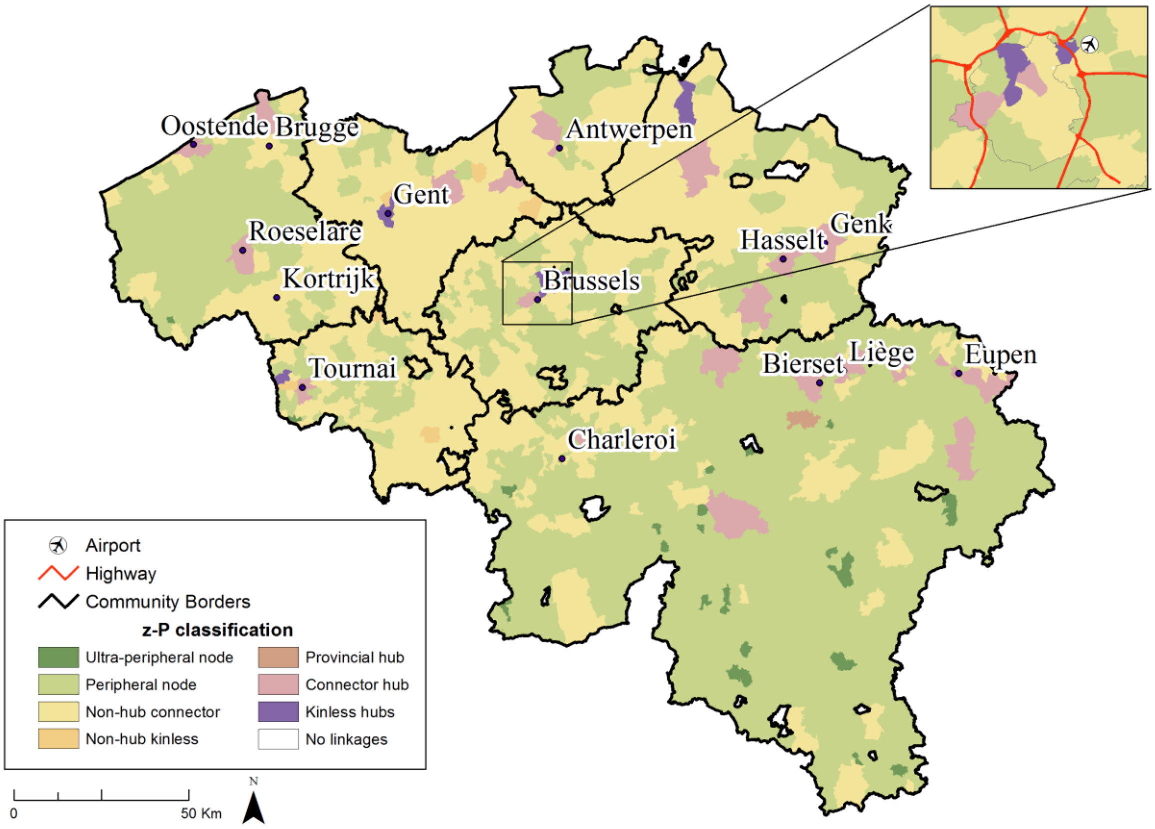
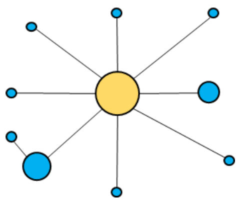
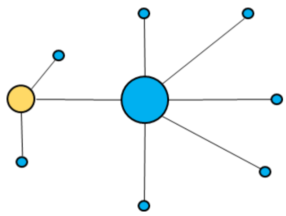
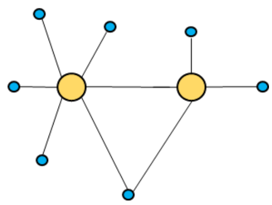


Figure 5: z-P parameter space of logistics communities (Fig 4)

