Discovering preferential patterns of sectoral trade networks*

Isabella Cingolani1, Carlo Piccardi, Lucia Tajoli

*Preliminary and incomplete. Please not quote

Abstract The aim of this work is to understand whether countries trade preferentially with a limited group of partners in selected sectors. For this purpose we have adopted a systemic approach through the use of tools pertaining to the complex network analysis field. In particular we have applied a pattern detection analysis based on lumped-Markov-chains to three intra-industry trade networks consisting of 222 countries for the year 2006 relative to textile, motor vehicles sectors. It is indeed of great interest to compare the organization of existing intra-industry trade flows between countries to unveil unknown properties of trade system in relation to different industrial characteristics. The methodology allows detecting two types of structures, communities and a hierarchy within which most of trade of members occurs. We found that both these two intra-industry trade networks are characterized by a similar evidence of preferential trade patterns: on the one hand they show few significant communities that define narrow sets of countries trading with each other as preferential destinations markets or supply sources; on the other hand, they are characterized on a greater extent by the presence of similar hierarchical structures, led by the largest economies acting as main supply sources or destination markets and fully absorbing the trade capacity of many other small countries. This evidence suggest that the way firms organize their international activities seems not be so much affected by industry characteristics as much as by economic asymmetries among countries. We have also observed that geographical distance influences trade partner selection acting like an hurdle to trade. Our results suggest also that distance matters more when firms export than when they import.

JEL CODE F14

Keywords Community analysis, Network analysis, Intra-industry trade

1 Introduction

Reveal characteristics of pattern of trade is perhaps the one of the main topic of international trade research history, if considered from very different perspectives. This work is essentially empirical in trying to discover patterns of trade, meaning that it is devoted to explore the reality of trade relations as they are revealed by trade statistics.

Network analysis applications to trade system have already reveal many characteristics of trade structure highlighting the complex nature of it. In particular, our main interest here is to discover real patterns of trade that are characterized by a certain degree of preferentiality by means of a local search algorithm based on information processing related to a random walker. This algorithm is able to detect structure with larger internal connectivity than external as well as structures characterize by the presence of nodes that

1 PhD Candidate in Management, Economics and Industrial Engineering, Politecnico di Milano
isabella.cingolani@polimi.it
are intensively connected to others without being the latter necessarily intensively connected to the former, that is another possible way in which the preferential trade could be organized. Indeed we are not searching necessarily for reciprocal preferential trade. Discovering how many and which countries are involved in, how much trade is characterized by preferences, would add other properties to the already known about organization of trading links.

This work is born in the attempt to conjunct suggestions from complex network research findings on trade network, and relative methods, with the recent and active interest on mapping trade along supply-chains in the clearest way. In particular the focus of this work is on discovering patterns of preferentiality in intra-industry trade, paying attention to possibly different outcomes, given the peculiar characteristics of chosen industries.

The work is organized as follows. In section 2 we start highlighting the some empirical regularities about distribution of trade among countries. We define preferential trade and illustrate the method applied to sectoral trade networks in section 3. In section 4 we describe dataset used and discuss the results; then we conclude in section 5.

2 Empirical regularities on preferential trade (TO BE COMPLETED)

The aim of this work is to analyse trade distribution among countries. Indeed trade is distributed unevenly both at country level and at pair and group level. There is a lot of empirical evidence that countries trade at very different levels of volume, value and number of partners. (WTO, 2008)

Heterogeneity is also a characteristic that affects the distribution of each country trade volume and value among existing trade partners. One of the most cited driving forces that shapes heterogeneity in trade is the geographical distance. Due to this countries seem to prefer trading with partners that are relatively close geographically (Head and Mayer, 2013). But many other reasons could be addressed in order to explain preferentiality patterns in trade: size of countries, level of development, relative factor endowments, relative productivities Depending on the theoretical framework one adopts and the specific aspects of reality addressed by it, one could suggest alternative explanations of the same empirical fact: the country preferentiality in trade decisions.

..... List of existing empirical evidence on preferentiality: south-south trade, north-north trade, north south trade

...plot with heterogeneous distribution of trade (by country and groups of countries)

The empirical research to date has tended to focus mainly on explaining trade shares among groups of countries by level of development, geographical positioning or participation to PTAs. Results suggest that
countries trade more with countries that are more close, but evidence for the other two reasons is not so clear (Hanson, 2012). For example countries seem to trade both with similar countries in terms of developmental stage and also with dissimilar ones.

Misleading interpretation could occur especially when aggregate flows of trade are analyzed, taking into account intra-industry trade and inter-industry trade simultaneously for all sectors. Sectors differ each other in terms of pattern of trade and what is valid at aggregate level could be totally different at industry level when searching how much is important the share of trade hold by the core countries (Nicita and Olarreaga, 2007). Discovering also that length of global supply chains may vary a lot sector by sector suggest us that the pattern of trade varies across the sectors (Baldwin et al, 2013). We know also that the diffusion of offshoring it is not the same across industries (WTO, 2008).

Finally trade is performed mainly within sectors. Also in this case there is a lot of evidence that trade occurs mainly within-sectors and that intermediate inputs trade weight as much as trade in finished products. (De Backer and Miroudot, 2014; Baldwin et al, 2013).

...Our elaboration from WIOD Tables for the last 10 years...

3. Pattern of trade via network analysis

The last years have witnessed the emergence of a large body of contributions addressing international trade issues from a complex network perspective (Li et al. 2003; Serrano and Boguñá 2003; Reichardt and White 2007; Serrano et al. 2007; Bhattacharya et al. 2008, 2007, Tzekina et al. 2008; Fagiolo et al. 2008; De Benedictis and Tajoli, 2011).

Many works were conducted within a network analysis approach to explore the *structural properties* of the network of international trade relations (De Benedictis and Tajoli, 2011; Fagiolo et al., 2008, 2011; Bhattacharya et al., 2008) that is to give a concrete image to the manner economic globalization works. In these works it can be recognized a common trait about trade system pattern: centralization and heterogeneous distribution of links and traded value on them that result in hierarchical structures where the most central countries are often advanced economies or growing giant economies like China.

