Maastricht University

UNITED NATIONS UNIVERSITY

UNU-MERIT

Working Paper Series

#2020-005

Mapping industrial patterns and structural change in exports

Charlotte Guillard

Maastricht Economic and social Research institute on Innovation and Technology (UNU-MERIT)
email: **info@merit.unu.edu** | website: **http://www.merit.unu.edu**

Boschstraat 24, 6211 AX Maastricht, The Netherlands
Tel: (31) (43) 388 44 00

UNU-MERIT Working Papers
ISSN 1871-9872

Maastricht Economic and social Research Institute on Innovation and Technology
UNU-MERIT

UNU-MERIT Working Papers intend to disseminate preliminary results of research carried out at UNU-MERIT to stimulate discussion on the issues raised.

Mapping industrial patterns and structural change in exports*

Charlotte Guillard†

January 20, 2020

Abstract

This paper proposes a new methodology for identifying patterns in the organisation of industries and their evolution over time, based on the temporal network structure of the product space. To do this, I apply a community detection algorithm on 5-year snapshots of the product space from 1975 to 2000. This exercise enables us to identify different clusters of related products and to follow their evolution over time. I find that the product space is highly modular, that is it contains well delimited clusters of products. The community structure and its evolution show that the factors explaining industrial patterns and structural change are more complex than the traditional divide between low, medium and high-tech industries. Several common drivers can be identified to explain the emergence and evolution of different communities including the experience in a technological domain, factor abundance, scale economies as well as global value chains and vertical integration. Moreover, I find that technological domains and boundaries between industries are not always clear-cut and can evolve over time.

Keywords: Structural change, Capabilities, Economic Complexity, Networks, Community Structure, Exports.

JEL codes: O11, O14, O33, O25, P40, E14.

---

*I am grateful to Robin Cowan and Pierre Mohnen for their useful comments.

†Institute for Innovation and Public Purpose, UCL, BETA, Université de Strasbourg and UNU-MERIT.

Email: guillard@merit.unu.edu

1 Introduction

In a globalised environment, export patterns and their evolution play a central role in countries' economic development. What you export matters, not only because some products are associated with higher productivity levels than others but also because they open the way to greater opportunities. Countries differ in the set of capabilities they have and products differ in the set of capabilities they require to be made. Assuming that the set of products a country has a comparative advantage in defines well what it can produce, looking at export data gives an indication of the set of capabilities a country has through the types of products it exports. Furthermore, one way in which products are related is through the type of capabilities they share. Because knowledge is cumulative, countries build on existing knowledge and capabilities and economic development is strongly path dependent (Weitzman, 1998, Arthur, 2009). It is therefore easier to start exporting new products that have capabilities in common with products already being produced in the economy so that fewer new capabilities have to be mastered. I rely on the product space methodology to measure product relatedness in terms of production capabilities (Hidalgo et al., 2007). The product space corresponds to a network of products where, roughly speaking, two products are linked if they are both exported by the same country. A country's "location" in the product space is defined by the set of products it exports. Structural change, then, can be seen as a movement through the product space. In principle, a country can move from anywhere to anywhere in the space. In practice, however, movement to nearby products will be simpler because of the way products are related to each other in the product space through the capabilities they require.

Existing literature suggests that products tend to cluster around a specific set of capabilities and form industries and sectors. These clusters reflect important differences in terms of knowledge and production processes (Pavitt, 1984, Malerba, 2002, Lall et al., 2004, Arthur, 2009). However, they are not completely isolated as each firm is associated to specific complementary activities and linkages with other firms within and between industries. The presence of industry interlinkages implies that the development of specific industries can act as bottlenecks or opportunities for the industrial prospects of countries. In addition, the boundaries between industries evolves over time for several reasons.

To address the limitations of industry taxonomies (Pavitt, 1984, Leamer, 1984, Lall et al., 2004), I present a new dynamic product classification based on the temporal network structure of the product space. I apply a community detection algorithm (Rosvall and Bergstrom, 2011) on 5-year snapshots of the product space from 1975 to 2000 to find this product classification. This exercise enables us to identify different clusters of related products and to follow their evolution over time. I find that the product space is highly modular, that is it contains well delimited clusters of products. The community structure and its evolution show that the factors explaining industrial patterns and structural change are more complex than the traditional divide between low, medium and high-tech industries. Several common drivers can

be identified to explain the emergence and evolution of different communities including the experience in a technological domain, factor abundance, scale economies and vertical integration. In addition, I find that the community structure is fairly stable over time and tends to stabilise even more from 1990 onwards.

This paper is organised as follows. In the first section, I review the literature on the modular structure of the economy and industry idiosyncrasies. In the second section, I describe the data and the construction of the temporal network. In the third section, I detail the methodological steps to uncover the community structure of the product space. In the fourth section, I analyse the results of the dynamic community structure. Finally, I conclude in the last section.

2 Background and literature review

Relying on an evolutionary framework, Malerba (2002) highlights the importance of accounting for the modular structure of the economy by introducing the concept of sectoral systems of innovation. He argues that the characteristics of knowledge and the nature of the innovation process varies across industries. This explains the fact that firms experience different innovation and productivity dynamics depending on the industry they belong to. Furthermore, by sharing a common set of knowledge bases, firms in the same industry have common learning processes, technologies and production processes (Nelson and Winter, 1982). They also share the same types of complementarities with other knowledge (internal or external to the industry), and similar relationships with other actors. The nature of the learning process also depends on the characteristics of knowledge, which is industry specific. Because the search process relies heavily on the characteristics of the knowledge base of firms, the potential to exploit technological opportunities greatly differs across industries and firms (Nelson and Winter, 1982 and Dosi et al., 1995). This, in turn, affects their potential for technological progress and transformation. Differences across industries are also expressed in the nature and the extent of their linkages with other industries. The existence of *technological complementarities* constitutes a triggering factor to the development of technologies and industries (Hirschman, 1958, Rosenberg, 1976 and Rosenberg, 1979). These complementarities are likely to cause *technological interdependences* between apparently unrelated processes and industries. Technological progress within a technological regime often relies on the improvement of major components via internal replacement and structural deepening (Nelson and Winter, 1982, Arthur, 2009), which are likely to create a *technological imbalance.*

Technological imbalance can originate from the need to improve a product or a production process in the same industry or another (Rosenberg, 1976). Therefore, the presence of complementarities between industries is central to explaining technological and economic development and its direction in some

countries as well as further technological backwardness in others.

Without focusing solely on technological development, the existence of backward and forward linkages between industries provides a stimulus for the development of related economic activities (Hirschman, 1958). On the one hand, domestic production of inputs is accompanied by an active stimulus for their use by local producers, that is they will attempt to create *forward linkages* with potential demand. On the other hand, the development of economic activities depends on the availability of inputs on the domestic market, giving rise to *backward linkages*. The growing role played by global value chains in the last decades shows that these linkages can be distorted. It is therefore important to consider industries' idiosyncrasies in the potential fragmentation of their production processes. Lall et al. (2004) highlights the differences in fragmentation intensity across industries and the important factors: (*i*) high technical divisibility of production processes and components in terms of scale, skills and technological needs; (*ii*) access to low-skilled labour at lower cost; (*iii*) labour intensive production process; (*iv*) simple production tasks; (*v*) high value to weight ratio goods.

Several authors have analysed the structure of industries and products (exports) in order to classify industrial goods into homogenous groups according to their technological content. Pavitt (1984) is one of the first to empirically study innovative firms' differences across industries. Although this classification provides rich information on the differences between firms and industries, only innovative firms are considered. In the present study, I am interested in both innovative and non-innovative firms. The classifications of Leamer (1984) and Lall (2000) cover a wider part of the economy, relying on the measurement product relatedness using export data. However, both of these methods rely on a certain amount of ad-hoc adjustments, in particular in Lall (2000). In addition, the lack of details on the algorithm applied in Leamer (1984) makes this method impossible to reproduce. For these reasons, I propose a new methodology that can be reproduced on any disaggregated export data and does not rely on any ad-hoc adjustments.

Countries differ in the set of capabilities they possess and products differ in the set of capabilities they require to be made. More precisely, the production of a good is the result of a combination of capabilities (Weitzman, 1998, Arthur, 2009). Some products are more complex in that they require a larger and more diverse set of capabilities that only a few countries possess. Products also differ in terms of their technological domain (Arthur, 2009), which can evolve over time. Different products can also share similar capabilities, and in that way, they can be related. Most classical and new trade models consider products as being organised in a continuous sequence in which adjacent ones are fairly similar but the extremes (i.e., very low-tech relative to very high-tech) are very different. The position of products on this continuous sequence depends on their technology level. This representation has several implications. First, products are related according to one dimension only (technology). Second, since the sequence is continuous, there is no discontinuity in the effort of countries to introduce new products to their basket. Third, a country

should add products according to the sequence and cannot start exporting products that are not directly linked to products already in its basket. As a consequence, according to this representation, there is one unique path to development, that is from low to high-tech products. Conversely, one could imagine that products are related through a network and that their links are defined according to many dimensions, not only technology. The product space is the first empirical attempt to represent products' relationships as a network and not as a continuum (Hidalgo et al. (2007)). This network is derived from a network of exports where countries are linked to products for which they have a comparative advantage. It is assumed that the set of products a country has a comparative advantage in defines well what it can produce, and indirectly, the extent of its capabilities. In other words, the export network gives an idea of the set of capabilities a country has through the products it exports. Further, one can assume that if two products are often found together in the export basket of countries, their production is likely to involve common capabilities. Based on this assumption, we can derive a network of products: the product space. In that representation, a given set capabilities is not unique to a product but is shared among several products. The focus here is not on any particular aspect of the similarity between products as capabilities correspond to different dimensions of product relatedness including inputs, markets to which products are exported, their R&D intensity, institutions and infrastructures required for the goods to be exported.

In the next section, I describe the methodology developed to empirically analyse the modular structure of the economic system and its evolution using export data.

3 Methodology

3.1 Network construction and data

Countries have a unique set of capabilities and product are associated to a set capabilities to be produced. Thanks to these relationships between countries and products, we can construct a network of products using export data by following the product space methodology developed in Hidalgo et al. (2007).[1] In this setting, I assume that if a country has a comparative advantage in a product, then

[1]Network theory offers a wide range of methodologies to measure the similarity between elements in a system based on the structure of their interactions. The standard way to do a one-mode projection would be use structural equivalence, i.e. the weight of the link between two products is a count of the number of countries that export both. However, this simple measure of structural equivalence suffers from a bias in favour of commonly exported products. Several measures have been proposed to correct for this bias. I focus on three common measures of similarity for which values range between 0 and 1: cosine similarity, Jaccard index and "proximity index" introduced in Hidalgo et al. (2007). They are highly correlated as shown in Table 3. However, in two specific cases, there are important differences between the measures.

- In the first case, the set of countries exporting product A is a subset of countries exporting product B (exported by 16 countries for example - see Figure 1a). Let's consider a situation in which the number of countries exporting product A increases and any country adding product A to its export basket always export product B. When using the proximity and Jaccard indexes, the similarity between the two products will grow linearly with the set of countries exporting product A. Conversely, the similarity derived from the Cosine index will increase at a decreasing rate when an increasing number of countries export product A. Yet, there is no reason to believe that this relationship should be concave.
- In the second case, the two products are exported by the same number of countries (for instance, 12 countries - see Figure 1b), but not necessarily the same set. When measured with the cosine or the proximity indexes, the similarity

it has the capabilities to produce it and that the set of products it exports then reflects the extent of its capabilities. Now, if products are often exported together then it is reasonable to assume that they require a similar set of capabilities.

To map industrial patterns and structural change, I construct several snapshot of the product space based on export data aggregated by 5-year period from 1975 to 2000 (1975-1979, 1980-1984, 1985-1989, 1990-1994, 1995-1999, 2000).

The networks described above are constructed using gross export data from COMTRADE cleaned by Feenstra et al. (2005). These data are disaggregated at the 4-digits level for the period 1962 to 2000 and include 108 to 134 countries depending on the year. For the sake of comparison across time, I only use data from 1975 to 2000 and the 111 countries that are present throughout the period of analysis. As highlighted by Lall (2000), export performance is key to economic development for several reasons. On the one hand, countries can benefit from economies of scale and specialisation through exports. On the other hand, the increased competitivity of the globalised world makes export performance a good indicator of the productive efficiency of countries in the manufacturing sector. Although it is an imperfect and partial representation of economic activities and their interdependencies across nations, trade values remain the most complete and disaggregated empirical measure in terms of time and space.

I assume that when a country exports a good p, it has the capabilities to produce it. When a large number of countries export two goods p and p', we can infer that the competences underlying their production are related. Identifying pairs of related products is fairly straightforward but when it comes to uncovering a larger set of related products it becomes more complicated. Community detection is used to identify groups of similar products. In network theory, a community consists of a group nodes tightly connected to each other and less connected to the rest network.

3.2 The backbone of the product space

As shown in Table 1, the snapshots of the product space at different periods in time are very dense (between 83.55% and 91.65% of the potential links are actual links). The density of these networks is tightly linked to their construction. In particular, ubiquitous products are connected to most other products even though the weights attached to them are extremely low due to the normalisation. However, we are interested in the capabilities that help countries develop. If all countries have capabilities X, associated with the production of product P, then clearly X is not the capability that distinguishes development

between the two products will increase linearly with the number of countries the pair of products has in common. Conversely, this relationship is convex when using the Jaccard index. This means that the similarity will increase with the number of countries the two products have in common and at an increasing rate. Here again, there is no reason to believe that this relationship should be convex.

For these reasons, I use the proximity index as a measure of similarity between products.

potential from lack thereof. Even if X is necessary for development, it is other capabilities that make the difference. For this reason, it is important to downplay X, hence the need to filter out some of the links in the product space. At the same time, the presence of edges with extremely low weights makes it complicated to discern the most significant connections and analyse their structure to identify important capabilities. Community detection algorithms are constructed for and tested on low-density networks since the density of most observed networks ranges between almost 0% to roughly 7%. The high density of our product space results in an unreasonable community structure when applying most algorithms to the raw network, giving product groups that contrast sharply with most industrial classifications. One solution to this problem is to reduce the number of links by uncovering the most important connections between products and so lowering the density of the network. However, it is important that the network reduction method keeps the relevant structural features of the original network. The weights of the snapshots of the product space follow a lognormal distribution , which behaves similarly to a power-law indicating the presence of a core-periphery structure. In this study, the network reduction method should preserve this property of the networks.

Table 1: Size and density of snapshots of the product space

Product space snapshots	Number of nodes	Number of edges	Density (%)
1975-1979	769	246710	83.55
1980-1984	783	276431	90.29
1985-1989	783	275234	89.90
1990-1994	783	274707	89.73
1995-1999	781	273477	89.79
2000	769	270628	91.65

To circumvent the problems of traditional network reduction methods[2], Serrano et al. (2009) construct a threshold considering weight heterogeneity at the node level. This method selects the most significant links in a network as determined by a disparity filter. This filter is a p-value representing the probability of observing a normalised weight larger or equal to the weight under the null hypothesis. In the null hypothesis, normalised weights are defined using a random model which accounts for the degree of each node. This probability is defined for a given edge attached to a given node and depends on the degree of the node and the normalised weight of the edge. A homogeneous significance level is chosen to filter out edges for which the p-value is above the significance level. Every edge involves two nodes, and so has two p-values. The edge is kept if it is significant for at least one of the two associated nodes. As

[2]Among network reduction methods for weighted networks, the application of a global threshold on the weights, which consists in withdrawing the edges of the network for which the weight is below a global threshold, is often used. However, this method performs very poorly when applied to networks with weights and strength following heavy tailed distributions (i.e. networks in which links and weights are not distributed normally) since it systematically discards nodes with low strength. With this method, one can only look at the information present at a higher aggregation level (which are arbitrarily defined by the threshold). This problem is amplified when weights are correlated locally leading to a network with groups of isolated nodes. This correlation is high when dealing with core-periphery networks such as the the product space. And indeed, using this method leads to the isolation of a large proportion of the nodes even when using a low threshold, as revealed in Table 4. The reason is that the median of the distribution of the maximum weight of each node is around 0.6 in these networks. Therefore, for any threshold of 0.6 or higher, at least half of the nodes in the networks become isolated.

one reduces the significance level, more edges are filtered out. However, below a critical level, the graph becomes disconnected. I redefine the condition under which edges are kept using the lower bound of the significance level. Further details on this method and the definition of lower bound of the significance level are provided in Appendix 8.