5. Discussion

The combination of the various indicators provides a synthesis of the complex network of logistics buyer-supplier relations within Belgium. The analysis of the employment values and buyer-supplier linkages prove that logistics activities are not just scattered over the country, we detect spatial densities of both employment and inter-firm relations. Despite its labelling as 'dispersing across regions' (Porter, 2003), there remain geographical patterns in the structure of the logistics network (Hesse & Rodrigue, 2004). By using an approach similar to some authors who measure concentration by agglomeration indices (e.g. Guillain & le Gallo, 2010), the results of the LISA would be used to indicate logistics clusters. In that case, places like Kortrijk and Mechelen would be classified as logistics clusters, however these places have a within-module degree under 2.5, indicating a low importance within their community. The combination of traditional agglomeration indices and parameters from complex network analysis (Attachment 2), provides the opportunity to create a typology of logistics concentrations.

Table 4: Typology of logistics concentration

Indicators	*Cluster*	*Cluster spillover*	*Polycentric cluster*
LISA	High-High	High-High	High-High/N.S.
Within-module degree z	>2.5	<=2.5	~2.5
Participation coefficient P	>0.3	>0.3	>0.3
Network structure	Hub and spoke	Linked to cluster(s)	Multi-point
Network configuration			
Examples	Antwerpen, Brussels, Gent, Roeselare	Mechelen, Kortrijk	Hasselt-Genk Liège-Bierset Oostende-Zeebrugge

The nodes identified as clusters in Table 4 combine employment densities with tight buyer-supplier linkages. These clusters are the logistics clusters that are traditionally identified in studies by Rivera et al. (2014), Sheffi (2012) and van den Heuvel (2014a). They are characterized by significant higher concentrations of employment (high-high), strong connections to other nodes in their community (z-score > 2.5) together with a decent amount of links to other communities (P-value > 0.3). Topologically this results in a hub and spoke network with a center of strong employment and many local (i.e. within-community) flows, but well connected to the surrounding areas as well. Examples in Belgium are parts of the major cities Brussels, Antwerp, Gent but also of the regional city of Roeselare, who serve as connector hubs for their regions in addition to the higher densities of logistics employment.

The economic strength of these clusters may lead to large inequalities between the core and peripheral regions, but it can also yield spill-over clusters. These are local concentrations of logistics activities, but with weaker local relations indicated by lower z-values. These spill-over clusters are less important because they are overshadowed by and focused on the nearby big cluster. However, they probably do take advantage of the growing congestion within the big cluster. Examples in Belgium are the region around Mechelen (mainly linked to Brussels) and Kortrijk (linked to Roeselare).

A third type of logistics clusters are the regions with several nearby concentrations, we call them polycentric clusters. Some of those regions may have significant concentrations of employment, but since the activities are spread over multiple locations the LISA can be not significant. However the concentration of logistics companies in those multiple locations serve via their buyer-supplier linkages as connector hubs for the region. Examples in Belgium of these polycentric clusters can be found in Hasselt-Genk, Liège-Bierset and Oostende-Zeebrugge.

While the logistics clusters in Antwerpen, Brussels, Gent and Roeselare would also have been identified using a traditional approach of only agglomeration indices, the methodology would have omitted or overrepresented other clusters. The polycentric clusters would not have been identified, the spill-over clusters like Mechelen and Kortrijk would have been classified like Brussels and Antwerpen. Remarkably, notwithstanding our quantitative geographical approach compared to the qualitative business approach from Markusen (1996), the clusters and spill-over clusters identified in Table 4 resemble very well her hub and spoke and satellite districts. This shows that her business approach can be verified quantitatively.

6. Conclusions

The outcome of this paper is a typology of logistics clusters resulting from the combination of both agglomeration indices and tools inspired by complex network analysis. The case of logistics in Belgium was studied through both quite accessible employment data and an innovative dataset of buyer-supplier linkages between logistics firms within the country. This provides a unique opportunity to analyze inter-firm relationships in combination with the already well documented co-location approach in order to detect geographical clusters. We demonstrate the value of community detection techniques when working with big relational data in economic geography, mapping the outcomes of community detection provides a valuable tool for analyzing inter-firm linkages. The combination of studying both co-location and inter-firm linkages should give the fuzzy term cluster more depth in future research.