But more important for research it seemed to be the discovery of community or *modular structures* that is pattern where can be recognize a geographical or cultural closeness probably due to economic advantage in terms of cost reduction of transactions and increasing national welfare (He e Deem, 2010; Tzekina et al., 2008; Reyes et al.,2009; Piccardi and Tajoli, 2012). Others recent contributions highlight the strong
core/periphery profile of a great number of trade relations (Ermann and Shepelyansky, 2012; Della Rossa et al, 2013; Landi and Piccardi, 2014).

The issue of discovering how the trade relation are organized has grown in importance in light of recent developments in studying supply-chain trade (Lejour et al, 2012, Baldwin and Lopez-Gonzalez, 2013, De Backer and Miroudot). Indeed, from these studies it emerges the urgency of mapping the way countries trade each other taking into account the nature of the trade (in intermediate inputs or in finished products) and the economic and geographical “position” of partners.

This work is born in the attempt to conjunct suggestions from complex network research findings on trade network, and relative methods, with the recent and active interest on mapping trade along supply-chains in the clearest way. In particular the focus of this work is on discovering patterns of preferentiality in intra-industry trade, paying attention to possibly different outcomes, given the peculiar characteristics of chosen industries.

From the general overview of trade statistics describing the pattern of trade among countries we are already aware that trade is not homogenously distributed: there are countries that trade more than others and there are groups of countries that are characterized by not irrelevant shares of trade realized within the group in respect to their total trade. But this evidence could be more in deep analyzed, in particular with respect to the latter case.

The parameters that are usually used to classify countries into groups are related to the concept of membership: to a Preferential Trade Agreement or whichever other type of agreement, to a geographical region or sub-region and so on. Generally speaking, this type of choice is inducted by the expectation that “membership” or, in other words, “closeness” could represent the incentive to trade more with partners that are also members of the group. But this is only a part of the story.

The other part of the story is that country could trade more with partners that, although not members of a pre-determined group, are “attractive” for other reasons. Depending on trade models we are considering, many driving forces could be addressed to explain cross-industry and intra-industry trade among countries: the best known examples of drivers of cross-industry trade are the existence of differences in relative factors endowments or productivities. On the other hand, similarity in dimensions of aggregate demand could be the main driver of intra-industry trade as international trade models grounded on monopolistic competition hypotheses suggest. In any case, the dimension of demand is an important factor shaping trade decisions of exporters.

Being the focus of this work on pattern of preferential intra-industry trade, in addition to partners’ market dimension to explain it, we will take into account also other driving forces that could be easily linked to the
concept of Global Value Chains. Indeed, trade statistics we will use to study pattern of preferential intra-
industry trade are not exclusively representative of finished products but also include trade in intermediate
inputs that are exchanged by industrial and service sectors located in different countries. This trade is
driven by foreign demand of intermediate inputs transformed into intermediate output or finished product.
The destinations for them could be local or, again, foreign. Even if brief, this description gives the intuition
that intra-trade could be driven also by the position of countries and their partners along sectoral global
value chains.

If we define *preferential* that trade of goods and services a country prefers to “export to” or “import from”
one or few partners, then *trade shares* can be easily interpreted as signals of such preferentiality when they
are concentrated on one or few partners. So, when we seek to collect simultaneously for all countries
shares of their total exports or total imports that are respectively “directed to” or “coming from” foreign
partners, the result is a *network of trade shares*. In respect to it, our interest here is discovering *preferential
trade patterns*, that is, how trade shares are distributed across partners in the network.

One way of thinking to preferential trade is that it could occur meanly within sub-networks that have the
property of including countries that trade a significant share of their total trade within the boundaries
defined by the sub-network. The patterns of these sub-networks, if they exist, cannot be easily predict: they
could be organized in many different ways, not excluding the above mentioned property about percentage
of their total trade realized within. Indeed one possible configuration is that each member of the sub-
network trades equal shares across all other member of the sub-network, resulting into a dense and
homogenously distributed structure of trade shares. But it could be true the configuration in which some
members concentrate their trade on one or few of others members, giving as outcome a sparse and
heterogeneous structure of trade shares.

The *attractiveness* of countries as foreign “demand” or “supply” is a key factor to consider when
preferentiality is under inquiry. It is not unlikely that preferential trade pattern is affected by trade *position*
of particular countries. Indeed, it’s not a novelty that world trade network is characterized by few leaders,
with prominent position in terms of trade and economic size; they trade relatively more in absolute term
and have relatively more trade partners; they have a relatively bigger intermediate and final demand and a
well provided and competitive sources of supply, depending on the industry. Therefore, preferential trade
could be configured again as a sub-network characterized by the presence of some members that
concentrate their trade on one or few *attractive* ones (because they are leaders), but in turn it could be that
those attractive members address their preference out of the sub-network, to the rest of the
world/network.
Method
In order to detect sub-network we have employed the approach to community analysis proposed by Landi and Piccardi (2014), which has the advantage to take into account the directionality of links when searching for structures with the property of having comparatively stronger internal connectivity. Indeed the latter could refer to the links related to exports or imports only or to both categories. Landi and Piccardi (2014) work introduces the notions of in-, out-, and in-/out-community in order to correctly classify the directedness of the interaction of a sub-network with the rest of the system. Furthermore, they extend the scope of community analysis by introducing the notions of in-, out-, and in-/out-pseudo-community. They are sub-networks having strong internal connectivity but also important interactions with the rest of the system, the latter taking place by means of a minority of its nodes. The various types of (pseudo-)communities are qualified and distinguished by a suitable set of indicators and, on a given network, they can be discovered by using a “local” searching algorithm based on information processing related to a random walker.