The density of the network is significantly reduced when applying this multi-scale backbone technique while the network stays entirely connected with $\alpha \geq \alpha_{LB}$ (Table 2). I consider the significance level generating the network with the lowest density and for which all nodes stay connected so as to be able to apply a community detection algorithm. The significance levels of each network snapshot, α_{LB}, are presented in Table 2. The reduced networks have a density ranging from 3.40% to 4.39% as shown in Table 2, which is considerably lower than that of the original networks.

Table 2: Network statistics of snapshots of the product space reduced using Serrano (2009)'s multi-scale backbone technique

	α_{LB}	%TW	%N	%E	D (%)	C
1975-1979	1	100	100	100	83.54	1
	0.09176	12.10	100	4.96	4.14	1
1980-1984	1	100	100	100	90.29	1
	0.10233	8.20	100	3.76	3.40	1
1985-1989	1	100	100	100	90.11	1
	0.09698	9.22	100	4.06	3.66	1
1990-1994	1	100	100	100	91.35	1
	0.11161	10.98	100	4.80	4.39	1
1995-1999	1	100	100	100	91.48	1
	0.10272	10.13	100	4.43	4.05	1
2000	1	100	100	100	93	1
	0.09995	8.47	100	3.81	3.54	1

Edge removed if weight$<\alpha_{LB}$; TW (%), the percentage of total weights left; N (%), the percentage of nodes left; E (%), the percentage of edges left; D, the density of the network; C, the number of components.

3.3 The community structure of the product space

In the product space, links between products reflect some similarity in terms of capabilities and knowledge. In addition, I argue that products are likely to be clustered according to a common set of capabilities they share. In network analysis, community detection algorithms enable to identify these clusters. A community or cluster is defined as a group of nodes more tightly connected together than with the rest of the network.

A wide range of methods have been developed in the last years to detect the community structure of networks.[3] Many of those are based on a community quality function called modularity (Newman and Girvan, 2004). The modularity function compares the fraction of edges inside the communities of the real network relative to the fraction of edges inside the communities of a random version of that network. The maximisation of this function enables one to find the community structure of a network. However, the maximisation of this function suffers from several important drawbacks. First, it is assumed that a random network does not have any community structure. If this assumption fails to be true, then the

[3] For an extensive review of community detection algorithms see Fortunato (2010)

modularity function is no longer a good measure to evaluate the quality of the community structure. However, it has been shown that random networks may have a community structure (Guimerà et al., 2004, Reichardt and Bornholdt, 2006). Second, it has a resolution limit (Fortunato and Barthélemy, 2007). This means that it does not detect clusters that are small with respect to the network, even when they are well defined communities (cliques). The reason is that under the null model, it is assumed that each node can interact with any other node of the network. However, in practice, nodes have a limited horizon. The third limitation has to do with the fact that the modularity function has many local optima near the global optimum and these lead to very different partitions (Good et al., 2010).

I use the Infomap algorithm (Rosvall and Bergstrom, 2008; 2010; 2011) because it does not have these drawbacks. In particular, compared to other algorithms, it performs very well in detecting small communities (Lancichinetti and Fortunato, 2009). This method is based on random walks on the network and information theory. It is that assumed that when a random walker enters a cluster, it will spend a long period of time inside before leaving since nodes inside a true cluster are more densely connected. The quality function is called the map equation and corresponds to the length of the code that describes the path of the random worker on a given network. This code is composed of the codename of the cluster and the codename of the vertex in which the random walker is. When the random walker changes cluster, the code indicates the codename of the cluster the random walker is leaving, the codename of the cluster it is entering and the codename of the new vertex. The strongest community structure corresponds to the one for which the code length is minimised. If the community structure is not well defined, then the random walker will often go in and out of the defined clusters, which will lead to a longer code length. There is a trade off between having a few community codenames along with longer node codenames and many community codenames conjointly with short node codenames.

4 Results

The evolution of the community structure of the product space from 1975 to 2000 reveals several export patterns linked to different sets of industries. In particular, this exercise shows the presence of modern development patterns including global value chains and vertical integration[4] as well as other traditional ones such as factor abundance (Heckscher-Ohlin theorem) and technology (Ricardo, 1817).

The main patterns of evolution in the composition of communities are represented with alluvial diagrams in Figures 3 to 7. In addition to these figures, the interpretation of the results requires quantitative measures to identify communities in terms of the types of products they contain. We can describe the constitution of a community by reference to the industries of the products it contains. That is the content

[4] I always refer to vertical integration at the country level and not at the firm level.

of a community is identified by observing the share of products from each industry it contains as displayed in Figure 2. The set of industries are defined according to the aggregation of goods from the SITC at 2 digits in terms of the industry and of the stage of production (whenever the information is available at such a disaggregated level). For example, textile fibres (SITC 26), textile material manufacturing (SITC 65) and textile manufacturing (SITC 83, 84, 85) are part of the textile industry but belong to three different stages of production. The details on the aggregation are presented in Table 5. Communities are hereafter labelled according to the set of industries strongly present in them.

The composition of communities also evolves over time as individual communities split into several modules or merge with others. The community structure is identified for each snapshot of the produce space (each 5-year period) and so we do not have at this stage information on the correspondence between clusters across the different periods of time. To analyse the evolution of the composition of communities, I measure the extent of product overlap between two communities at time t and $t+1$ using the Jaccard index as suggested in Lancichinetti and Fortunato (2012). The Jaccard index is calculated as follows:

$$J_{q_t,q'_{(t+1)}} = \frac{\mid S_{q_t} \cup S_{q'_{t+1}} \mid}{\mid S_{q_t} \cap S_{q'_{t+1}} \mid} \tag{1}$$

with S_{q_t} corresponding to the set of products in community q at time t.

When the overlap is large, many products are in both cluster q at time t and cluster q' at time $t+1$. This index enables us to find the correspondence between clusters from one period to the next. In addition, the extent of the overlap provides information on the continuity of a community throughout time as well as changes in its composition. A relatively high overlap between one community at time t and several communities at time $t+1$ indicates a split into several modules while a significant overlap of several communities at time t with one module at time $t+1$ reveals a merger. Figure 8 displays heatmaps of the overlap between communities across time.

The electronics community

More than 50% of electronics products are grouped into one community. This module also includes on average more than 50% of goods from photographic apparatus, optical goods and watches manufacturing industries (SITC 88). This association is not surprising given the growing importance of electronics in the making of such products, indicating that internal replacement and structural deepening is taking place in such technologies. The relationship between the electronics and machinery industries is limited and decreasing over time as depicted in the alluvial diagram in Figure 3 and 4. According to Figure 2, the share of electronics products located in the machinery community[5] goes from 32% in 1975-1979 to 21% in 2000. In addition, the ties between the electronics and machinery communities are weak as

[5]During periods 1975-1979 and 1985-1989, the machinery industry corresponds to several communities.

the product overlap between them over the period of analysis is almost inexistent (Figure 8). The use of disaggregated data at the industry level is therefore crucial to better grasp the different patterns of structural change within the manufacturing sector.

As shown in Figure 8, this module is relatively stable over the period of analysis (1975-2000) with an average product overlap between one period to the next being close to 53%. However, the electronics community experiences two important changes over the period studied. While it is almost exclusively composed of medium to high-tech products during the period 1975-1979, the alluvial diagram in Figure 3 shows that, from 1980 onwards, it incorporates crude and manufactured materials for the textile industry (textile fibres - SITC 26 and textile yarns and fabrics - SITC 65) as well as low-tech manufacturing (SITC 89). This merge is also noticeable in the heatmap in Figure 8 as the product overlap between the electronics community in the 1975-1979 period and the metal and textile manufactures in the 1980-1984 period amounts to 10%. This evolution provides evidence for the acceleration of the fragmentation of the production process in the electronics industry during the 1980s. This representation also shows that countries that have entered the electronics global value chain were for the most part little developed in the textile industry as textile manufacturing products (SITC 84, 85, 86) are not part of the electronics community and the product overlap between the two communities over time is never above 3% (Figure 8). The presence of the least complex stages of production in the textile industry within the electronics community can be explained by the fact that the fragmentation of the production process in the electronics industry implies the delocalisation of labour intensive and low skill tasks (such as assembly) to countries where wages are particularly low in the interest of reducing costs.

The share of products from low-tech industries grows until 1995-1999, after which the community splits into two modules in 2000. The electronics community in the period 1995-1999 has an important product overlap with two communities in the next period (2000) as shown in Figure 8. While the electronics community remains strong with 59% of products overlapping between 1995-1999 and 2000, it also has a strong overlap (20% of the products) with a newly formed community composed of textile materials and metal manufactures. Figure 2 provides information on the composition of the two modules. The largest one includes mostly medium to high-tech products as well as some products from low-tech manufacturing industries while the second one is composed of the least complex industries (crude and material manufacturing for the textile industry). Such an evolution can reflect different events. Some countries previously dedicated to low value-added stages of the production process have upgraded their export basket by focusing on the electronics industry. This period also follows the Asian crisis of 1997, which has significantly affected economies and thereby exports within the region.

Finally, there is a noticeable difference between medium and high-tech industries (SITC 75, 76, 77 and 88) relative to low-tech industries (SITC 26, 65 and 89) in terms of the concentration of industries across communities as measured by the entropy of industry shares across communities (Figure 9). While

products from the former are highly concentrated into few communities with an average entropy of 0.35, products from the later are much more spread as entropy is close to 0.6 on average. In other words, countries associated these low-tech industries have different export patterns and having a comparative advantage in these industries is not a sufficient condition to take part in the electronics GVC.

The machinery community

The machinery community is composed of machinery products, most products from the controlling instruments industry (SITC 87) as well as chemicals goods, especially plastics (SITC 5 and 58). Chemical products associated with the machinery industry consist mainly of plastics (SITC 58) and some organic chemicals (SITC 51), dyeing and colouring materials (SITC 53) as well as chemical materials and products (SITC 59). The presence of controlling instruments and chemicals in the machinery module means that countries with a comparative advantage in machinery also have a comparative advantage in those two industries. Despite the fact that these industries are associated with different technological domains, they have common features explaining their co-occurrence in the export basket of countries. The production and export of goods from these industries involves complex processes and requires a great amount of resources to dedicate to R&D, which explains the predominance of industrialised countries in this community. Furthermore, vertical integration also partially explains the high occurrence of both machinery and chemical products in the export basket of countries. According to input-output data from Timmer et al. (2015), in 1995, almost 80% of the production from the chemical industry consists of intermediate consumption used in other industries and a large share is connected to the machinery industry. It is also worth noting that there are important differences between chemical products (ISIC 24) and plastic and rubber products (ISIC 25). While the machinery industry is an important source of income for the plastics and rubber industry as it absorbs close to 20% of its production of inputs, it is limited for the chemical industry (consuming less that 4% of its production of inputs). This explains the preponderance of plastic products within the chemical industry present in the machinery community.

The machinery industry is split into several communities at the first period of analysis (1975-1979) but merges into fewer communities over time as shown in Figure 4. Towards the last three periods of analysis, a large share of products from the chemical industry move to the machinery community (60% in 1990-1994, 49% in 1995-1999 and 41% in 2000). The composition of the machinery community is also affected by an increase of the share of metal manufacturing products (SITC 67, 68 and 69) over time. As it is the case for the plastic industry, input-output data from Timmer et al. (2015) point to important linkages between the machinery and metal industries. In 1995, basic and fabricated metals amount to 20% of intermediary consumption used in the production of machinery goods. The machinery industry[6] also represents the largest source of income for the industries of basic and fabricated metals (21%) during

[6]The machinery industry consists of groups 29, 24 and 35 from the ISIC (Revision 3.1) and the chemical industry includes groups 24 and 25.

the same year. Another common feature highlighted by Pavitt (1984) is that these two industries are both scale-intensive. The appearance of metal manufactured products within the machinery community can be induced by the change in the composition of the export basket of countries exporting a large number of products from the industries attached to this community or by the upgrade of the export basket of countries with a comparative advantage in metal industries. Although the results from this analysis cannot provide a detailed explanation for this evolution, the overwhelming presence of industrialised economies among the set of countries with a comparative advantage in at least two metal manufacturing products within the machinery community suggests that vertical integration, rather than a fragmentation of the production process, better explains these trends.[7] This is in line with the characteristics of the machinery and chemicals industries described in Lall et al. (2004). The fragmentation of the production process is limited by the high value to weight ratio of products and parts in the machinery industry and by the limited divisibility of production processes in the chemical industry. Indeed, on average more than 70% of parts and components from the machinery industry are in the machinery community.

The machinery community also sees its share of textile materials and fibres grow from 1990-1994. Much like for the metal manufacturing industries, this change is unlikely to be linked to a fragmentation of the production process within the machinery industry as it is the case for the electronics industry since industrialised countries have the highest level of export diversity in textile products present in the machinery community.[8]

The textile and food community

A large share of products from textile and food industries form a community during most of the period of analysis. The strong co-occurrence of products from these two industries can be attributed to the fact that they share common characteristics as they are both labour intensive Lall et al., 2004 and supplier-dominated (Pavitt, 1984). Some degree of vertical integration seems to take place in the textile industry as products ranging from crude textile materials (animal skins - SITC 21 and textile fibres - SITC 26), to manufactured textile materials (leather - SITC 61 and textile yarns and fabrics - SITC 65) and to textile manufactures (SITC 84, 85 and 86) are present in the same community. In other words, countries that have a comparative advantage in textile will likely be exporting products from several levels of the value chain. This module is the most stable over time as its average product overlap across time is close to 70% (Figure 8).

There is one noticeable change during the period 1980-1984, where the textile and food as well as the textile manufacturing communities merge (Figure 8 and 5). While during the period 1975-1979, the textile

[7] 63% of countries with a comparative advantage in at least two metal manufacturing products within the machinery community during the period 1985-1989 and 65% in 2000.

[8] During the period 1990-1994, more than 60% of countries with a comparative advantage in at least two textile products located in the machinery community are industrialised countries against less than 25% in the electronics community during the same period.

industry within this module is mainly represented by crude and material textile products, the module concentrating most of the textile manufacturing industry merges with the textile and food community from 1980 onwards. This merge can have several explanations. One possibility is an upgrade of the export basket of countries with a strong focus on agro-based products (including textile materials and fibres) towards textile manufacturing. Another potential explanation is a diversification of textile manufacturing exporters with the incorporation of agro-based products. A more detailed analysis would be required to understand the causes of this merge. Despite the important structural change occurring during the period 1975-1979, the high and increasing overlap of the set of products belonging to this module from one period to the next indicates its growing stability over time. As shown in Figure 8, it is always over 50% and increases with time reaching 79% between 1995-1999 and 2000. The increased cohesion of this community over time reflects the static composition of the basket of exporters associated with this community, which translates into an increased resistance to structural change towards modernisation.