Logistics activities within Belgium are concentrated in some places and they are less relevant in others. Three different types of logistics concentrations are identified, providing a new typology of the logistics sector in Belgium. Concentrations of logistics employment that serve as the center of a large region are identified as clusters as discussed by Sheffi (2012). Related to these centers, spill-over clusters appear in the nearby hinterland. While they are also characterized by higher concentrations of employment, their relationships are mostly directed to a big cluster nearby. Finally, polycentric clusters act as a third type of logistics concentration, although a less prominent co-location, they do play an important in tying together the firms in the region and connecting them to the overall network. This cluster typology could be helpful in smart specialization strategies in order to detect possible regions for endogenous growth of the logistics sector. Nowadays in Belgium almost all regions apply for investments in logistics, we can expect more agglomeration advantages in the identified clusters, of course depending on the scale of the cluster.

This paper demonstrates how inter-firm linkages can be studied in a quantitative way. Adding this information to the result of co-location analysis brings us closer to the detection of industrial clusters. Yet, though the inter-firm linkages together with co-location provide evidence of the presence of Marshallian externalities, little is known of the quality and sustainability of the identified clusters. No differentiation has been made amongst job categories, neither between different inter-industry linkages. This remains a limitation, especially from a cluster life cycle assessment perspective. The challenge therefore remains to link the observed cluster effects to specialization and knowledge spill-overs, which should be the subject of further research. In addition, it would be interesting to apply the proposed methodology to new cases in other economic sectors and regions and to include international relationships to assess their impact on the spatial structure

Acknowledgements

This research was supported by the Flemish Government via the Urban Logistics and Mobility project (ULM) financed by IWT-SBO. We are grateful to the National Bank of Belgium for providing data.

References

Akyelken, N., & Keller, H. (2014). Framing the Nexus of Globalisation, Logistics and Manufacturing in Europe. *Transport Reviews*, 34(6), 674–690.

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis*, 27(2).

Arbia, G. (2001). The role of spatial effects in the empirical analysis of regional concentration. *Journal of Geographical Systems*, 3, 271–281.

Balland, P. A., De Vaan, M., & Boschma, R. (2012). The dynamics of interfirm networks along the industry life cycle: The case of the global video game industry, 1987-2007. *Journal of Economic Geography*, 13(5).

Barrios, S., Bertinelli, L., Strobl, E., & Teixeira, A. C. (2009). Spatial Distribution of Manufacturing Activity and its Determinants: A Comparison of Three Small European Countries. *Regional Studies*, 43(5), 721–738.

Bathelt, H., Malmberg, A., & Maskell, P. (2004). Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography*, 28(1), 31–56.

Bell, G. G. (2005). Clusters, networks, and firm innovativeness. *Strategic Management Journal*, 26(3), 287–295.

Benneworth, P., Danson, M., Raines, P., & Whittam, G. (2003). Confusing Clusters? Making Sense of the Cluster Approach in Theory and Practice. *European Planning Studies*, 11(5), 511–520.

Bertinelli, L., & Decrop, J. (2005). Geographical Agglomeration: Ellison and Glaeser's Index Applied to the Case of Belgian Manufacturing Industry. *Regional Studies*, 39(5), 567–583.

Blondel, V., Guillaume, J., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of community hierarchies in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 10, 1–6.

Blondel, V., Krings, G., & Thomas, I. (2010). Regions and borders of mobile telephony in Belgium and in the Brussels metropolitan zone. *Brussels Studies*, 42.

Bresnahan, T. F., & Gambardella, A. (2004). B*uilding High-Tech Clusters: Silicon Valley and Beyond*. Cambridge University Press.

Croitoru, A., Wayant, N., Crooks, A., Radzikowski, J., & Stefanidis, A. (2014). Linking cyber and physical spaces through community detection and clustering in social media feeds. *Computers, Environment and Urban Systems*, 53, 47–64.

CSCMP (Council Of Supply Chain Management Professionals). (2016). Supply Chain Management Definitions. Retrieved April 6, 2016, from https://cscmp.org/supply-chain-management-definitions.