The standard model is the following: at each (discrete) time step, the walker which is in node $i$ randomly selects one of the out-links $i \rightarrow j$ with a probability proportional to the link weight and, accordingly, follows the selected out-link to reach the neighbor $j$. The induced notion of community is that of a sub-network with a large escape time, i.e., such that the walker at each step has a large probability of remaining within the sub-network. If the network is directed, however, this approach introduces a bias, in that such a community has surely weak out-links towards the rest of the network, but nothing excludes that it has strong in-links too, i.e., links with large weights pointing from the outside towards the community. For example, think about the network of trades among countries, where the weight of $i \rightarrow j$ is the value of the export from country $i$ to country $j$. According to the above standard notion, the countries of a strongly significant community would export much more within the community rather than outside. But this does not exclude that they might have large import flows from the outside, which is in contrast with the idea of a community as a set weakly connected to the rest of the network. This calls for the new, distinct notions of out- and in-community: the former is a sub-network whose nodes direct most of their out-strength to within the community rather than to the rest of the network, whereas the latter is such that nodes receive most of their in-strength from within the community rather than from the outside. A sub-network with both features will be denoted as in/out-community.

The other new notion introduced by this paper is that of pseudo-community. Several case studies reveal the existence of peculiar structures, namely “star-like” sub-networks in which most of the nodes direct most of their out-strength within the sub-network (often towards a single “central” node), but a few of them (typically the “central” node only) mainly direct their out-strength to the outside. Examples are found in the worldwide air transportation network, where regional “hubs” collect all the traffic originating from domestic
airports and forward it to the rest of the world. In a sub-network with such features, a random walker has a small escape time due the large out-strength from the “central” node to the outside: the sub-network cannot therefore be qualified as a (out-)community. Yet, such as a structure is worth to be revealed and classified, since it has a special form of strong intra-connectivity: we will denote it as *out-pseudo-community*. Dually, an *in-pseudo-community* will be a sub-network where most of the nodes receive most of their in-strength from the community rather than from the rest of the network, but a few nodes have instead a large in-strength from the outside. A sub-network with both features will be denoted as *in/out-pseudo-community*.

Two indicators will be used to quantify the out-properties of a sub-network $S$, namely the *persistence probability* $\alpha_S$ and the *average internal strength* $\beta_S$. Their values will be used to possibly classify $S$ as out-community or out-pseudo-community. Dually, another two indicators $\alpha'_S$ and $\beta'_S$ will be used to quantify the in-properties of $S$, and to possibly classify it as in-community or in-pseudo-community. If we combine the in-and out-features, namely we consider the value of the 4-tuple $(\alpha_S, \beta_S, \alpha'_S, \beta'_S)$, we discover that a directed network can contain eight different non-trivial types of structures, i.e., sub-networks with peculiar properties: *out-community; in-community; in/out-community; out-pseudo-community; in/pseudo-community; in/out-pseudo-community; in-community/out-pseudo-community*. Each one of them corresponds to a specific combination of in-/out- as well as in-/inter-connectivity.

Consider a directed, weighted network with nodes $N = \{1, 2, ..., n\}$ and weight matrix $W = [w_{ij}]$, i.e. $w_{ij} > 0$ denotes the weight of the link of $i \rightarrow j$, which is set to 1 when the network is binary (i.e. unweighted), while $w_{ij} = 0$ if the link $i \rightarrow j$ does not exist. Denote $s_{i}^{\text{out}} = \sum_j w_{ij}$ and $s_{i}^{\text{in}} = \sum_j w_{ji}$, which reduce to *out- and in-degree* when the network is binary.

An $n$–state discrete time Markov chain can be associated to the network in a standard fashion. For that we denote by $p_{ij} = \frac{w_{ij}}{s_{i}^{\text{out}}}$ the probability that, at each step a random walker which is in node $i$ jumps to $j$, so that the probability $\pi_{i,t}$ of finding the walker in node $i$ at time $t$ is governed by $\pi_{t+1} = \pi_t P$ with $\pi_t = (\pi_{1,t}, \pi_{2,t}, ..., \pi_{n,t})$ and $P = [p_{ij}]$. Assuming that $P$ is an irreducible matrix, the stationary probability distribution $\pi = \pi P$ is unique and strictly positive $\pi_i > 0$ for all $i$.

Let us now denote by $S$ the sub-network formed by a subset $N_S \subset N$ of nodes of the original network, and by all the links of the latter connecting pairs of nodes of $S$. Sub-networks are candidates to be (pseudo-)communities depending on the value assumed by *persistence probabilities* $\alpha_S$, $\alpha'_S$ and the *average internal strengths* $\beta_S, \beta'_S$.
The persistence probability $\alpha_S$ is the probability that a random walker which is in any of the nodes of $S$ at time $t$, remains in $S$ at time $t + 1$. If we assume that Markov chain $\pi_{t+1} = \pi_t P$ is in the stationary state $\pi$, then we have:

$$\alpha_S = \sum_{i \in N_S} \frac{\pi_i}{\Pi_S} \sum_{j \in N_S} p_{ij} = \sum_{i \in N_S} \frac{\pi_i}{\Pi_S} \sum_{j \in N_S} \frac{w_{ij}}{s_{out}^i}$$

where $\Pi_S = \sum_{i \in N_S} \pi_i$ is the aggregate stationary probability of the sub-network $S$ and $\alpha_S$ goes from 0 to 1.

The persistence probability $\alpha_S$ is a measure of cohesiveness of $S$ (indeed, the expected escape time from $S$ is $1 - \alpha_S$) and it proves to be an effective tool for the structural analysis of networks. As it is apparent from (1), $S$ is a convex combination of the fraction of the out-strength of the nodes of $S$ that is directed within $S$: the coefficient of the term $i$ is the (normalized) stationary probability $\pi_i$, a well-known measure of centrality of node. Notice that the exact value of $S$ cannot be computed without knowing $\pi$, hence the entire network.