Finally, there are some notable differences between the industries present in this module in term of their dispersion across communities as illustrated in Figure 9. While products from the textile manufacturing industry are mostly concentrated in the food and textile community (entropy is 0.23 on average), products from the food industry, textile fibres industry and especially textile materials industry are much more spread across communities (their average entropy is respectively 0.53, 0.56 and 0.61).

The petroleum community

The petroleum community is identified in most of the periods of analysis and regroups most products from the petroleum industry. The presence of inorganic chemicals within this community shows that some petroleum exporters have upgraded their exports through vertical integration. Indeed, crude and refined petroleum is at the bottom of the chemical value chain, directly followed by basic chemicals (inorganic) and polymers. However, results also show that this integration is limited as polymers, which would require greater knowledge and infrastructure, are almost inexistent from the petroleum community.[9]

According to the results depicted in Figure 8, the product content of this community becomes more and more stable over time as the overlap between product sets from one period to the next increases and reaches 66% between period 1995-1999 and 2000.[10] As depicted in Figure 6, the community is broken down into several modules only during the years following the second oil crisis of 1979 (1980-1984), which is likely to have disturbed export flows in this industry.

The low level and limited variability of the entropy of shares of petroleum products across communities (0.21 on average) confirm the concentration of petroleum products in one or a few communities (Figure 9). Overall, this outcome makes the petroleum industry rather peculiar relative to other resource-based

[9] Only 1 is found during 1985-1989, 3 during 1990-1994, 1 during 1995-1999 and 2 during 2000.

[10] During 1990-1994, the petroleum community merges with the one of iron and steel. Note that over this period, many communities tend to merge.

industries, for which products are often dispersed across communities and unstable over time.

The wood and paper communities

The wood community mostly consists of crude wood and wood manufacturing, indicating some degree of vertical integration from countries with abundant wood resources. The wood community is maintained over the period of analysis as evidenced by its relatively high product overlap from one period to the next, shown in Figure 8 (close to 45%).[11]

Some exporters of wood manufacture focus solely on wood while others diversify into resource-based manufactures such as metal and paper or into textile and food manufacture. The export pattern relating wood and other resource-based manufacture only holds for two periods as the community related to resource-based products is discontinued from the period 1985-1989 onwards. The presence of wood manufacturing in the textile and food community decreases from 18% during the period 1975-1979 to 9% during the rest of the periods of analysis[12].

Crude wood products have similar export patterns with wood manufactures but some important differences are worth mentioning. Contrasting with wood manufacturing, crude wood products are concentrated into a fewer communities as indicated by the lower level of entropy in Figure 9. Furthermore, there is a shift of crude wood products from communities including wood manufacturing products (wood module and textile and food module) towards the paper community. The share of crude wood products located in the paper community increases over time from 11% during the period 1975-1979 to 33% in 2000.

Vertical integration in the paper industry ranges from crude wood to pulp wood and to paper products as all these products are present in the same community. The paper community is sustained over the period of analysis with a product overlap of above 30% on average (Figure 8). The entropy of shares across communities depicted in Figure 9 provides evidence of the high uniformity of export patterns within this industry and increasingly so towards the high-end of the value chain (paper products).[13]

A pattern of diversification is identified from period 1990-1994 onwards as the metal and paper modules merge into one community. Figure 8 shows that this newly formed community is stable over time as the products' overlap is over 40% between periods from 1990 onwards.

These results contrast with Leamer (1984), where the paper and wood industries are in the same product class. The present study reveals the boundaries present in most countries between the wood and

[11] The wood community is broken down during the period 1990-1994 with most products shifting to the textile and food community. However, there is tendency for communities to merge during this period, which is likely to be linked to a change in the data rather than a relevant evolution of export patterns. In addition, most products shift back to the wood community from 1995 onwards.

[12] With the exception of period 1990-1994, which is subject likely to be subject to irregularity in the data

[13] The set of paper products is the most concentrated across communities with an average entropy of shares slightly above 0.15. Backward industries producing the inputs for the paper industries (pulp wood and crude) are also concentrated within a few communities over the period of analysis relative to other resource-based products.

paper industries, especially at the manufacturing level.

The metal communities

There exist differences within the metal industry regarding the type of metal and the stage of production. Crude metal products tend to be less dispersed across communities than metal manufacturing products as evidenced by its lowest entropy of shares across communities reported in Figure 9. Natural resources are characterised by export concentration in a limited set of countries. Crude and metal manufacturing exports (SITC 28 and 68) are often found in the same communities, attesting to the limited upgrade of crude metal exporters.

Iron and steel products are regrouped in one community that can be identified in each period of analysis while other metal products are found in a relatively wide range of communities. That is, we can expect a greater degree of specialisation in the export of iron and steel than other metals. Nonetheless, the product overlap from one period to the next in the iron and steel community is on average relatively low (31%) and variable, ranging from 20 to 40%, as presented in Figure 8, which means that some change in the composition of the basket of iron and steel exporters occurs.[14] The alluvial diagrams in Figure 7 and Figure 2 show that from 1985 onwards, around 20% of the products from the iron and steel industry shift to machinery communities on average.[15] A common feature of industries related to the manufacture of metal is their increasing presence in machinery communities. By 2000, above 25% of products relative to metal manufacturing industries (SITC 67, 68 and 69) are located in the machinery community. The share of goods relative to manufacturing metal products (SITC 69) has experienced the greatest increase as it has been multiplied by almost 5 from 1975 to 2000 while iron and steel products remain the predominant metals within the machinery communities. This evolution is likely to be due to the use of metals, especially iron and steel, as inputs in this industry. Local value chains linking metal and machinery products exclude crude metals because those require the abundance of such resources in the country and involve much less complex or no manufacturing process.

Another evidence of structural change is worth mentioning. During the period 1990-1994, the merge between the metal and paper communities results in the formation of a new community. Although part of the metal products (both crude and worked) leave this community to form a new one almost exclusively composed of metal goods during the period 1995-1999, the module associating metal and paper products remains until 2000 with an average product overlap of 44% from one period to the next (Figure 8).

[14] As it is the case of many communities during the period 1990-1994, a merge occurs between the iron and steel module and the petroleum one.

[15] In the period 2000, iron and steel products are found in two communities with a predominance of machinery goods.

5 Discussion: export patterns, structural change and factors of production

Overall, the community structure and its evolution show that the factors explaining industrial patterns and structural change are more complex than the traditional divide between low-tech and medium to high-tech industries. Below I examine the community structures that emerge and identify a set of potential common drivers that explain the emergence and evolution of communities.[16] Clusters can form due to *(i)* the experience in a technological domain, *(ii)* the abundance in production factors, *(iii)* the need for large scale economies and *(iv)* vertical integration.

The community corresponding to electronics products (SITC 75, 76, 77) and photographic equipment, optical and watches (SITC 88) forms around their *experience in a technological domain*, which corresponds to the one associated with the phenomenon of the electron. In addition, this association reveals the growing importance of electronics in the making of photographic equipment, optical goods and watches, indicating that internal replacement and structural deepening taking place in such technologies. Also, the electronics and machinery industries are associated with an *experience in different technological domains*. Their location in distinct communities (Figure 3 and 4) illustrates the importance of accounting for heterogeneity within the manufacturing sector. Moreover, the relationship between the electronics and machinery industries is limited and decreasing over time as depicted in the alluvial diagram in Figure 8.

The relative *abundance in production factors* also influences the structure of communities. The results reveal the importance of several factors. *Natural resources* are a determinant driver in several cases. There exists a community associated with petroleum products, one regrouping iron and steel goods, several with a predominance in metal exports and another focused on the export of wood products. The community corresponding to textile and food products appears due to the abundance of *low-skilled labour* (Lall et al., 2004).[17] The formation of the machinery community characterised by a high concentration and predominance of high to medium-tech industries (machinery, controlling instruments and chemicals) is likely to be associated with important *technological capabilities* and *financial resources* as they involve knowledge intensive and complex processes and require a great amount of R&D.

Furthermore, *vertical integration* seems to be a fundamental determinant in the community structure and its evolution during the period of analysis. We can observe several cases of vertical integration within

[16] A more detailed analysis of the results is presented in section **??**.

[17] Another common characteristic shared by these two industries is that they are supplier-dominated (Pavitt, 1984).

countries.

Some degree of vertical integration within countries is noticeable and deepening around the machinery industry with several industries involved: chemicals, iron and steel and metal. The association of machinery and metal products could also be driven by *the need for large scale economies.*

Vertical integration at the local level is also present in industries involving less complex processes. The set of chemicals (mostly inorganic) present in the petroleum community increases over time, indicating that some petroleum exporters have upgraded their exports through vertical integration. However, this integration is limited as polymers are are almost nonexistent in this community. This can be explained by the fact that polymers are at the high end of the value chain and so their production would require greater knowledge, technology and infrastructure.

There exist two integrated value chains associated with the export of wood, which evolves over time. One community groups most wood products including crude and manufacturing wood while another represents the paper industry as it comprises crude wood, pulp wood and paper products. The composition of these two communities evolves as crude wood products tend to shift from the wood towards the paper community.

From 1980 onwards, the electronics industry and other low-tech industries are merged into a single community, indicating the presence of vertical integration at the global level. This evolution provides evidence for the acceleration of the fragmentation of the production process in the electronics industry during the 1980s. This representation also shows that countries that have entered the electronics global value chain were for the most part little developed in the textile industry as textile manufacturing products are not part of the electronics community and the product overlap between the two communities over time is never above 3% (Figure 8). The presence of the least complex stages of production in the textile industry within the electronics community can be explained by the main objectives of the fragmentation of the production process: reducing costs. Due to this, the fragmentation of production implies the delocalisation of labour-intensive and low skill tasks (such as assembly) to countries where wages are particularly low. This community splits again in 2000 to form two modules, one including mostly medium to high-tech industries and another composed of low-tech industries. Such an evolution can reflect different events. Some countries previously dedicated to low value-added stages of the production process may have upgraded their export basket by focusing on the electronics industry. This period also follows the Asian crisis of 1997, which significantly affected economies and thereby exports within the region. However, a deeper analysis would be needed to shed light on the causes of this evolution.

By contrast, there is no evidence of fragmentation of the production process in the community associated with the machinery, controlling instruments and chemical industries. This result is in line with the characteristics of the machinery and chemicals industries described in Lall et al. (2004). The fragmentation of the production process is limited by the high value to weight ratio of products and parts in

the machinery industry and by the limited divisibility of production processes in the chemical industry. Indeed, on average more than 70% of parts and components from the machinery industry are in the machinery community. However, this result should be taken with caution as the sample excludes Eastern Europe countries, which is where a large part of the global value chain (GVC) in the machinery industry has taken place after 1989.

Finally, I have argued that firms cluster into industries according to a set of common capabilities, technologies and knowledge. However, these boundaries are likely to differ across industries. Recall that a community is defined here as a set of products that are likely to be exported together, which I assume provides information on their relatedness in terms of the capabilities they require. An industry refers to a group of products derived from a traditional and ad-hoc classification: the SITC (see details in Table 5). I look at the dispersion of products from each industry across communities to test the strength of the boundaries between industries. If products from one industry are clustered into one community, then it is likely that this industry has fairly strong boundaries. The reason for this is that the countries exporting them are likely to export other products from the same industry (according to the definition of the cluster). Conversely, if products from one industry are spread across several communities, then this industry has softer boundaries. A high dispersion of the products of a given industry can also be a sign that the capabilities required within a industry are heterogeneous. In other words, the capabilities required to make one product can be different from the ones to make another product from the same industry. In what follows, I describe the tools used in the analysis of the results.

Results show that while some industries are represented by one or two communities, others have their products distributed across a wide range of communities. I calculate the entropy of shares of products from industry i across q communities to quantify the dispersion of an industry's products across a set of communities as follows:

$$H_q^i = \frac{1}{Q} \sum_{q=1}^{Q} s_q^i \ln s_q^i \tag{2}$$

with $s_q^i = \frac{x_q^i}{\sum_{q=1}^{Q} x_q^i}$

This measure provides an indication on the strength of the boundaries of each industry. The results are presented in a heatmap in Figure 9. When H_q^i is low, the products of the industry are found in one or a few communities and exporters within this industry are characterised by relatively uniform export patterns. This industry displays strong boundaries, suggesting that it would not be straightforward for countries to start exporting products from this industry if they do not have any in their existing export basket.

The degree of uniformity of export patterns differs between low-tech relative to medium and high-tech industries. This is to be expected as low-tech industries require a less diverse set of capabilities. As a

consequence, it is easier to start exporting products from these industries even when the export basket of countries does not contain any of their products. Low-tech industries such as textile, wood, paper and furniture tend to be characterised by a wide range of export patterns. Figure 9 shows that the entropy is often high in these low-tech industries, meaning that products are distributed across a large number of communities. Conversely, the entropy of medium to high-tech industries such as machinery, controlling instruments, electronics, optical goods and plastics (4 out of 6 periods) is much lower in general, with the exception of chemicals. These results indicate that such industries tend to exhibit a relatively homogenous export patterns. In other words, countries exporting products from these industries are likely to focus on them. Low entropy is also found in most resource-based industries including petroleum, coal, wood and pulp wood, metal and, in half of the periods, iron and steel. The general high uniformity of exports found in these industries is not surprising as the large majority of countries abundant in natural resources find it difficult to develop and diversify their production and export basket. Overall, these patterns are relatively constant over the period of analysis.

6 Conclusion and limitations

Uncovering the community structure of the product space and its evolution enables us to identify potential paths for structural change by providing information on the linkages between products and industries. This exercise enables us to map several industries and groups of industries according to common characteristics, thereby reinforcing the empirical evidence of the importance of industry differences (Malerba, 2002) at a global level. Technological development is an important determinant of export patterns as medium to high-tech products tend to be grouped together, as evidenced by the community comprising machinery goods, controlling instruments and chemicals and the electronics community. Results also indicate that the structure of export patterns also depends on bounds defined by domains of knowledge and technology, partly explaining the location of medium to high-tech products in different communities. This study also provides empirical evidence of the evolution of the"redomaining" of some technologies towards electronics over time. Furthermore, according to the results, industry differences cannot be solely explained by technological complexity and domains but also by factor abundance and vertical integration at the local and global levels. On the one hand, the abundance of low skilled labour in some countries is at the source of the formation of a community regrouping the food and textile industries. In addition, several modules are characterised by products associated with a certain number of natural resources. On the other hand, most communities gather products associated with the same value chains indicating the manifestation of vertical integration, which tends to intensify with time. This is true for all industries, independently of their technology intensity, as value chains are found in the ma-

chinery, petroleum, wood and paper communities. While in these cases, vertical integration is likely to take place within the country, the presence and composition of the electronics community shows evidence of an acceleration of the fragmentation of the production process within the industry from the 1980s onwards. This analysis also provides information on the export profile of countries that have succeeded to enter this global value chain. Finally, the analysis presented here provides evidence of the relatively high heterogeneity of export patterns *within* some industries, especially low-tech ones, such as those of textile and food.