Delgado, M., Porter, M. E., & Stern, S. (2010). Clusters and entrepreneurship. *Journal of Economic Geography*, 10(4), 495–518.

Delgado, M., Porter, M. E., & Stern, S. (2016). Defining clusters of related industries. *Journal of Economic Geography*, 16(1), 1–38.

Ding, Y., & Fotheringham, A. S. (1992). The integration of spatial analysis and GIS. *Computers, Environment and Urban Systems*, 16(1), 3–19.

Duranton, G., & Overman, H. G. (2005). Testing for Localization Using Testing Data Micro-Geographic. *Review of Economic Studies*, 72(4), 1077–1106.

Duschl, M., Scholl, T., Brenner, T., Luxen, D., & Raschke, F. (2015). Industry-Specific Firm Growth and Agglomeration. *Regional Studies*, 49(11), 1822–1839.

Ellison, G., & Glaeser, E. L. (1997). Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy*, 105(5), 889.

Ellison, G., Glaeser, E. L., & Kerr, W. R. (2010). What Causes Industry Agglomeration ? Evidence from Coagglomeration Patterns. *American Economic Association*, 100, 1195–1213.

Flämig, H., & Hesse, M. (2011). Placing dryports. Port regionalization as a planning challenge - The case of Hamburg, Germany, and the Süderelbe. *Research in Transportation Economics*, 33(1), 42–50.

Fortunato, S. (2010). Community detection in graphs. *Physics Report*, 75(174), 103.

Giuliani, E. (2010). Clusters, networks and economic development: an evolutionary economics perspective. In R. Boschma & R. Martin (Eds.), *The Handbook of Evolutionary Economic Geography*. Cheltenham: UK: Edward Elgar.

Guillain, R., & le Gallo, J. (2010). Agglomeration and dispersion of economic activities in and around Paris: An exploratory spatial data analysis. *Environment and Planning B: Planning and Design*, 37(6), 961–981.

Guimerà, R., & Amaral, L. A. N. (2005). Functional cartography of complex metabolic networks. *Nature*, 433(7028), 895–900.

Hesse, M. (2008). *The city as a terminal: The urban context of logistics and freight transport*. Burlington, VT: Ashgate Publishing Company.

Hesse, M. (2015). Selling the region as hub. In J. Cidell & D. Prytherch (Eds.), *Transport, Mobility, and the Production of Urban Space*. Routledge.

Hesse, M., & Rodrigue, J. P. (2004). The transport geography of logistics and freight distribution. *Journal of Transport Geography*, 12, 171–184.

Hoffmann, J., Hirsch, S., & Simons, J. (2015). Identification of spatial agglomerations in the German food processing industry. *Papers in Regional Science*, 96(1), 139-162

Krugman. (1991). *Geography and Trade*. MIT press.

Lagneaux, F. (2008). Economic Importance of Belgian Transport Logistics. *National Bank of Belgium Working Paper.*

Lancichinetti, A., & Fortunato, S. (2009). Community detection algorithms: a comparative analysis. Physical Review E, 80.

Lancichinetti, A., Radicchi, F., Ramasco, J. J., & Fortunato, S. (2011). Finding statistically significant communities in networks. *PloS One*, 6(4).

Levy, R., & Talbot, D. (2015). Control by proximity: Evidence from the "Aerospace Valley" Competitiveness Cluster. *Regional Studies*, 49(6), 955–972.

Lundequist, P., & Power, D. (2002). Putting Porter into Practice? Practices of Regional Cluster Building: Evidence from Sweden. *European Planning Studies*, 10(6), 685–704.

Markusen, A. (1996). Sticky Places in Slippery Space : A Typology of Industrial Districts. *Economic Geography*, 72(3), 293–313.

Martin, R., & Sunley, P. (2003). Deconstructing clusters: chaotic concept or policy panacea? *Journal of Economic Geography*, 3(1), 5–35.