So sub-networks qualified as out-communities will be characterized by large values of $\alpha_S$. Substituting $p_{ij} = \frac{w_{ij}}{s_{out}^i}$ with $p'_{ij} = \frac{w_{ij}}{s_{in}^i}$ we define the corresponding in-indicator $\alpha'_S$ which allows us to characterize in-communities with the combination of the fraction of the in-strength of the nodes of $S$ that is directed within $S$.

Measuring the persistence probability alone may fail in revealing some interesting structures in the network: nodes could be intensively connected to nodes which in turn are more intensively connected to the rest of the network. Due to the large probability of escaping from $S$ by the latter, $\alpha_S$ is likely to be small. Yet, the structure appears to be definitely interesting and worth to be revealed. We try to capture it by associating a second indicator to $S$, called average internal strength $\beta_S$

$$\beta_S = \frac{1}{n_S} \sum_{ij \in N_S} p_{ij} = \sum_{i \in N_S} \frac{1}{n_S} \sum_{j \in N_S} \frac{w_{ij}}{s_{in}^i}$$

where $n_S$ is the number of nodes of $S$. The quantity $0 \leq \beta_S \leq 1$ is simply the arithmetic mean, over the nodes of $S$, of the fraction of the out-strength directed internally to $S$ (we recall that $\alpha_S$ is a weighted mean of the same quantities). Thus $\beta_S$ will be large when most of the nodes of $S$ direct most of their out-strength within $S$, although a few others could do the opposite yielding a small $\alpha_S$. The latter case describes well the behavior of an out-pseudo-community.

As before substituting $p_{ij} = \frac{w_{ij}}{s_{out}^i}$ with $p'_{ij} = \frac{w_{ij}}{s_{in}^i}$ we define the corresponding in-indicator $\beta'_S$ that will be large when most of the nodes of $S$ direct most of their in-strength within $S$.

By combining these four indicator is now possible to define eight type of structures as much are different plausible combination of values assumed by the 4-upla ($\alpha_S$, $\beta_S$, $\alpha'_S$, $\beta'_S$). We define out-community that
structure with the 4-upla \((\alpha_S, \beta_S, \alpha'_S, \beta'_S)\) tends to assume values \((1,1,-0)\), in-community when it tends to be \((-0,1,1)\), in-out-community when it tends to be \((1,1,1)\), out-pseudo-community when it tends to be \((0,1,-0)\), in-pseudo-community when it tends to be \((-0,1,1)\), in-out-pseudo-community when it tends to be \((0,1,0,1)\), in-community/out-pseudo-community when it tends to be \((0,1,1,1)\), and out-pseudo-community/in-community when it tends to be \((1,1,0,1)\).

The point now is to fix a quality threshold \(\varepsilon > 0\) to values of indicators in order to analyze and classify a concrete sub-networks. Fixing it at 0,5 it means for example that significant out-communities will be that one with the 4-upla \((\alpha_S, \beta_S, \alpha'_S, \beta'_S)\) assuming values \((\alpha_S \geq 0,5, \beta_S \geq 0,5, -, \beta'_S \leq 0,5)\) and so on. In economic terms it means that significant out-communities will be those within which countries export with a probability higher than 0.5 \((\alpha_S \geq 0,5)\), if \(\alpha_S \geq 0,5\) it necessarily holds that on average they trade within the structure a friction \(\beta_S \geq 0,5\) but it is not avoided the case they trade, on average, a large friction of their imports by countries that are outside the structure \((\beta'_S \leq 0,5)\). \(\alpha'_S\) in this case is redundant and not considered to characterize the structure.

Table 1 would be a reference schema where type of structure and characterizing indicators are connected by assigning at each of the latter a threshold of significance equal to 0.5. Indeed it is not implausible to assume that preferences in trading are signaled by (weighted or arithmetic) a mean of trade shares realized within each structures \(S\), higher than 50%.

| Type of structures               | \(\alpha_S\) | \(\beta_S\) | \(\alpha'_S\) | \(\beta'_S\) |
|---------------------------------|--------------|--------------|---------------|--------------|
| in-out-communities              | \(\geq0,5\)  | \(\geq0,5\)  | \(\geq0,5\)   | \(\geq0,5\)  |
| in-comm./out-ps-comm.           | \(\leq0,5\)  | \(\geq0,5\)  | \(\geq0,5\)   | \(\geq0,5\)  |
| in-communities                  | \(-\)        | \(\leq0,5\)  | \(\geq0,5\)   | \(\geq0,5\)  |
| in-ps-communities               | \(-\)        | \(\leq0,5\)  | \(\geq0,5\)   | \(\geq0,5\)  |
| in-ps-comm./out-comm.           | \(\geq0,5\)  | \(\geq0,5\)  | \(\leq0,5\)   | \(\geq0,5\)  |
| in/out-ps-communities           | \(\leq0,5\)  | \(\geq0,5\)  | \(\leq0,5\)   | \(\geq0,5\)  |
| out-ps-communities              | \(\leq0,5\)  | \(\geq0,5\)  | \(-\)          | \(\leq0,5\)  |
| out-communities                 | \(\geq0,5\)  | \(\geq0,5\)  | \(-\)          | \(\leq0,5\)  |

Table 1 Constraints applied to indicators characterizing different type structures in order to classify them as significant

Having defined the types of structures we are interested in, we need an algorithm to discover them in a given network. They use a local approach, similar in spirit to a few recent proposals (Bagrow and Boltt, 2005; Clauset, 2005; Piccardi, 2011). Differently from many community analysis methods, this algorithm
does not yield a partition of the network, i.e., a node might be not included in any structure. This seems perfectly reasonable, however, as it is not uncommon to discover strongly connected groups of nodes even in networks which, overall, do not possess a definite clusterized structure. Forcing a partition in such networks places side by side high and low quality clusters, often without the capability of discriminating among them. On the other hand, the above algorithm highlights significant structures only, and fully allows for overlapping, as one node may belong to more than one (pseudo)community. But there is also a “second level” of over-lapping, remarkably, since a node could be at the same time part of structures of different types, e.g., an out-community and an in-pseudo-community, sharing these memberships with different sets of partners.