Several limitations are worth noting. Gross export data comprise important measurement errors including double counting and does not account for product quality. In addition, the level of disaggregation of this data does not allow to distinguish between parts and components and final products. Countries' capabilities are thereby imperfectly estimated, and by consequence, some of the links between industries are biased. This bias could be limited by looking at the value added share of exports data. However, these data should be available at a high level of product disaggregation and cover a wide range of countries and years. Lastly, although this analysis has enabled us to identify important factors explaining export patterns and their evolution, further work is needed to shed light on countries' export trajectory to explain changes in the community structure over time. For instance, it is important to explain the causes of the incorporation of metal products within the machinery community. Is this change the result of an upgrade of the export basket of metal exporters or of a diversification of industrialised countries?

References

Arthur, W. B. (2009). *The nature of technology: What it is and how it evolves*. Simon and Schuster.

Balassa, B. (1965). Trade liberalisation and "revealed" comparative advantage1. *The Manchester School 33*(2), 99–123.

Dosi, G., O. Marsili, L. Orsenigo, and R. Salvatore (1995). Learning, market selection and the evolution of industrial structures. *Small Business Economics 7*(6), 411–436.

Feenstra, R. C., R. E. Lipsey, H. Deng, A. C. Ma, and H. Mo (2005). World trade flows: 1962-2000. Technical report, National Bureau of Economic Research.

Fortunato, S. (2010). Community detection in graphs. *Physics Reports 486*(3), 75–174.

Fortunato, S. and M. Barthélemy (2007). Resolution limit in community detection. *Proceedings of the National Academy of Sciences 104*(1), 36–41.

Good, B. H., Y.-A. De Montjoye, and A. Clauset (2010). Performance of modularity maximization in practical contexts. *Physical Review E 81*(4), 046106.

Guimerà, R., M. Sales-Pardo, and L. A. N. Amaral (2004). Modularity from fluctuations in random graphs and complex networks. *Phys. Rev. E 70*, 025101.

Hidalgo, C. A., B. Klinger, A.-L. Barabási, and R. Hausmann (2007). The product space conditions the development of nations. *Science 317*(5837), 482–487.

Hirschman, A. O. (1958). *The strategy of economic development*, Volume 10. Yale university Press New Haven.

Jaffe, A. B., M. Trajtenberg, and R. Henderson (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *the Quarterly journal of Economics 108*(3), 577–598.

Lall, S. (2000). The technological structure and performance of developing country manufactured exports, 1985-98. *Oxford development studies 28*(3), 337–369.

Lall, S., M. Albaladejo, and J. Zhang (2004). Mapping fragmentation: electronics and automobiles in east asia and latin america. *Oxford Development Studies 32*(3), 407–432.

Lancichinetti, A. and S. Fortunato (2009). Community detection algorithms: A comparative analysis. *Phys. Rev. E 80*, 056117.

Lancichinetti, A. and S. Fortunato (2012). Consensus clustering in complex networks. *Scientific reports 2*.

Leamer, E. E. (1984). *Sources of international comparative advantage: Theory and evidence.* MIT press Cambridge, MA.

Leicht, E. A., P. Holme, and M. E. Newman (2006). Vertex similarity in networks. *Physical Review E 73*(2), 026120.

Lundvall, B.-Å. (1992). *National systems of innovation: Toward a theory of innovation and interactive learning.* Pinter Publishers.

Malerba, F. (2002). Sectoral systems of innovation and production. *Research policy 31*(2), 247–264.

Nelson, R. R. and G. Winter, Sidney (1982). *An evolutionary theory of economic change.* Cambridge, MA: Harvard University Press.

Newman, M. E. and M. Girvan (2004). Finding and evaluating community structure in networks. *Physical review E 69*(2), 026113.

Pavitt, K. (1984). Sectoral patterns of technical change: towards a taxonomy and a theory. *Research policy 13*(6), 343–373.

Reichardt, J. and S. Bornholdt (2006). Statistical mechanics of community detection. *Phys. Rev. E 74*, 016110.

Ricardo, D. (1817). *On The Principles of Political Economy and Taxation.* John Murray.

Rosenberg, N. (1976). *Perspectives on technology.* Cambridge University Press.

Rosenberg, N. (1979). Technological interdependence in the american economy. *Technology and Culture 20*(1), 25–50.

Rosvall, M. and C. T. Bergstrom (2008). Maps of information flow reveal community structure in complex networks. *PNAS 105*(4), 1118–1123.

Rosvall, M. and C. T. Bergstrom (2010). Mapping change in large networks. *PloS one 5*(1), e8694.

Rosvall, M. and C. T. Bergstrom (2011). Multilevel compression of random walks on networks reveals hierarchical organization in large integrated systems. *PloS one 6*(4), e18209.

Serrano, M. Á., M. Boguná, and A. Vespignani (2009). Extracting the multiscale backbone of complex weighted networks. *PNAS 106*(16), 6483–6488.

Timmer, M. P., E. Dietzenbacher, B. Los, R. Stehrer, and G. J. Vries (2015). An illustrated user guide to the world input–output database: the case of global automotive production. *Review of International Economics 23*(3), 575–605.

Weitzman, M. L. (1998). Recombinant growth. *Quarterly Journal of Economics*, 331–360.

Winter, S. G. (1984). Schumpeterian competition in alternative technological regimes. *Journal of Economic Behavior & Organization 5*(3-4), 287–320.

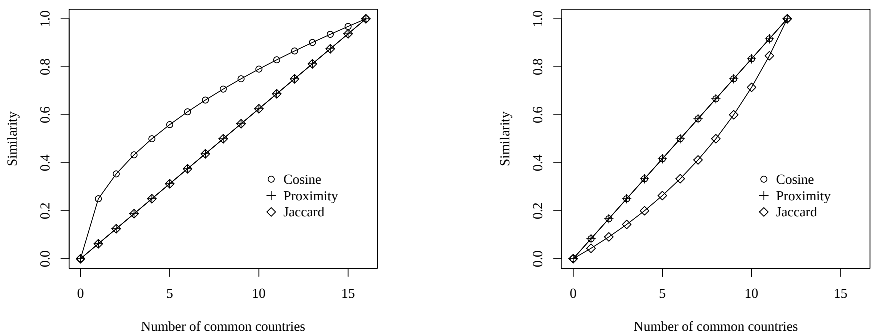
7 Appendix

Table 3: Correlation between the similarity measures

Period	Cosine - Proximity	Proximity - Jaccard	Cosine - Jaccard
1960-1964	0.968	0.982	0.984
1965-1969	0.967	0.981	0.986
1970-1974	0.951	0.979	0.979
1975-1979	0.942	0.977	0.967
1980-1984	0.962	0.979	0.985
1985-1989	0.965	0.980	0.986
1990-1994	0.965	0.980	0.986
1995-1999	0.963	0.980	0.985
2000	0.960	0.979	0.986

Figure 1: Simulation of two specific cases in which similarity measures differ

(a) Case 1 **(b)** Case 2



Note: In case 1, Product A is exported by 16 countries and product B by countries that also export product A. In case 2, the two products that are each exported by 12 countries

8 Network reduction method

In Serrano et al. (2009)'s network reduction method, as one reduces the significance level, more edges are filtered out and below a critical level, the graph becomes disconnected. This provides a natural value at which to set the significance level α. In what follows I define the critical value of α as the smallest value consistent with a connected network. For that condition to be fulfilled, each node should have at least one edge. For the degree of each node to be at least equal to one, the significant level should be higher than the p-value $\alpha_{i,j}$ corresponding to the most significant edge of the node that has the least significant edge among all the vertices. The lower bound of the significance level denoted α_{LB} is defined

as follows:

$$\alpha_{LB} = \max_i\{\min_j\{\alpha_{i,j}^{min}\}\} \tag{3}$$

where $\alpha_{i,j}$ corresponds to the p-value (or disparity filter) of the link between node i and j.

For the graph to stay connected, the significance level should be superior or equal to α_{LB}. I can then redefine the condition under which edges are kept using the lower bound of the significance level as :

$$\tilde{w}_{ij} = \begin{cases} w_{ij}, & \text{if } \alpha_{i,j}^{min} \leq \alpha_{LB}. \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

Table 4: Network statistics of each snapshot of the product space reduced using a global threshold on weights

Period	α	TW (%)	N (%)	E (%)	D (%)	C
1975-1979	0	100	100	100	83.55	1
	0.2	68.33	100	43.28	36.16	1
	0.4	19.85	98.44	8.04	6.72	14
	0.6	2.42	54.10	0.63	0.53	385
	0.8	0.98	14.82	0.20	0.17	680
	1	0.85	5.33	0.17	0.14	735
1980-1984	0	100	100	100	90.29	1
	0.2	73.36	100	49.71	44.88	1
	0.4	22.88	98.72	10.32	9.32	11
	0.6	2.06	57.47	0.66	0.60	364
	0.8	0.09	9.07	0.02	0.02	732
	1	0	0	0	0	783
1985-1989	0	100	100	100	89.90	1
	0.2	70.23	100	46.21	41.54	1
	0.4	20.41	98.47	8.82	7.92	13
	0.6	2.03	55.17	0.62	0.56	392
	0.8	0.15	9.45	0.04	0.03	724
	1	0	0	0	0	783
1990-1994	0	100	100	100	89.73	1
	0.2	68.62	100	44.35	39.79	1
	0.4	18.54	98.85	7.79	6.99	11
	0.6	1.74	53.51	0.52	0.47	404
	0.8	0.08	7.41	0.02	0.02	741
	1	0	0	0	0	783
1995-1999	0	100	100	100	89.79	1
	0.2	68.87	99.87	44.71	40.14	2
	0.4	18.39	97.82	7.77	6.98	18
	0.6	1.54	54.29	0.46	0.41	399
	0.8	0.11	6.91	0.03	0.02	741
	1	0	0.26	0	0	780
2000	0	100	100	100	91.65	1
	0.2	70	100	46.69	42.79	1
	0.4	17.39	98.05	7.56	6.93	17
	0.6	1.17	49.41	0.36	0.33	424
	0.8	0.07	5.98	0.02	0.01	736
	1	0	0.26	0	0	768

Edge removed if weight$<\alpha$; TW (%), the percentage of total weights left; N (%), the percentage of nodes left; E (%), the percentage of edges left; D, the density of the network; C, the number of components.

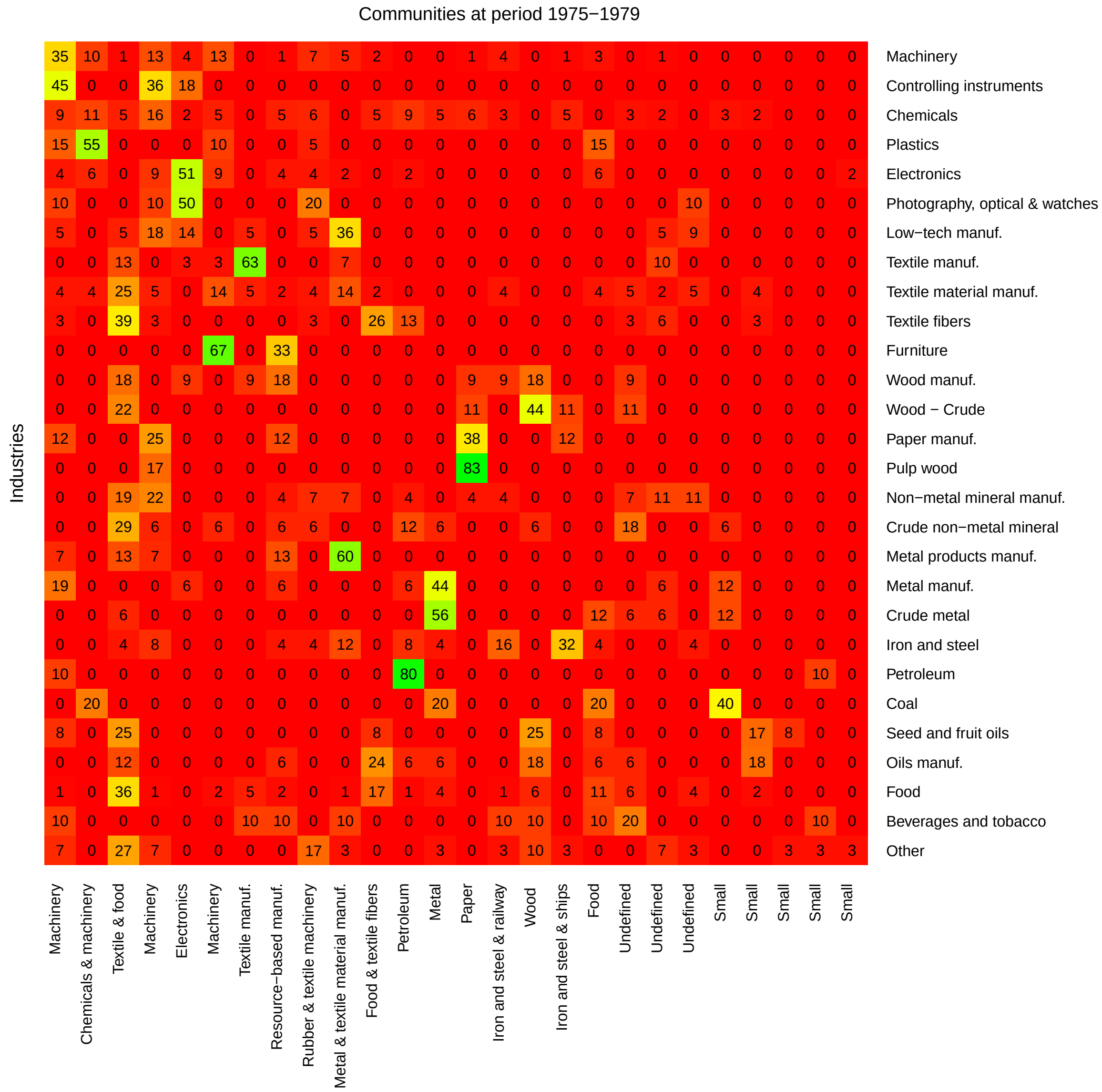
9 Results

Table 5: Definition of the set of industries in terms of SITC Revision 2 at 2 digits

	Industry	SITC 2 digits
1	Machinery	71, 72, 73, 74, 78, 79
2	Controlling instruments	87
3	Chemicals	51, 52, 53, 54, 55, 56, 57, 59
4	Plastics	58
5	Electronics	75, 76, 77
6	Optical goods	88
7	Prints	89
8	Textile - Manufacture	83, 84, 85
9	Textile - Material	61, 65
10	Textile - Crude	21, 26
11	Wood - Manufacture	63
12	Wood - Crude	24
13	Furniture	81, 82
14	Paper - Manufacture	64
15	Paper - Crude	25
16	Non-metal mineral - Manufacture	66
17	Non-metal mineral - Crude	27
18	Metal - Manufacture	68, 69
19	Metal - Crude	28
20	Iron and steel	67
21	Petroleum	33
22	Coal	32
23	Oils - Seed and fruit	22
24	Oils - Manufacture	41, 42, 43
25	Food	00, 01, 02, 03, 04, 05, 06, 07, 08, 09
26	Drinks and tobacco	11, 12

Figure 2: Share of products from each industry across communities (in %)

(a) Period 1975-1979



Note: The entries of the matrix represent the share of products from each industry across communities (in %).