Martin, R., & Sunley, P. (2007). Complexity thinking and evolutionary economic geography. *Journal of Economic Geography*, 7(5), 573–601.

Mathys, C., & Van Kerckhoven, M. (2012). Economic Importance of Belgian Transport Logistics. *National Bank of Belgium Working Paper,* Brussels.

Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review*, 45(2), 167–256.

Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*, 69(2).

O'Connor, K. (2010). Global city regions and the location of logistics activity. *Journal of Transport Geography*, 18(3), 354–362.

Porter, M. A., Onnela, J., & Mucha, P. J. (2009). Communities in networks. *Notices of the AMS*, 56(9), 1082–1097.

Porter, M. E. (1998). Clusters and the new economics of competition. *Harvard Business Review*, 76(6), 77–90.

Porter, M. E. (2003). The Economic Performance of Regions. *Regional Studies*, 37(6-7), 545–546.

Riguelle, F., Thomas, I., & Verhetsel, A. (2007). Measuring urban polycentrism: A European case study and its implications. *Journal of Economic Geography*, 7(2), 193–215.

Rivera, L., & Sheffi, Y. (2016). The benefits of logistics clustering. *International Journal of Physical Distribution & Logistics Management*, 46(3), 242–268.

Rivera, L., Sheffi, Y., & Welsch, R. (2014). Logistics agglomeration in the US. *Transportation Research Part A: Policy and Practice*, 59, 222–238.

Rodrigue, J. P., & Notteboom, T. (2010). Comparative North American and European gateway logistics: The regionalism of freight distribution. *Journal of Transport Geography*, 18(4), 497–507.

Sheffi, Y. (2012). *Logistics Clusters*. Cambridge, MA: MIT Press.

The World Bank. (2017). Logistics Performance Index - LPI Results 2016.

Tranos, E., Gheasi, M., & Nijkamp, P. (2015). International migration: a global complex network. *Environment and Planning B: Planning and Design*, 42(1), 4–22.

van den Heuvel, F., de Langen, P. W., van Donselaar, K. H., & Fransoo, J. C. (2012). Identification of Employment Concentration Areas. *European Planning Studies*, 22(1), 204–226.

van den Heuvel, F., de Langen, P. W., van Donselaar, K. H., & Fransoo, J. C. (2014). Proximity matters: synergies through co-location of logistics establishments. *International Journal of Logistics Research and Applications: A Leading Journal of Supply Chain Management*, 17(5), 377-395.

van den Heuvel, F., Rivera, L., van Donselaar, K. H., de Jong, A., Sheffi, Y., de Langen, P. W., & Fransoo, J. C. (2014). Relationship between freight accessibility and logistics employment in US counties. *Transportation Research Part A: Policy and Practice*, 59(2012), 91–105.

van den Heuvel, F., van Donselaar, K., de Langen, P., & Fransoo, J. (2016). Co-Location Synergies: Specialised Versus Diverse Logistics Concentration Areas. *Tijdschrift Voor Economische En Sociale Geografie*, 107(3), 331-346.

Vanthillo, T., Verhetsel, A., & Vanelslander, T. (2014). Subregionaal speerpuntbeleid in Vlaanderen. Steunpunt Ondernemen en Regionale Economie.

Verhetsel, A., Kessels, R., Goos, P., Zijlstra, T., Blomme, N., & Cant, J. (2015). Location of logistics companies: a stated preference study to disentangle the impact of accessibility. *Journal of Transport Geography*, 42, 110–121.