4. Data analysis: sectoral preferential patterns

Data
For the purpose of this work data on three sectoral specific trade among 222 countries were collected from CEPII-BACI database. In particular Textiles, Electronics and Motor Vehicles trade data were analyzed for the pre-crisis year 2006. BACI database implements HS classification to describe sectoral trade between pairs of countries. The first corresponds to section XI, the second and third are subsection respectively of the section XVI and XVII of HS 2002 classification. Measures of trade do not distinguish trade in finished products from trade in intermediate inputs. Using CEPII-BACI data we counted that in 2006 the trade among all 222 countries was about 543 millions for Textiles and 1.6 and 1 billions of dollars respectively for Electronics and Motor Vehicles sectors.

Discussion of results
Applying the local searching algorithm to sectoral trade data we have found evidence of existence of significant sub-networks in which preferential trade occurs. In particular, we have found that structures such as in-, out-, and in/out-pseudo-community dominate the pattern in which preferential trade is organized, independently from the sector we are considering. But also significant in-communities, out-communities and in-communities/out-pseudo-communities have been found. Generally, pseudo-communities are much numerous, more significant and bigger. Table 2 shows number and size statistics about significant structures we have found.
Table 2 – Number and sizes of significant structures found analyzing sectoral trade networks.

|                      | Textiles          |                      | Motor vehicles      |                      | Electronics        |
|----------------------|-------------------|----------------------|---------------------|---------------------|--------------------|
|                      | Number | Size (min-max) | Distribution of size | Number | Size (min-max) | Distribution of size | Number | Size (min-max) | Distribution of size |
| in-communities       | 1       | 6 (2*)         | 2-60                | 2       | 26-44          | 3-100               | 3       | 26-44          | 3-100               |
| out-communities      | 16      | 8 (8*)         | 5-50                | 5       | 14-17          | 3-50                | 2       | 14-17          | 3-50                |
| in-ps-communities    | 44      | 15*           | 2-27                | 120     | 15*           | 4-43                | 47      | 10*           | 4-45                |
| out-ps-communities   | 119     | 39*           | 2-39                | 85      | 38*           | 2-54                | 84      | 23*           | 3-41                |
| in-out-ps-communities| 38      | 16*           | 9-35                | 20      | 9*            | 3-31                | 35      | 7*            | 15-27               |
| in-comm./out-ps-comm | 5       | 4*            | 17-25               |         |                |                     | 5       | 4*            | 17-25               |

*Reduced number of structures or sub-networks. In this case overlapping sub-networks were merged when similarity among them resulted higher than a threshold fixed at 0.5. Given two sub-networks defined as two sets of node pairs, similarity between sub-networks has been measured as the ratio between the intersection of the two sets of nodes pairs and their union.

Figure 1 illustrates how these structures can appear in reality in sectoral trade networks we have analyzed (with simplifications). Here are reported only types of structures found analyzing data. It is not rare the case where there are more than one node relatively more central that the others, receiving relatively more preferences as destination market or source of supply. But asymmetries in position of nodes were spotted across all the significant structures we have found revealing that preference on trade is driven from some peculiarities embodied by the few countries chosen by many of the others as main trade partner.

To have a measure of the relevance of these structures we have to confront the value of trade occurring within these structures (not involving all countries) and the sectoral total value exchanged at world level. Indeed preferentiality in choosing trade partners affect a good proportion of sectoral world trade and, more importantly, for all those countries that are directly involved into these structures, preferential trade constitutes the majority of their trade.
Figure 1 Simplified topologies of significant structures found in the three sectoral networks analyzed

Table 3 illustrates this fact when the all significant structures of a particular type are taken simultaneously, with most of them overlapped, as considering trade occurring within meta-structures. This gives us a measure of the extension of the phenomenon of preferential trade, without saying anything about actual behavior of each type of structure.

| Textiles as share* of world trade | Textiles as share* of their trade | Motor Vehicles as share* of world trade | Motor Vehicles as share* of their trade | Electronics as share* of world trade | Electronics as share* of their trade |
|----------------------------------|----------------------------------|----------------------------------------|----------------------------------------|-------------------------------------|-------------------------------------|
| in-communities                   | ----                             | ----                                   | ----                                   | 26.11%                             | 70.80%                             |
| out-communities                  | ----                             | 0.69%                                  | 71.21%                                 | 0.45%                               | 50.97%                             |
| in-ps-communities                | 70.04%                           | 60.13%                                 | 63.21%                                 | 49.88%                             | 65.44%                             |
| out-ps-communities               | 30.48%                           | 52.64%                                 | 53.63%                                 | 59.61%                             | 61.82%                             |
| in/out-ps-communities            | 49.94%                           | 56.48%                                 | 58.21%                                 | 34.75%                             | 40.96%                             |
| in-comm./out-ps-comm.            | 59.42%                           | 62.38%                                 | ----                                   | ----                               | ----                               |

*structures are considered simultaneously (with overlapping)

Table 3 Trade occurring within meta-structures, combination of all found significant structures, by type.
Evidence on average behavior of each type of structure by sector can be seen from Table 4, where structures are analyzed separately to see how much are relevant respect to world sectoral overall trade and how much trade occur on average within each of them respect to all their trade. It is interesting to see that pseudo-communities, across all sectors, even if the most numerous, significant and bigger they trade on average not a big share of world sectoral trade, if confronted with in-communities/out-pseudo-communities in Textiles or in-communities in Electronics. Electronics is also characterized by more diversified preferential trade structures, with larger shares occurring within them respect the relative world trade.