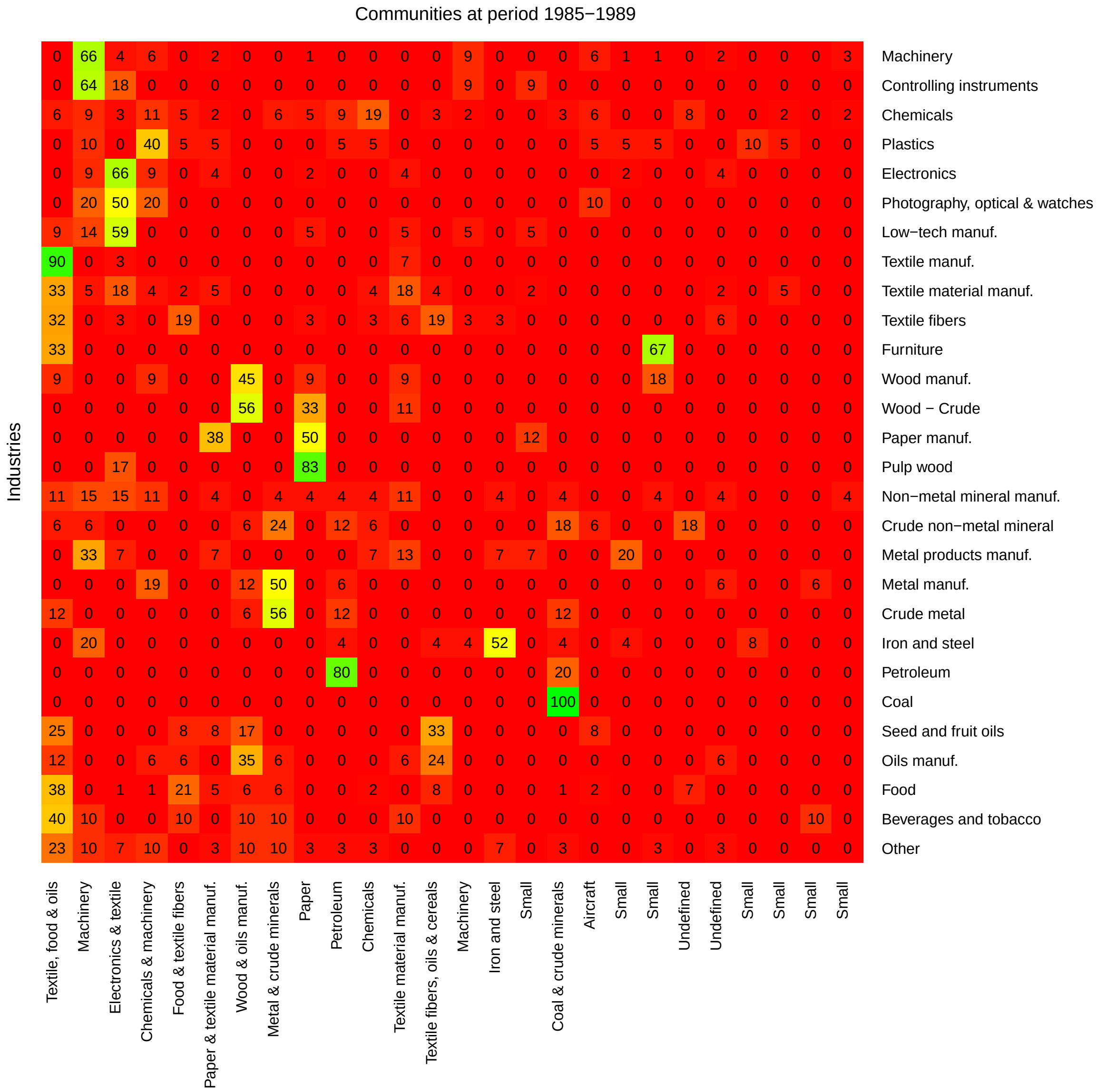
(b) Period 1980-1984

Communities at period 1980−1984

Industries	Machinery	Textile & food	Electronics & textile	Metal	Machinery	Undefined	Resource−based manuf.	Wood & oils manuf.	Paper	Iron and steel	Textile material manuf.	Chemicals	Cereals, crude metal & coal	Textile fibers	Metal manuf.	Textile material manuf.	Basic chemicals	Fertilizers & food	Small	Small	Small	Small	Small	Small	Small
Machinery	71	0	2	0	17	2	0	0	0	1	1	2	0	1	0	0	0	0	1	0	0	0	2	0	0
Controlling instruments	73	0	18	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Chemicals	22	8	2	9	2	8	5	0	3	2	0	11	2	2	3	3	8	5	5	0	0	0	2	2	0
Plastics	80	0	0	0	0	0	0	0	0	5	0	5	0	0	5	0	5	0	0	0	0	0	0	0	0
Electronics	30	0	55	0	0	0	0	0	0	2	0	2	2	0	0	0	0	0	4	0	0	0	0	2	2
Photography, optical & watches	30	0	60	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Low−tech manuf.	9	9	36	0	5	5	0	5	0	0	5	0	0	0	5	5	14	0	0	0	0	5	0	0	0
Textile manuf.	0	97	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Textile material manuf.	11	47	11	0	2	4	0	0	0	0	12	0	0	2	0	7	0	2	0	2	2	0	0	0	0
Textile fibers	3	32	6	19	0	3	0	0	0	0	0	3	0	10	0	6	0	0	0	10	6	0	0	0	0
Furniture	0	33	0	0	0	0	0	0	0	67	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Wood manuf.	9	9	0	0	0	0	18	36	9	9	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Wood − Crude	0	0	0	0	0	0	0	56	22	11	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Paper manuf.	25	0	0	0	0	12	12	0	38	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pulp wood	0	0	0	17	0	17	0	0	67	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Non−metal mineral manuf.	19	19	4	0	7	0	4	0	4	0	15	4	0	4	4	11	4	0	0	0	0	4	0	0	0
Crude non−metal mineral	12	24	0	6	0	6	6	0	12	0	0	12	6	6	0	0	6	6	0	0	0	0	0	0	0
Metal products manuf.	20	27	13	0	0	0	7	0	0	7	7	0	0	0	20	0	0	0	0	0	0	0	0	0	0
Metal manuf.	19	0	6	38	0	6	6	12	6	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Crude metal	0	6	0	69	0	0	0	0	0	0	6	0	12	0	0	0	0	0	0	0	6	0	0	0	0
Iron and steel	8	4	4	8	4	0	0	0	4	48	4	0	0	0	8	0	0	0	0	4	4	0	0	0	0
Petroleum	0	50	0	0	0	0	20	10	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0
Coal	20	0	0	0	0	20	0	0	0	0	0	0	60	0	0	0	0	0	0	0	0	0	0	0	0
Seed and fruit oils	8	58	0	0	0	8	0	8	0	0	0	8	8	0	0	0	0	0	0	0	0	0	0	0	0
Oils manuf.	12	18	0	6	0	6	0	29	0	0	6	18	0	0	0	0	0	0	0	6	0	0	0	0	0
Food	0	51	0	13	0	13	1	5	1	0	1	2	6	1	0	0	0	5	0	0	0	0	0	0	0
Beverages and tobacco	10	40	0	0	0	0	30	0	0	0	10	0	0	0	0	0	0	0	0	0	0	10	0	0	0
Other	20	23	3	7	7	3	0	13	7	0	0	0	0	10	0	3	0	0	0	0	0	3	0	0	0

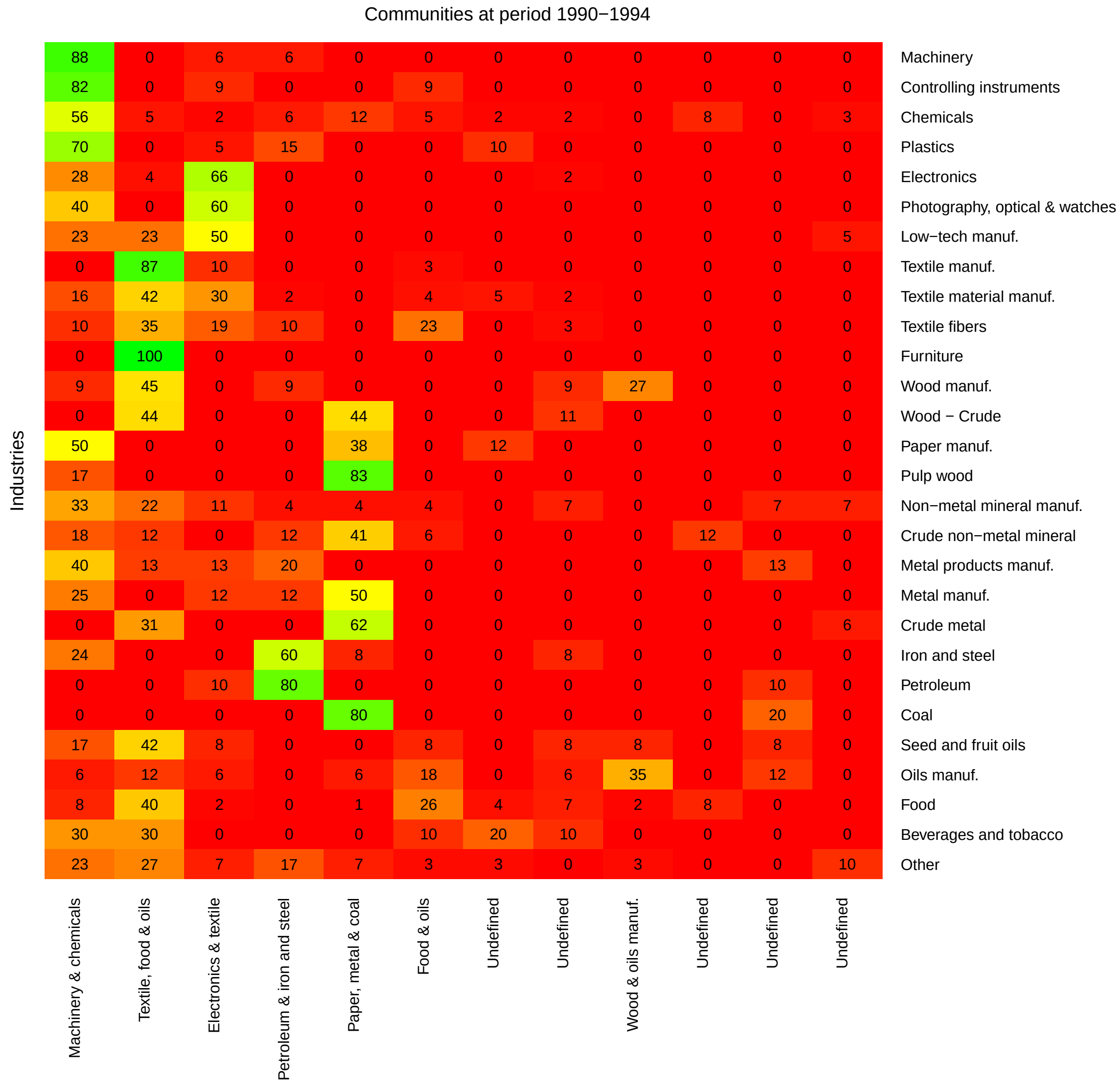
Note: The entries of the matrix represent the share of products from each industry across communities (in %).

(c) Period 1985-1989



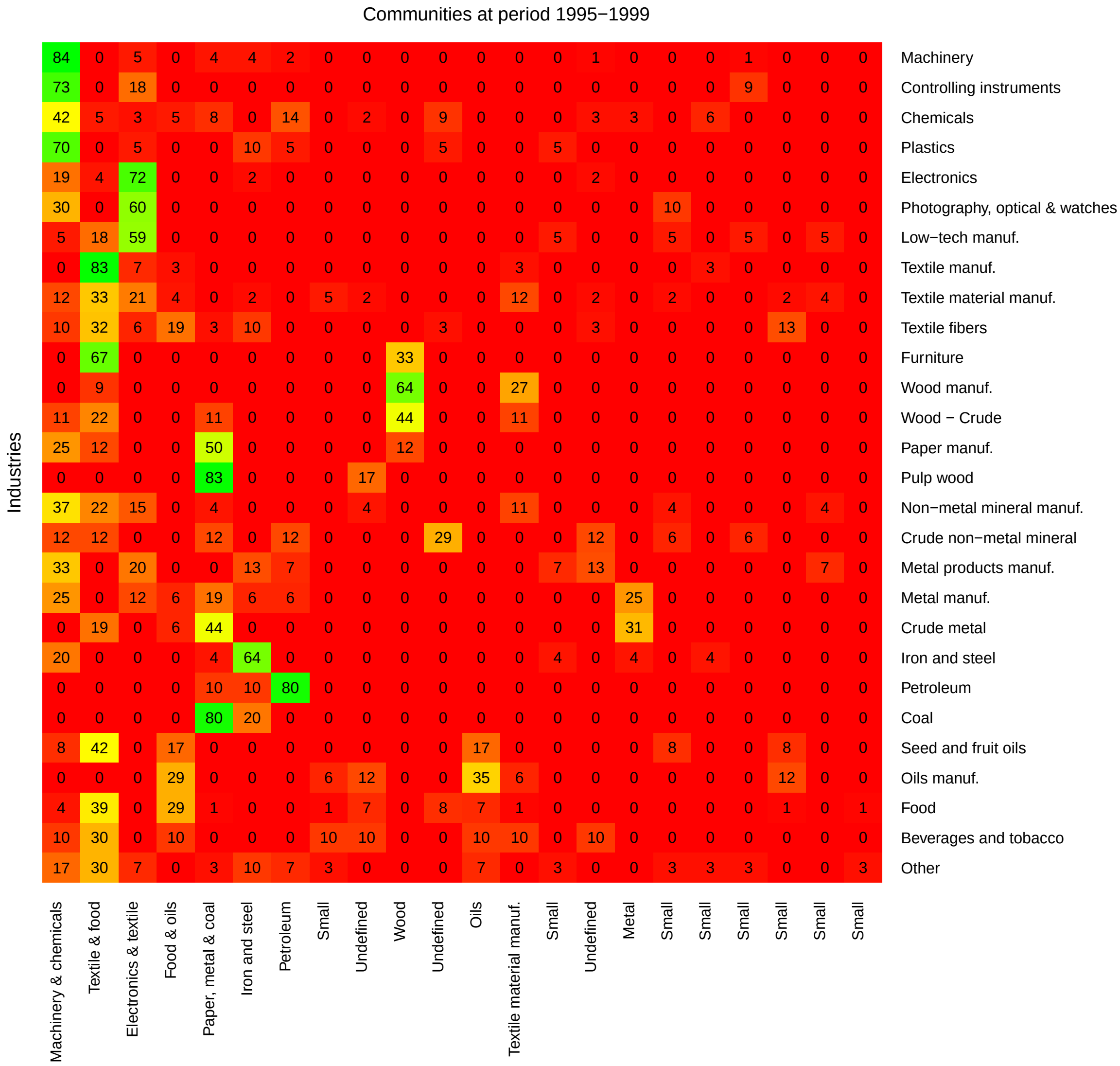
Note: The entries of the matrix represent the share of products from each industry across communities (in %).

(d) Period 1990-1994



Note: The entries of the matrix represent the share of products from each industry across communities (in %).

(e) Period 1995-1999



Note: The entries of the matrix represent the share of products from each industry across communities (in %).