Appendix A: Zip codes identified as connector or kinless hub

ZIP CODES	*z-P* CLASSIFICATION	TOTAL LINKAGES
ANTWERPEN (2030 RIGHT BANK)	Connector hub	86,603
ANTWERPEN (2000 CENTRE)	Connector hub	68,025
BRUSSELS (1000 CENTRE)	Kinless hubs	39,609
ZAVENTEM (NATIONAL AIRPORT)	Kinless hubs	35,077
DIEGEM (NATIONAL AIRPORT)	Kinless hubs	29,006
BRUSSELS (1020 LAKEN)	Kinless hubs	19,519
BRUSSELS (1070 ANDERLECHT)	Connector hub	17,144
ZEEBRUGGE	Connector hub	15,655
GENT	Kinless hubs	14,156
ROESELARE	Connector hub	13,452
BRUSSELS (1030 SCHAARBEEK)	Connector hub	10,904
OOSTENDE	Connector hub	10,830
HASSELT	Connector hub	9,913
GENK	Connector hub	8,918
TEMSE	Connector hub	8,308
TURNHOUT	Kinless hubs	8,185
LOKEREN	Connector hub	7,475
GEEL	Connector hub	6,600
TOURNAI (7522 WEST)	Kinless hubs	6,160
SINT-TRUIDEN	Connector hub	5,587
LIÈGE (4460 BIERSET AIRPORT)	Connector hub	5,521
LIÈGE(4020 NORTH)	Connector hub	5,267
EUPEN	Connector hub	4,937
LIÈGE (4400 FLEMALLE)	Connector hub	4,111
LIÈGE (4040 HERSTAL)	Connector hub	3,826
WELKENRAAT	Connector hub	3,256
LIÈGE (4000 CENTRE)	Connector hub	2,978
CHARLEROI	Connector hub	2,941
HERVÉ	Connector hub	2,437
HANNUT	Connector hub	2,297
TOURNAI (7500 CENTRE)	Connector hub	2,290
CINEY	Connector hub	2,014
MALMEDY	Connector hub	1,851

Appendix B: Characteristics of main nodes

ZIP CODES	LISA	LINK FREQUENCY	*z*-VALUE	*P*-COEFFICIENT	*Z-P* *CLASSIFICATION*	*TYPOLOGY*
ANTWERPEN	HH	>50,000	>2.5	0.56-0.69	Connector hub	Cluster
MECHELEN	HH	10-15,000	<2.5	0.77	Non-hub connector	Spill-over cluster
BRUSSEL	HH	>50,000	>2.5	0.50-0.79	Connector hub	Cluster
GENT	HH	10-15,000	>2.5	0.77	Kinless hub	Cluster
KORTRIJK	HH	5-10,000	<2.5	0.67	Non-hub connector	Spill-over cluster
ROESELARE	HH	10-15,000	>2.5	0.66	Connector hub	Cluster
HEIST-OP-DEN-BERG	HH	5-10,000	<2.5	0.76	Non-hub connector	Spill-over cluster
BIERSET	HH	<5000	>2.5	0.60	Connector hub	Polycentric cluster
GENK	N.S.	5-10,000	>2.5	0.68	Connector hub	Polycentric cluster
OOSTENDE	N.S.	10-15,000	>=2.5	0.74	Connector hub	Polycentric cluster
ZEEBRUGGE	N.S.	15-20,000	>2.5	0.74	Connector hub	Polycentric cluster
HASSELT	N.S.	5-10,000	>2.5	0.65	Connector hub	Polycentric cluster
LIEGE	N.S.	5-10,000	>2.5	0.60	Connector hub	Polycentric cluster
TOURNAI	N.S.	<5000	>2.5	0.60	Connector hub	-
EUPEN	N.S.	<5000	>2.5	0.70	Connector hub	-
GEEL	N.S.	5-10,000	>2.5	0.68	Connector hub	-
WESTERLO	HH	<5000	<2.5	0.70	Non-hub connector	-
LOKEREN	N.S.	5-10,000	>2.5	0.68	Connector hub	-
TEMSE	HH	5-10,000	>=2.5	0.74	Connector hub	Spill-over cluster
SINT-TRUIDEN	N.S.	5-10,000	>=2.5	0.62	Connector hub	-
HANNUT	N.S.	<5000	>=2.5	0.64	Connector hub	-
CINEY	LL	<5000	>2.5	0.53	Connector hub	-
FLEURUS	N.S.	<5000	>2.5	0.67	Connector hub	-
MALMEDY	N.S.	<5000	>2.5	0.40	Connector hub	-
BRUGGE	N.S.	5-10,000	<=2.5	0.24	Non-hub connector	-