|                  | Textiles | Motor Vehicles | Electronics |
|------------------|----------|----------------|-------------|
|                  | as share** of world trade | as share** of their trade | as share** of world trade | as share** of their trade | as share** of world trade | as share** of their trade |
| in-communities   | ----     | ----           | ----        | 22.09% | 67.84% |
| out-communities  | ----     | ----           | 0.09%      | 56.90% | 0.27%  | 51.78% |
| in-ps-communities| 8.85%   | 27.66%         | 5.10%      | 29.64% | 10.53% | 29.65% |
| out-ps-communities| 4.70% | 42.93%         | 2.50%      | 28.02% | 6.77%  | 32.93% |
| in/out-ps-communities | 10.42% | 30.50%        | 10.76%     | 34.31% | 13.32% | 35.35% |
| in-comm./out-ps-comm. | 25.02% | 51.68%        | ----       | ----   | ----   | ----   |

** structures are considered singularly (without overlapping): average value of relative trade shares

Table 4 Trade occurring within structures, by type.

To provide more insights on the phenomenon of preferential trade it is worth to see how is distributed the trade occurring inside the structures we have found if we cluster countries into two groups defined by level of GDP per-capita. We will call H those countries with a level of GDP higher than 11.115 (threshold suggested by World Bank to classify 2006 GDP per-capita levels as high), instead we indicate with L the rest of countries.
Table 5  Distribution of preferential trade by income level of participating countries. HL for example stands for H countries address preference to L countries.

As you can see from Table 5 there is mixed evidence on composition of preferential trade in terms of nature of relationships among participating countries. It seems depending mainly on sector the structures are referred. Only out-communities seem different from the others: they include only countries of L type, with a relative lower income per capita, that trade each other in an intensively as main destination markets of their finished products or intermediate inputs, but they are connected more intensively with the rest of the system by imports. Adding to this fact it emerges that in Textile the preferential pattern is leading by L addressing their preferences to H, across all the structure, it means, both when they import from or export to them. Instead in Motor vehicles sector it prevails the pattern in which H countries address their preferences to H. Finally in Electronics it seems that exists a more balanced pattern of preferences given to H countries by L and H.
Many observations can be done on this evidence: firstly, regarding that share of intra-industry trade that occur along preferential channels, not necessarily it concerns countries of the same income level, but a more complex pattern emerges, where a not irrelevant friction of this trade is fulfilled by countries of different level of income. Secondly, this evidence suggests the existence of some relation between distribution of trade by income level of participating countries and the nature of goods they trade: the statistics we use include both intermediates and finished products. It is plausible to think that trade among L and H is different in nature from the trade among H and L, but also from that occurring within H group or L group.

Another important matter is to understand how much these structures are affected by geographical distance, if there are differences across type of structure and across sectors. Below (Figure 2) you find a box plot of distribution of relative geographical distance of preferential partners weighted by share of trade (export or import depending on type of structure), by sector and by structures.

What you can see instantaneously is that preferentiality in choosing trade partners as main destination of exports is more affected by distance than when choosing sources of import. Indeed in the former case the preferences are addressed systematically to relatively closer partners in respect to the latter. Furthermore it seems that distance play a weaker or stronger role on preferences depending on sector we consider, but there is a mixed evidence on which sector is more affected by distance, depending on type of structure we are considering.
We can conclude this paragraph adding one more observation: this evidence underlines the importance to make distinction between export side and import side of trade decisions of countries when searching for the strength of trade frictions like geographical distance: indeed trade frictions act heterogeneously on volume of trade a country import from or export to a partner, depending not only from the characteristics of partners but more importantly, from the nature of traded goods. i.e. import vs export and finished goods vs intermediate inputs. Unfortunately we do not have the possibility to breakdown trade in its main component of intermediate and finished products to add substance to this discussion. But let us now do an example to clarify the importance of considering nature of trade in relation to distance frictions evaluation.

Consider in-communities in Electronics which represent 22.11% of world trade in that sector and 70.80% of their sectoral total trade (Table 3). In this structure trade we labeled as HH counts 12.19%, LH 30.29%, HL 31.63% and finally LL 25.89% (See Table 5). Relative geographical distance of preferential partners weighted by share of trade varies from 8 to 13 (given that the country most distant equal to 100). In the same sector there are also significant in-pseudo communities accounting for 49.88% and 65.44% respectively of world trade in this sector and of their sectoral total trade. In this structure HH trade counts 35.96%, LH 33.74% HL 17.84% and finally LL 12.46%. Relative geographical distance of preferential partners weighted by share of trade varies from 27 to 35. These structures differ in terms of behavior with the rest of the world system: in-communities are characterized by the presence of countries that import preferentially by some members of structures that in turn import preferentially from some of them but export preferentially with the rest of the system. In-pseudo-communities are characterized by the presence of countries that import preferentially by some members of their structures that in turn import preferentially from and export to the rest of the system. Now, suppose that the in-communities are mainly composed by countries that trade each other intermediate inputs to be transformed and then exported to the rest of the world. Instead in-pseudo-communities are characterized by the presence of small countries that import mainly finished products from bigger countries which mainly import from elsewhere intermediates to export again to the system as finished products. For the former the distance is more important because the trade links are more related to production process and faster communications are essential for not losing competitiveness in terms of production lead times. For the latter instead, the distance is relatively less important than the price competitiveness and/or product variety a small country reach importing from bigger countries, leaders of these sectors.

**Connecting trade preference to partners characteristics**

*Dependence, reciprocity and triangulation (to be completed)*

This paragraph would treat about correlations among trade shares of each pair of countries within the structures, revealing the importance of trade *dependence* when a country chooses a partner for its import or
exports, but also the fact that chosen partners tend to be those one that have relation mainly with the rest of the system, featuring no reciprocity in preferences (with the exception of out-communities). Also triangulation does not seem to be a characteristic of preferential trade relations: country does not trade more with countries that share a larger number of partners in trade. All this elements reveal the fact that preferential trade is essentially asymmetrical and with scarce cohesiveness.