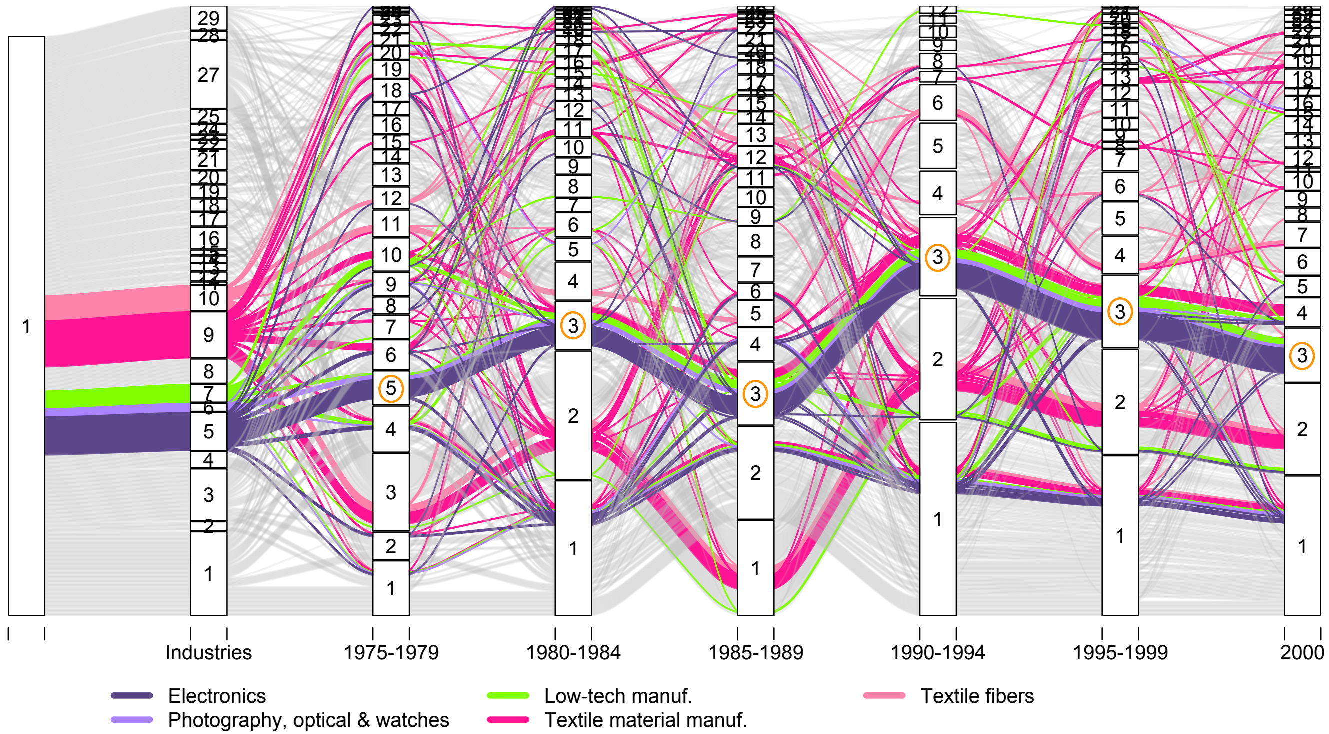
(f) Period 2000

Communities at period 2000

Industries	Machinery & chemicals	Textile & food	Electronics	Textile material & metal manuf.	Machinery	Paper & metal	Food, textile & oils	Food	Petroleum	Food & non−metallic minerals	Small	Iron and steel	Wood & furniture	Metal & chemicals	Small	Oils & food	Small	Food & textile	Textile fibers & material manuf.	Small	Undefined	Small	Small	Small	Coal	Small
Machinery	69	0	5	1	12	2	0	1	0	0	0	0	2	0	1	0	0	1	0	0	0	2	4	0	0	1
Controlling instruments	73	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0
Chemicals	38	3	3	2	3	0	3	2	6	2	2	3	0	12	0	0	6	0	0	0	9	0	0	6	0	0
Plastics	50	0	10	5	0	0	0	0	10	0	0	5	0	5	0	0	5	5	0	0	0	0	0	5	0	0
Electronics	21	4	66	6	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Photography, optical & watches	40	0	40	10	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0
Low−tech manuf.	9	18	36	18	9	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0
Textile manuf.	0	80	3	10	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0
Textile material manuf.	12	30	2	23	2	0	2	0	0	5	0	4	0	0	2	0	0	5	9	2	0	4	0	0	0	0
Textile fibers	3	23	10	3	0	6	16	3	0	3	0	6	0	0	0	0	3	0	13	3	0	6	0	0	0	0
Furniture	0	33	0	0	0	0	0	0	0	0	0	0	67	0	0	0	0	0	0	0	0	0	0	0	0	0
Wood manuf.	9	9	0	0	0	0	0	0	0	0	0	0	55	0	0	0	0	9	0	9	0	9	0	0	0	0
Wood − Crude	0	0	0	0	0	33	0	0	0	0	0	0	33	0	0	11	0	11	0	0	0	11	0	0	0	0
Paper manuf.	0	0	0	12	38	38	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0
Pulp wood	0	0	0	0	0	83	0	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Non−metal mineral manuf.	22	19	15	4	4	4	0	4	0	7	0	0	0	0	4	0	4	4	0	0	4	0	0	0	7	0
Crude non−metal mineral	24	12	0	0	6	18	0	0	6	18	0	0	0	0	0	0	6	0	6	0	0	0	0	6	0	0
Metal products manuf.	33	0	0	33	0	0	0	7	7	0	7	7	0	0	7	0	0	0	0	0	0	0	0	0	0	0
Metal manuf.	31	0	12	0	0	12	6	0	6	0	0	0	0	31	0	0	0	0	0	0	0	0	0	0	0	0
Crude metal	12	19	0	0	0	38	6	0	0	0	0	0	0	25	0	0	0	0	0	0	0	0	0	0	0	0
Iron and steel	8	0	0	0	8	8	0	0	0	0	4	28	8	4	0	0	0	12	12	0	4	0	0	4	0	0
Petroleum	0	0	0	0	0	0	0	0	80	0	0	0	0	0	0	0	0	10	0	0	0	0	10	0	0	0
Coal	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	60	0
Seed and fruit oils	8	25	0	0	0	0	8	0	0	8	0	8	0	0	0	17	0	17	8	0	0	0	0	0	0	0
Oils manuf.	12	0	0	0	0	0	12	6	0	6	0	18	0	0	0	35	0	0	6	0	0	6	0	0	0	0
Food	1	33	0	1	0	2	18	11	0	10	1	2	0	0	0	6	0	7	0	4	2	0	0	0	0	1
Beverages and tobacco	10	30	0	0	0	10	30	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0
Other	13	37	7	0	0	3	0	0	7	7	0	7	0	3	0	3	0	3	0	3	0	0	0	0	0	7

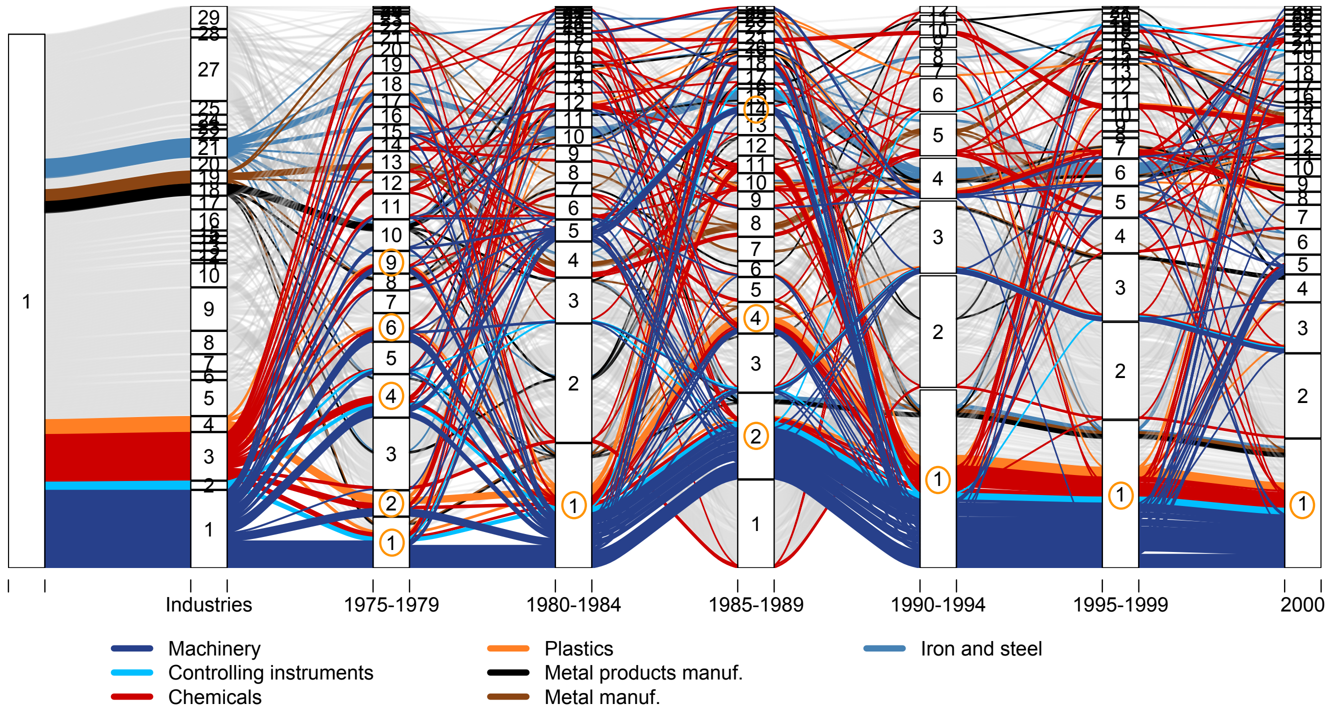
Note: The entries of the matrix represent the share of products from each industry across communities (in %).

Figure 3: Alluvial diagram of the evolution of the communities linked to the electronics industry



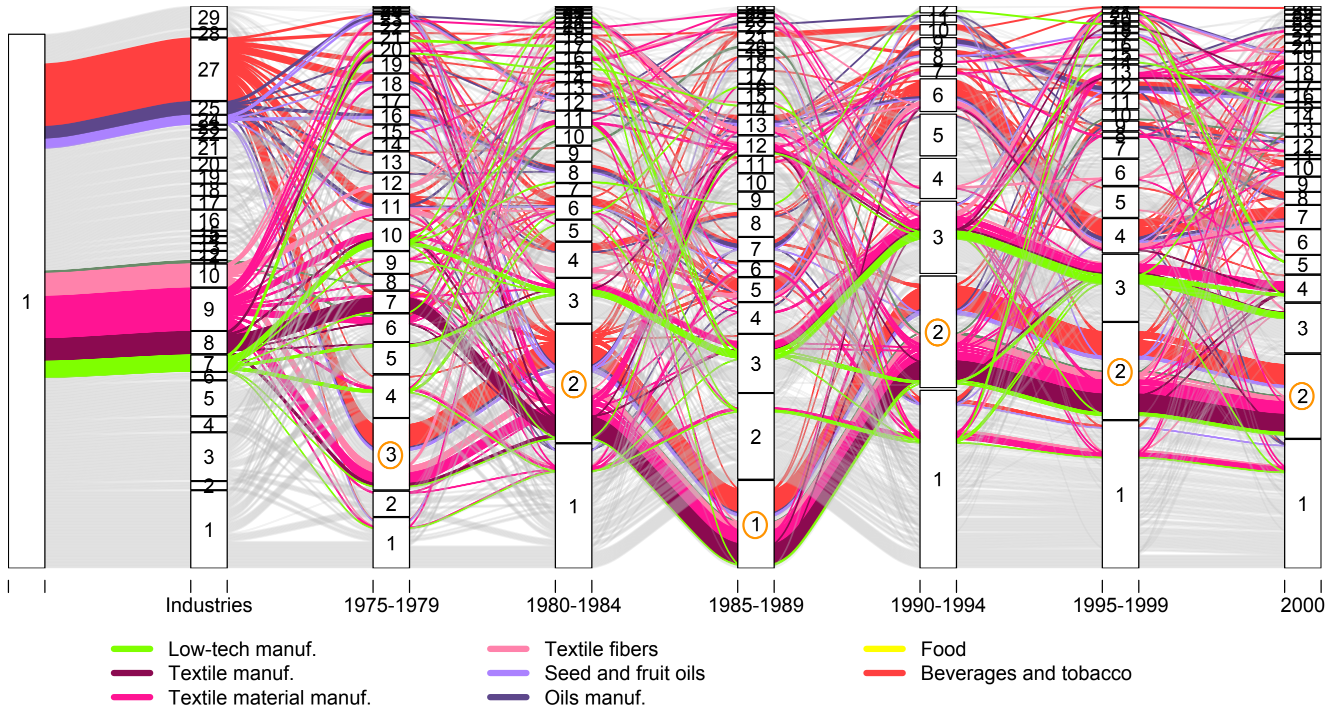
Notes: The block corresponding to "Industries" represent groups of products according the classification detailed in Table 5. Industries that are often found in the community with a predominance in electronics products are highlighted in colour so as to identify the composition of communities linked to the electronics industry. Blocks corresponding to "1975-1979", "1980-1984", "1985-1989", "1990-1995" and "2000" represent the communities of products detected for each 5-year period. The stream fields between the last five blocks represent changes in the composition of these clusters over time. The height of a block represents the size of the community and the height of a stream field represents the size of the product overlap between communities connected by the stream field. The communities circled in orange correspond to communities with a predominance in electronics products.

Figure 4: Alluvial diagram of the evolution of the communities linked to the machinery industry



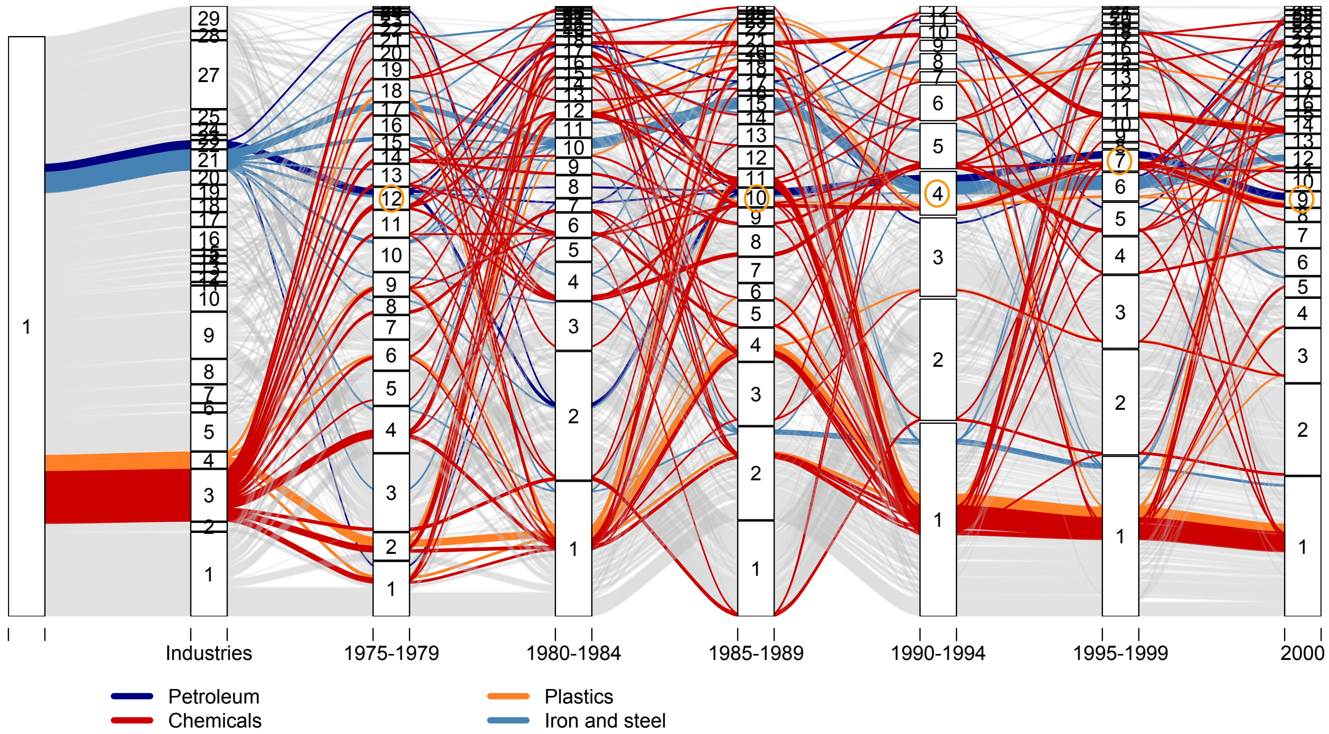
Notes: The block corresponding to "Industries" represent groups of products according the classification detailed in Table 5. Industries that are often found in the community with a predominance in machinery products are highlighted in colour so as to identify the composition of communities linked to the machinery industry. Blocks corresponding to "1975-1979", "1980-1984", "1985-1989", "1990-1995" and "2000" represent the communities of products detected for each 5-year period. The stream fields between the last five blocks represent changes in the composition of these clusters over time. The height of a block represents the size of the community and the height of a stream field represents the size of the product overlap between communities connected by the stream field. The communities circled in orange correspond to communities with a predominance in machinery products.

Figure 5: Alluvial diagram of the evolution of the communities linked to the textile industry



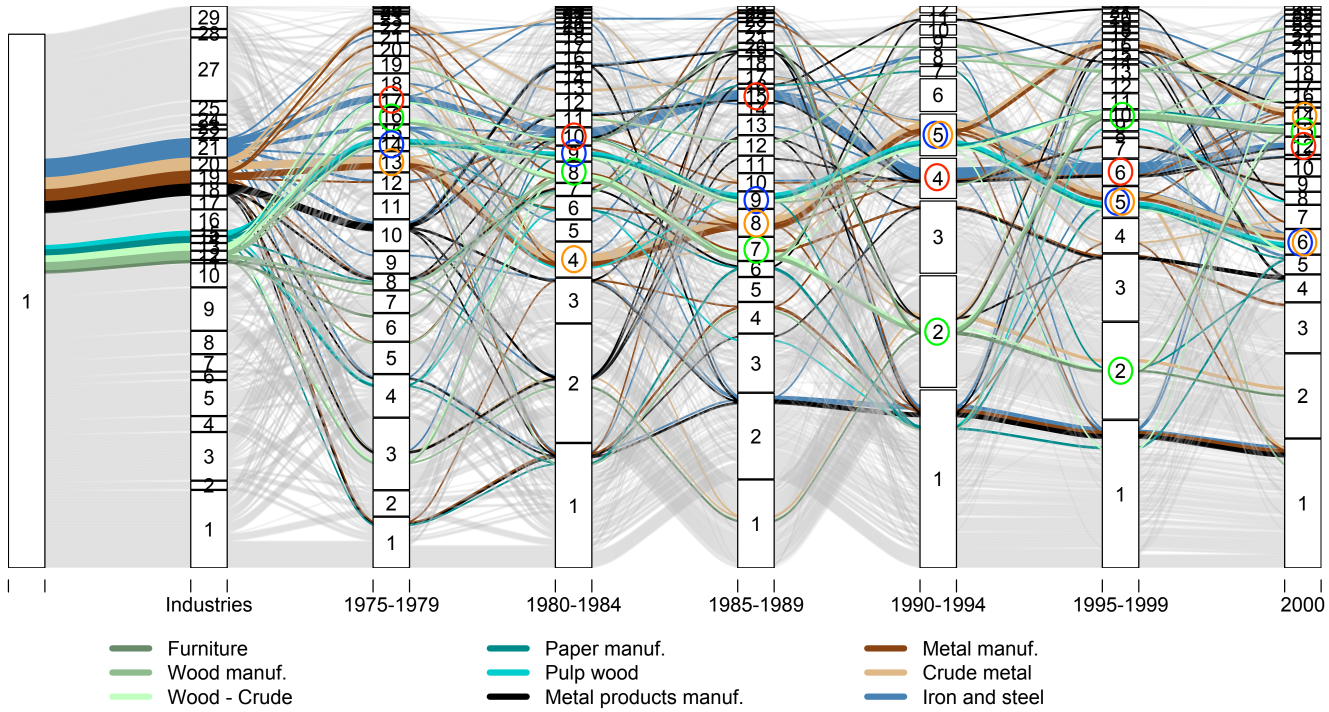
Notes: The block corresponding to "Industries" represent groups of products according the classification detailed in Table 5. Industries that are often found in the community with a predominance in textile products are highlighted in colour so as to identify the composition of communities linked to the textile industry. Blocks corresponding to "1975-1979", "1980-1984", "1985-1989", "1990-1995" and "2000" represent the communities of products detected for each 5-year period. The stream fields between the last five blocks represent changes in the composition of these clusters over time. The height of a block represents the size of the community and the height of a stream field represents the size of the product overlap between communities connected by the stream field. The communities circled in orange correspond to communities with a predominance in textile and food products.

Figure 6: Alluvial diagram of the evolution of the communities linked to the petroleum industry



Notes: The block corresponding to "Industries" represent groups of products according the classification detailed in Table 5. Industries that are often found in the community with a predominance in petroleum products are highlighted in colour so as to identify the composition of communities linked to the petroleum industry. Blocks corresponding to "1975-1979", "1980-1984", "1985-1989", "1990-1995" and "2000" represent the communities of products detected for each 5-year period. The stream fields between the last five blocks represent changes in the composition of these clusters over time. The height of a block represents the size of the community and the height of a stream field represents the size of the product overlap between communities connected by the stream field. The communities circled in orange correspond to communities with a predominance in petroleum products.

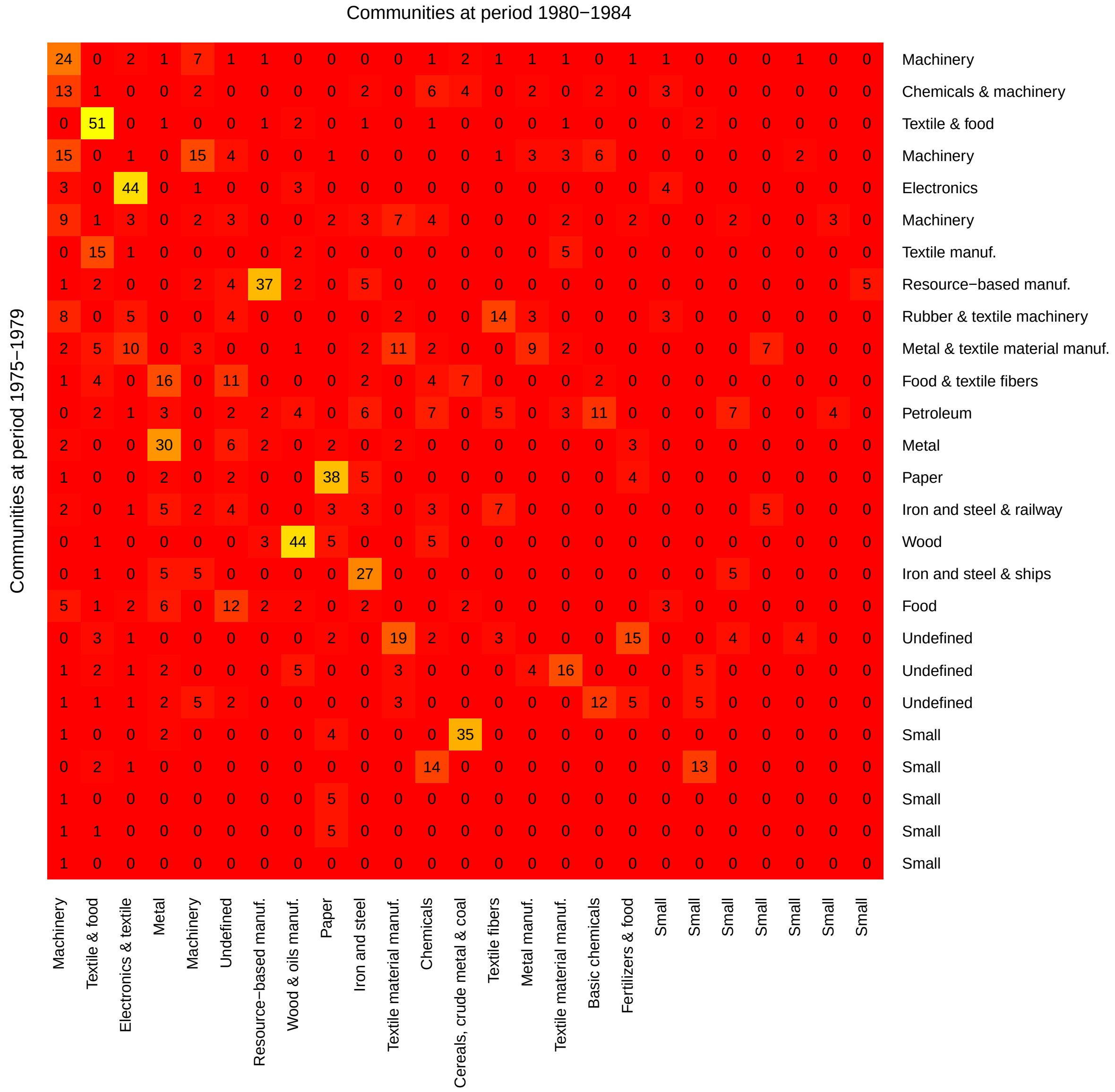
Figure 7: Alluvial diagram of the evolution of the communities linked to the wood, paper and metals industries



Notes: The block corresponding to "Industries" represent groups of products according the classification detailed in Table 5. Industries that are often found in the community with a predominance in wood, paper or metal products are highlighted in colour so as to identify the composition of communities linked to the wood, paper and metal industries. Blocks corresponding to "1975-1979", "1980-1984", "1985-1989", "1990-1995" and "2000" represent the communities of products detected for each 5-year period. The stream fields between the last five blocks represent changes in the composition of these clusters over time. The height of a block represents the size of the community and the height of a stream field represents the size of the product overlap between communities connected by the stream field. The communities circled in green, blue, orange and red correspond to communities with a predominance in wood, paper, metals, iron and steel products, respectively.

Figure 8: Product overlap between communities across time (in %)

(a) From period 1975-1979 to period 1980-1984



Notes: The entries of the matrix represent the product overlap between communities between period 1975-1979 to period 1980-1984 (in %). The overlap is calculated using the Jaccard index.

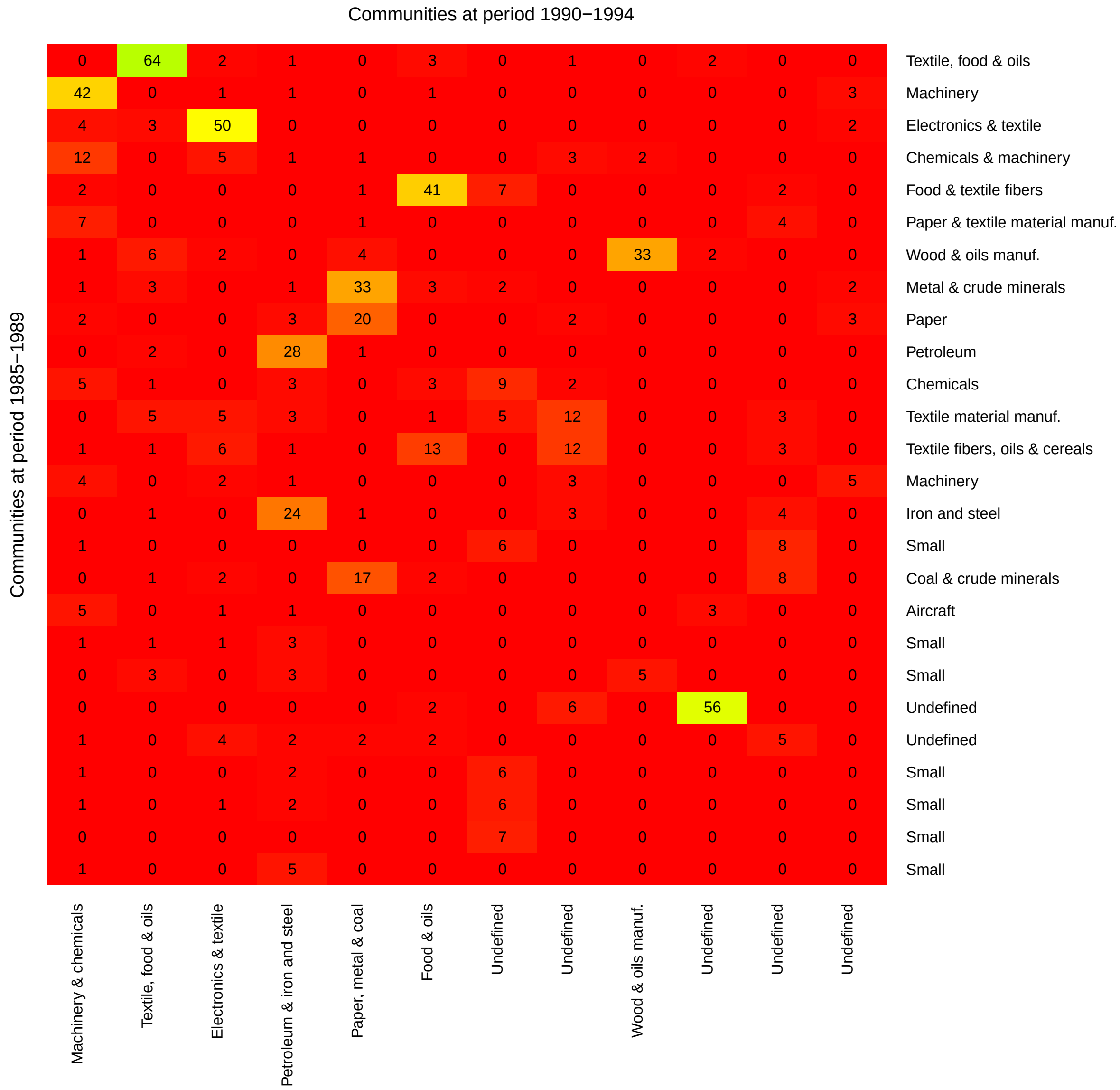
(b) From period 1980-1984 to period 1985-1989

Communities at period 1985−1989 (columns); Communities at period 1980−1984 (rows)