_Triangulation:_ the partner where a country exports or import is connected to common partners from where the county imports (sources) or export. Is it part of a chain where both the country and the partner take part?
Table 6

*Other partner attributes from network positioning: in/out degree and in-out strength, in-eigen, out-eigen, authority, hubness centrality (to be completed)*

This paragraph would be a last effort in this work in order to account for the positioning of partners in the trade network (centrality in our case) when searching for drivers of trade preferences. More a country is central in the entire network, more is preferred by countries.
### Table 7

#### Textiles

| Second Order Attributes | Higher Order Attributes |
|-------------------------|-------------------------|
| gdp                     | in-degree               | out-degree | in-strength | out-strength | in-eigen-centrality | out-eigen-centrality | authority | hubness |
| In-communities          | ---                     | ---        | ---         | ---          | ---               | ---                   | ---        | ---     |
| Out-ps-communities      | 0.452                   | 0.604      | 0.533       | 0.555        | 0.540             | 0.494                 | 0.589      | 0.608   |
| In/out-ps-communities (import) | 0.438                   | 0.646      | 0.685       | 0.552        | 0.642             | 0.653                 | 0.639      | 0.612   |
| In/out-ps-communities (export) | 0.451                   | 0.615      | 0.547       | 0.571        | 0.551             | 0.511                 | 0.608      | 0.582   |
| In-comm./out-ps-comm. (import) | 0.448                   | 0.666      | 0.726       | 0.571        | 0.683             | 0.718                 | 0.663      | 0.643   |
| In-comm./out-ps-comm. (export) | 0.503                   | 0.664      | 0.601       | 0.594        | 0.598             | 0.578                 | 0.663      | 0.645   |

#### Motor Vehicles

| Second Order Attributes | Higher Order Attributes |
|-------------------------|-------------------------|
| gdp                     | in-degree               | out-degree | in-strength | out-strength | in-eigen-centrality | out-eigen-centrality | authority | hubness |
| In-communities          | ---                     | ---        | ---         | ---          | ---               | ---                   | ---        | ---     |
| Out-communities         | 0.084                   | 0.058      | 0.017       | 0.089        | 0.014             | 0.012                 | 0.059      | 0.062   |
| In-ps-communities       | 0.420                   | 0.658      | 0.667       | 0.601        | 0.624             | 0.652                 | 0.647      | 0.645   |
| Out-ps-communities      | 0.389                   | 0.517      | 0.463       | 0.517        | 0.489             | 0.434                 | 0.501      | 0.501   |
| In/out-ps-communities (import) | 0.431                   | 0.673      | 0.690       | 0.610        | 0.650             | 0.670                 | 0.656      | 0.650   |
| In/out-ps-communities (export) | 0.437                   | 0.596      | 0.550       | 0.584        | 0.571             | 0.513                 | 0.576      | 0.572   |
| In-comm./out-ps-comm.   | ---                     | ---        | ---         | ---          | ---               | ---                   | ---        | ---     |

#### Electronics

| Second Order Attributes | Higher Order Attributes |
|-------------------------|-------------------------|
| gdp                     | in-degree               | out-degree | in-strength | out-strength | in-eigen-centrality | out-eigen-centrality | authority | hubness |
| In-communities          | 0.539                   | 0.765      | 0.809       | 0.641        | 0.746             | 0.809                 | 0.758      | 0.752   |
| Out-communities         | 0.214                   | 0.390      | 0.317       | 0.356        | 0.300             | 0.312                 | 0.406      | 0.397   |
| In-ps-communities       | 0.529                   | 0.770      | 0.802       | 0.670        | 0.758             | 0.793                 | 0.767      | 0.753   |
| Out-ps-communities      | 0.488                   | 0.668      | 0.618       | 0.629        | 0.620             | 0.594                 | 0.658      | 0.656   |
| In/out-ps-communities (import) | 0.509                   | 0.755      | 0.787       | 0.652        | 0.738             | 0.781                 | 0.757      | 0.740   |
| In/out-ps-communities (export) | 0.460                   | 0.658      | 0.602       | 0.625        | 0.609             | 0.586                 | 0.663      | 0.655   |
| In-comm./out-ps-comm.   | ---                     | ---        | ---         | ---          | ---               | ---                   | ---        | ---     |
5. Conclusions
We have applied a local search algorithm to three sectoral trade networks in order to reveal endogenous clusters of countries with the propensity of concentrating their trade in few partners, taking into account the directedness of trade links.

We have found that preferential trade is not an irrelevant part of total trade, i.e. countries generally prefer to concentrate their trade (exports or imports) in one or very few partners, often the most important economies of the world that are also the more connected to the system of trade relations.

This fact has as outcome that trade is mainly organized in asymmetrical, not reciprocal and very sparse structures, featuring few leading countries to which many other countries are dependent in exporting and importing goods. There is evidence in support of prevalence of structures with low cohesiveness and high centralization in any of the sector analyzed.

Sector characteristics determine differences in terms of which countries are the source of preferentiality and which other are the destinations of such preferentiality. Patterns of preferential trade varies with sectors, confirming the importance to analyze disaggregate data.

We have also found that distance is less important as trade friction when countries import rather than when they export preferentially but it could be also true that it may vary its strength when considering also the nature of the traded good (intermediate inputs or finished products).

Finally the position of trade partners on trade web in terms of their centrality, even if highly correlated with their economic dimension, it could be considered another strong driving force in shaping the pattern of trade preferences both when they export and import: all the dimensions of centrality analyzed are positively correlated with preferences, more than with the pure GDP.
References

Bagrow, James P., and Erik M. Bollt. "Local method for detecting communities." *Physical Review E* 72, no. 4 (2005): 046108.

Baldwin, Richard, and Javier Lopez-Gonzalez. "Supply-chain Trade: A Portrait of Global Patterns and Several Testable Hypotheses." *The World Economy* (2014).