Communities at period 1980−1984	Textile, food & oils	Machinery	Electronics & textile	Chemicals & machinery	Food & textile fibers	Paper & textile material manuf.	Wood & oils manuf.	Metal & crude minerals	Paper	Petroleum	Chemicals	Textile material manuf.	Textile fibers, oils & cereals	Machinery	Iron and steel	Small	Coal & crude minerals	Aircraft	Small	Small	Undefined	Undefined	Small	Small	Small	Small
Machinery	0	40	6	16	1	5	1	0	1	0	3	0	1	2	0	1	2	6	0	0	0	2	1	1	0	0
Textile & food	66	0	1	0	3	0	1	0	0	5	2	5	5	0	1	1	1	1	1	0	2	0	0	0	0	0
Electronics & textile	0	3	50	3	0	0	0	0	0	0	0	4	2	0	0	0	0	0	0	0	0	1	0	0	0	0
Metal	1	1	0	0	14	0	1	48	3	4	0	0	0	0	0	0	2	0	0	0	2	0	0	0	0	0
Machinery	1	9	0	0	0	0	0	0	0	0	0	0	0	24	0	0	0	2	3	3	0	0	3	0	0	6
Undefined	1	2	1	1	17	16	2	2	0	2	4	0	0	0	0	3	2	2	0	0	0	0	0	0	0	0
Resource−based manuf.	0	1	0	2	4	6	2	2	0	3	0	0	0	0	0	5	3	0	0	9	0	0	0	0	12	5
Wood & oils manuf.	0	0	0	0	0	0	71	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Paper	0	1	0	0	2	3	2	4	40	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Iron and steel	0	0	0	0	0	0	0	0	12	2	0	0	0	0	41	0	0	0	7	11	0	0	0	0	0	0
Textile material manuf.	1	1	4	2	0	0	0	2	0	0	0	18	0	0	5	0	0	0	0	0	3	0	0	4	0	0
Chemicals	0	1	1	2	0	0	0	0	2	2	10	0	12	0	0	0	0	6	0	4	6	7	0	0	0	0
Cereals, crude metal & coal	0	0	1	0	5	0	0	0	0	0	0	0	8	0	0	0	33	0	0	0	0	0	0	0	0	0
Textile fibers	0	1	0	6	0	0	0	2	0	0	0	9	0	0	11	0	3	0	0	0	0	0	0	0	0	0
Metal manuf.	0	2	0	0	0	0	0	2	0	0	3	0	0	9	0	0	0	0	13	0	0	0	8	0	0	0
Textile material manuf.	0	0	2	0	2	0	0	0	0	0	0	5	3	4	0	0	0	0	0	5	4	14	0	6	0	0
Basic chemicals	1	1	2	2	0	0	0	0	0	9	13	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0
Fertilizers & food	1	0	0	0	0	4	3	0	0	0	0	0	3	0	0	0	0	0	0	0	28	0	0	0	0	0
Small	0	0	2	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	6	0	0	0	22
Small	0	1	0	0	0	0	0	0	0	0	0	0	14	0	0	0	0	0	0	0	0	0	0	10	0	0
Small	0	0	0	0	0	0	0	0	0	4	4	3	3	0	0	0	0	0	0	0	0	0	12	0	0	0
Small	0	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0
Small	0	1	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Small	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	0	0	0	0	0	0	0

Notes: The entries of the matrix represent the product overlap between communities between period 1980-1984 to period 1985-1989 (in %). The overlap is calculated using the Jaccard index.

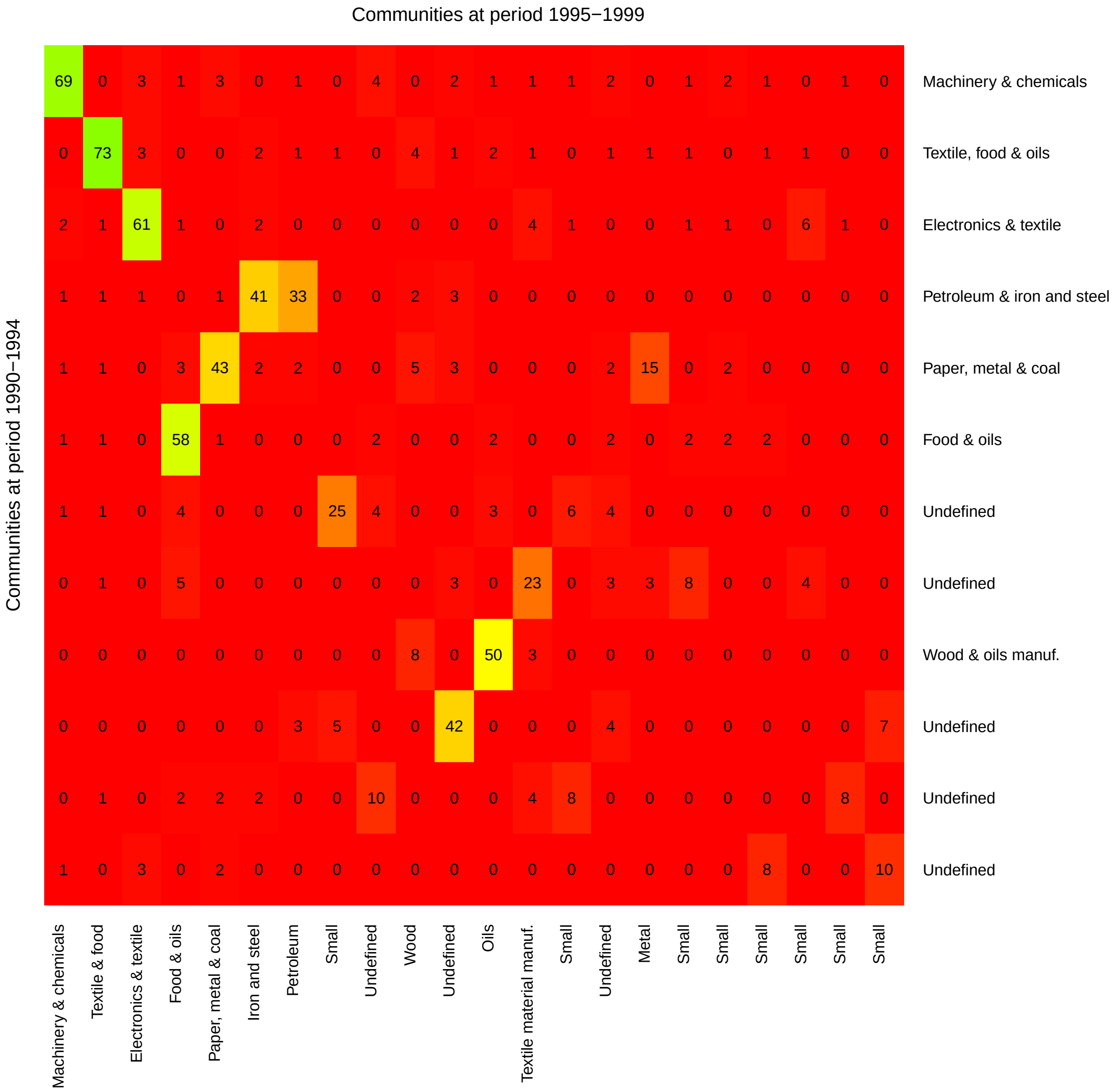
(c) From period 1985-1989 to period 1990-1994



Communities at period 1985−1989	Machinery & chemicals	Textile, food & oils	Electronics & textile	Petroleum & iron and steel	Paper, metal & coal	Food & oils	Undefined	Undefined	Wood & oils manuf.	Undefined	Undefined	Undefined
Textile, food & oils	0	64	2	1	0	3	0	1	0	2	0	0
Machinery	42	0	1	1	0	1	0	0	0	0	0	3
Electronics & textile	4	3	50	0	0	0	0	0	0	0	0	2
Chemicals & machinery	12	0	5	1	1	0	0	3	2	0	0	0
Food & textile fibers	2	0	0	0	1	41	7	0	0	0	2	0
Paper & textile material manuf.	7	0	0	0	1	0	0	0	0	0	4	0
Wood & oils manuf.	1	6	2	0	4	0	0	0	33	2	0	0
Metal & crude minerals	1	3	0	1	33	3	2	0	0	0	0	2
Paper	2	0	0	3	20	0	0	2	0	0	0	3
Petroleum	0	2	0	28	1	0	0	0	0	0	0	0
Chemicals	5	1	0	3	0	3	9	2	0	0	0	0
Textile material manuf.	0	5	5	3	0	1	5	12	0	0	3	0
Textile fibers, oils & cereals	1	1	6	1	0	13	0	12	0	0	3	0
Machinery	4	0	2	1	0	0	0	3	0	0	0	5
Iron and steel	0	1	0	24	1	0	0	3	0	0	4	0
Small	1	0	0	0	0	0	6	0	0	0	8	0
Coal & crude minerals	0	1	2	0	17	2	0	0	0	0	8	0
Aircraft	5	0	1	1	0	0	0	0	0	3	0	0
Small	1	1	1	3	0	0	0	0	0	0	0	0
Small	0	3	0	3	0	0	0	0	5	0	0	0
Undefined	0	0	0	0	0	2	0	6	0	56	0	0
Undefined	1	0	4	2	2	2	0	0	0	0	5	0
Small	1	0	0	2	0	0	6	0	0	0	0	0
Small	1	0	1	2	0	0	6	0	0	0	0	0
Small	0	0	0	0	0	0	7	0	0	0	0	0
Small	1	0	0	5	0	0	0	0	0	0	0	0

Notes: The entries of the matrix represent the product overlap between communities between period 1985-1989 to period 1990-1994 (in %). The overlap is calculated using the Jaccard index.

(d) From period 1990-1994 to period 1995-1999



Notes: The entries of the matrix represent the product overlap between communities between period 1990-1994 to period 1995-1999 (in %). The overlap is calculated using the Jaccard index.

(e) From period 1995-1999 to period 2000

Communities at period 2000 (columns); Communities at period 1995−1999 (rows)

Communities at period 1995−1999	Machinery & chemicals	Textile & food	Electronics	Textile material & metal manuf.	Machinery	Paper & metal	Food, textile & oils	Food	Petroleum	Food & non−metallic minerals	Small	Iron and steel	Wood & furniture	Metal & chemicals	Small	Oils & food	Small	Food & textile	Textile fibers & material manuf.	Small	Undefined	Small	Small	Small	Coal	Small
Machinery & chemicals	64	0	2	1	7	1	0	1	0	0	0	0	0	0	0	0	1	4	0	0	1	0	1	2	0	1
Textile & food	0	79	1	4	0	0	1	0	0	2	0	1	2	0	0	0	1	1	1	2	0	0	0	0	0	0
Electronics & textile	2	0	59	20	0	0	0	0	0	0	0	0	0	0	1	0	0	0	3	0	0	0	0	1	0	0
Food & oils	1	0	0	2	0	3	48	3	0	5	0	1	0	0	0	0	2	6	2	2	2	0	0	0	0	0
Paper, metal & coal	3	0	0	0	3	45	0	0	0	0	0	2	0	3	0	0	0	2	0	0	2	2	0	0	7	0
Iron and steel	0	0	0	1	7	0	0	0	4	0	5	23	6	4	0	0	0	0	6	0	2	7	8	0	0	0
Petroleum	1	0	1	0	0	0	0	0	61	2	3	0	2	2	0	0	3	0	0	0	6	0	0	0	0	0
Small	0	1	0	0	0	0	3	0	0	4	10	7	0	0	0	0	0	0	5	0	0	0	0	0	0	0
Undefined	1	0	0	0	0	0	0	38	0	0	0	3	0	0	0	0	0	3	0	5	0	0	0	0	0	0
Wood	0	0	0	0	0	5	0	0	0	0	0	0	45	0	0	4	0	0	0	5	0	0	0	0	0	0
Undefined	2	2	0	0	0	2	0	6	0	17	0	2	0	14	0	0	0	0	0	0	0	0	0	0	0	0
Oils	1	1	0	0	0	0	4	0	0	0	0	0	0	0	0	65	0	0	0	0	0	0	0	0	0	0
Textile material manuf.	1	0	0	0	0	0	2	0	0	8	0	0	0	0	0	0	0	11	6	0	0	23	0	0	5	0
Small	0	0	0	2	0	0	0	0	0	0	0	4	0	0	10	0	0	4	0	7	0	0	0	0	0	0
Undefined	1	0	0	2	3	0	0	0	0	6	0	3	0	0	13	0	0	0	0	11	5	0	0	0	0	0
Metal	0	0	0	0	0	5	2	0	0	0	0	0	0	39	0	0	0	0	0	0	0	0	0	0	0	0
Small	1	0	0	0	3	0	0	0	0	0	0	7	0	0	0	5	7	0	0	0	0	0	0	8	0	0
Small	1	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	7	7	0	0	6	0	0	8	0	0
Small	0	0	0	0	7	0	0	0	0	8	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0
Small	0	1	3	0	0	2	0	0	0	3	0	0	0	0	0	4	0	0	14	0	0	0	0	0	0	0
Small	1	0	0	5	0	0	0	5	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0
Small	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	50

Notes: The entries of the matrix represent the product overlap between communities between period 1995-1999 to period 2000 (in %). The overlap is calculated using the Jaccard index.

Figure 9: Entropy of industry shares across communities

	1975–1979	1980–1984	1985–1989	1990–1994	1995–1999	2000
Machinery	0.64	0.31	0.41	0.14	0.22	0.38
Controlling instruments	0.31	0.23	0.31	0.18	0.23	0.23
Chemicals	0.83	0.81	0.79	0.48	0.58	0.66
Plastics	0.38	0.23	0.61	0.27	0.32	0.51
Electronics	0.52	0.36	0.38	0.25	0.25	0.29
Photography, optical & watches	0.41	0.27	0.37	0.2	0.27	0.36
Low−tech manuf.	0.58	0.64	0.42	0.35	0.41	0.49
Textile manuf.	0.36	0.04	0.12	0.14	0.2	0.21
Textile material manuf.	0.75	0.54	0.62	0.44	0.6	0.64
Textile fibers	0.5	0.59	0.56	0.48	0.57	0.68
Furniture	0.19	0.19	0.19	0	0.19	0.19
Wood manuf.	0.6	0.53	0.46	0.41	0.26	0.42
Wood − Crude	0.43	0.34	0.28	0.29	0.43	0.44
Paper manuf.	0.44	0.44	0.29	0.29	0.36	0.37
Pulp wood	0.14	0.26	0.14	0.14	0.14	0.14
Non−metal mineral manuf.	0.66	0.71	0.78	0.57	0.52	0.7
Crude non−metal mineral	0.63	0.69	0.62	0.48	0.59	0.62
Metal products manuf.	0.36	0.55	0.57	0.45	0.53	0.5
Metal manuf.	0.48	0.53	0.43	0.36	0.53	0.47
Crude metal	0.4	0.31	0.38	0.25	0.36	0.44
Iron and steel	0.63	0.56	0.45	0.32	0.34	0.66
Petroleum	0.19	0.37	0.15	0.19	0.19	0.19
Coal	0.4	0.29	0	0.15	0.15	0.29
Seed and fruit oils	0.54	0.4	0.49	0.5	0.47	0.59
Oils manuf.	0.62	0.57	0.54	0.56	0.47	0.56
Food	0.64	0.48	0.56	0.49	0.5	0.62
Beverages and tobacco	0.65	0.43	0.52	0.45	0.59	0.45
Other	0.69	0.64	0.71	0.57	0.66	0.63

Note: Industry shares correspond to the number of products in each community relative to the total number of products within the industry.

The UNU-MERIT WORKING Paper Series

2020-01 *Debating the assumptions of the Thirlwall Model: A VECM analysis of the Balance of Payments for Argentina, Brazil, Colombia, and Mexico* by Danilo Spinola

2020-02 *The La Marca Model revisited: Structuralist Goodwin cycles with evolutionary supply side and balance of payments constraints* by Danilo Spinola

2020-03 *Uneven development and the balance of payments constrained model: Terms of trade, economic cycles, and productivity catching-up* by Danilo Spinola

2020-04 Time-space dynamics of return and circular migration: Theories and evidence by Amelie F. Constant

2020-05 *Mapping industrial patterns and structural change in exports* by Charlotte Guillard