Bhattacharya, Kunal, Gautam Mukherjee, Jari Saramäki, Kimmo Kaski, and Subhrangshu Sekhar Manna. "The international trade network: weighted network analysis and modelling." *Journal of Statistical Mechanics: Theory and Experiment* 2008, no. 02 (2008): P02002.

Clauset, Aaron. "Finding local community structure in networks." *Physical review E* 72, no. 2 (2005): 026132.

De Backer, Koen, and Sebastien Miroudot. *Mapping Global Value Chains*. No. 159. OECD Publishing, 2013.

De Benedictis, Luca, and Lucia Tajoli. "The world trade network." *The World Economy* 34, no. 8 (2011): 1417-1454.

Della Rossa, Fabio, Fabio Dercole, and Carlo Piccardi. "Profiling core-periphery network structure by random walkers." *Scientific reports* 3 (2013).

Ermann, Leonardo, and Dima L. Shepelyansky. "Ecological analysis of world trade." *Physics Letters A* 377, no. 3 (2013): 250-256.

Fagiolo, Giorgio, Javier Reyes, and Stefano Schiavo. "On the topological properties of the world trade web: A weighted network analysis." *Physica A: Statistical Mechanics and its Applications* 387, no. 15 (2008): 3868-3873.

Lejour, Arjan, Hugo Rojas-Romagosa, and Paul Veenendaal. "Identifying hubs and spokes in global supply chains with redirected trade in value added."

Miroudot, Sébastien, and Koen De Backer. *Mapping Global Value Chains*. Working Paper (forthcoming), OECD, Paris, 2013.

Hanson, Gordon H. "The Rise of Middle Kingdoms: Emerging Economies in Global Trade." *Journal of Economic Perspectives* 26, no. 2 (2012): 41-64.

He, Jiankui, and Michael W. Deem. "Structure and response in the world trade network." *Physical review letters* 105, no. 19 (2010): 198701.

Head, Keith, and Thierry Mayer. "What separates us? Sources of resistance to globalization." *Canadian Journal of Economics/Revue canadienne d'économique* 46, no. 4 (2013): 1196-1231.

Nicita, Alessandro, and Marcelo Olarreaga. "Trade, production, and protection database, 1976–2004." *The World Bank Economic Review* 21, no. 1 (2007): 165-171.
Landi, Pietro, and Carlo Piccardi. "Community analysis in directed networks: In-, out-, and pseudocommunities." *Physical Review E* 89, no. 1 (2014): 012814.

Li, Xiang, Yu Ying Jin, and Guanrong Chen. "Complexity and synchronization of the world trade web." *Physica A: Statistical Mechanics and its Applications* 328, no. 1 (2003): 287-296.

Piccardi, Carlo. "Finding and testing network communities by lumped Markov chains." *PloS one* 6, no. 11 (2011): e27028.

Piccardi, Carlo, and Lucia Tajoli. "Existence and significance of communities in the world trade web." *Physical Review E* 85, no. 6 (2012): 066119.

Piccardi, Carlo, and Lucia Tajoli. "The Relevance and Significance of Preferential Trade in the World Trade Network." (2012).

Fagiolo, Giorgio, Javier Reyes, and Stefano Schiavo. "World-trade web: Topological properties, dynamics, and evolution." *Physical Review E* 79, no. 3 (2009): 036115.

Reichardt, Jörg, and Douglas R. White. "Role models for complex networks." *The European Physical Journal B-Condensed Matter and Complex Systems* 60, no. 2 (2007): 217-224.

Serrano, Ma Angeles, and Marián Boguñá. "Topology of the world trade web." *Physical Review E* 68, no. 1 (2003): 015101.

Squartini, Tiziano, Giorgio Fagiolo, and Diego Garlaschelli. "Randomizing world trade. II. A weighted network analysis." *Physical Review E* 84, no. 4 (2011): 046118.

Tzekina, Irena, Karan Danthi, and Daniel N. Rockmore. "Evolution of community structure in the world trade web." *The European Physical Journal B-Condensed Matter and Complex Systems* 63, no. 4 (2008): 541-545.

World Trade Organization, "WTO Report 2008- Trading in globalize world", (2008)
Appendix

Description of sectors composition by HS 2002 classification

SECTION XI

TEXTILES AND TEXTILE ARTICLES

50 Silk.

51 Wool, fine or coarse animal hair; horsehair yarn and woven fabric.

52 Cotton,

53 Other vegetable textile fibres; paper yarn and woven fabrics of paper yarn.

54 Man-made filaments.

55 Man-made staple fibres.

56 Wadding, felt and nonwovens; special yarns; twine, cordage, ropes and cables and articles thereof.

57 Carpets and other textile floor coverings.

58 Special woven fabrics; tufted textile fabrics; lace; tapestries; trimmings; embroidery.

59 Impregnated, coated, covered or laminated textile fabrics; textile articles of a kind suitable for industrial use.

60 Knitted or crocheted fabrics.

61 Articles of apparel and clothing accessories, knitted or crocheted.

62 Articles of apparel and clothing accessories, not knitted or crocheted.

63 Other made up textile articles; sets; worn clothing and worn textile articles; rags.

SECTION XVII

VEHICLES, AIRCRAFT, VESSELS AND ASSOCIATED TRANSPORT EQUIPMENT

[...]

87 Vehicles other than railway or tramway rolling-stock, and parts and accessories thereof.

[...]

SECTION XVI

MACHINERY AND MECHANICAL APPLIANCES; ELECTRICAL EQUIPMENT; PARTS THEREOF; SOUND RECORDERS AND REPRODUCERS, TELEVISION IMAGE AND SOUND RECORDERS AND REPRODUCERS, AND PARTS AND ACCESSORIES OF SUCH ARTICLES

[...]

85 Electrical machinery and equipment and parts thereof; sound recorders and reproducers, television image and sound recorders and reproducers, and parts and accessories of such articles,

[...]
